LOBSTER: Local Orbital Projections, Atomic Charges, and Chemical Bonding Analysis from Projector-Augmented-Wave-Based DFT

Ryky Nelson,[1] Christina Ertural,[1] Janine George,[2] Volker L. Deringer,[3] Geoffroy Hautier,[2] and Richard Dronskowski[1,4,5]*

Correspondence to: Richard Dronskowski (E-mail: *drons@HAL9000.ac.rwth-aachen.de*)

[1] *Institute of Inorganic Chemistry, RWTH Aachen University, Landoltweg 1, 52056 Aachen, Germany*

[2] *Institute of Condensed Matter and Nanosciences, Université Catholique de Louvain, Chemin des étoiles 8, 1348 Louvain-la-Neuve, Belgium*

[3] *Department of Chemistry, Inorganic Chemistry Laboratory, University of Oxford, Oxford OX1 3QR, UK*

[4] *Jülich-Aachen Research Alliance, JARA-HPC, RWTH Aachen University, 52056 Aachen, Germany*

[5] *Hoffmann Institute of Advanced Materials, Shenzhen Polytechnic, 7098 Liuxian Blvd, Nanshan District, Shenzhen, China*

ABSTRACT

We present an update on recently developed methodology and functionality in the computer program LOBSTER (Local Orbital Basis Suite Towards Electronic-Structure Reconstruction) for chemical-bonding analysis in periodic systems. LOBSTER is based on an analytic projection from projector-augmented wave (PAW) density-functional theory (DFT) computations [*J. Comput. Chem.* **2013**, *34*, 2557], reconstructing chemical information in terms of local, auxiliary atomic orbitals and thereby opening the output of PAW-based DFT codes to chemical interpretation. We demonstrate how LOBSTER has been improved by taking into account time reversal symmetry, thereby speeding up the DFT and LOBSTER calculations by a factor of 2. Over the recent years, the functionalities have also been continually expanded, including accurate projected densities of states (DOS), crystal orbital Hamilton population (COHP) analysis, atomic and orbital charges, gross populations, and the recently introduced k-dependent COHP. The software is offered free-of-charge for non-commercial research.

Introduction

The notion of a "chemical bond", specifically in the framework of molecular orbital (MO) theory, is ingrained in any chemist's education, and the simplest and most iconic representative of that is the H_2 molecule with its bonding (σ_g stabilizing) and antibonding (σ_u^* destabilizing) orbital interactions. The same conceptual approach is used, very widely and arguably successfully, in many areas of solid-state chemistry and materials science. Despite perpetual controversies surrounding the very idea of a chemical bond (and the trivial fact that a chemical bond is not a quantum-mechanical observable), the usefulness and predictive power of chemical-bonding models has led to substantial scientific progress, in particular, new chemical matter by rational syntheses. And despite the increasing role of automation and "big data" in

chemical research, one may argue that more, not less, of this insight is needed because human beings are committed to understanding.

In principle, bonding and antibonding orbital interactions are defined for solids just as well as for isolated molecules. Quantifying orbital overlap to understand bonding has been pioneered for solids in the 1980s, in Hughbanks' and Hoffmann's iconic Crystal Orbital Overlap Population (COOP) technique.[1] Related ideas were then pursued in the first-principles domain where weighting of orbital interactions by the associated Hamiltonian matrix element leads to the Crystal Orbital Hamilton Population (COHP).[2] The COHP in particular has been implemented in programs including the venerable TB-LMTO-ASA suite[3] and more recently in the CRYSTAL software.[4]

One thing in common that the aforementioned programs share is the *locality* of the basis sets em-

ployed for their calculations. However, many of to-day's DFT programs widely use delocalized plane-wave (PW) basis sets in combination with the pseudo-potential approach, which hinder the development of chemical bonding tools in such programs. Using a projection method which was similar to the method originally developed by Sánchez-Portal et al.,[5] and used early on in plane-wave calculations,[6] LOBSTER enables the chemical-bonding analysis from PW-based wavefunctions. It is worth mentioning that currently the projection method in LOBSTER has been designed to work with the projector-augmented wave (PAW) method,[11] which also marks the difference with the method developed in Ref. [5]. The reason for this is that the PAW method has been very widely implemented in many PW-based DFT programs due to its efficiency and accuracy, making a powerful advance over many previous plane-wave pseudopotential schemes. Among the many applications of LOBSTER in recent years, we may highlight but a few: it has been used to chemically understand molecules on nano-particle surfaces,[7] unconventional phases from crys-tal-structure prediction,[8] or the reactivity of zeo-lites[9] or oxide electrodes.[10] Furthermore, all the de-tails of the projection method and of the local basis sets implemented in LOBSTER can be found in our previous publications[12] and the User's Guide distrib-uted with the program for any technical questions. In the present work, we describe developments in LOB-STER that have been implemented and used over the recent years.
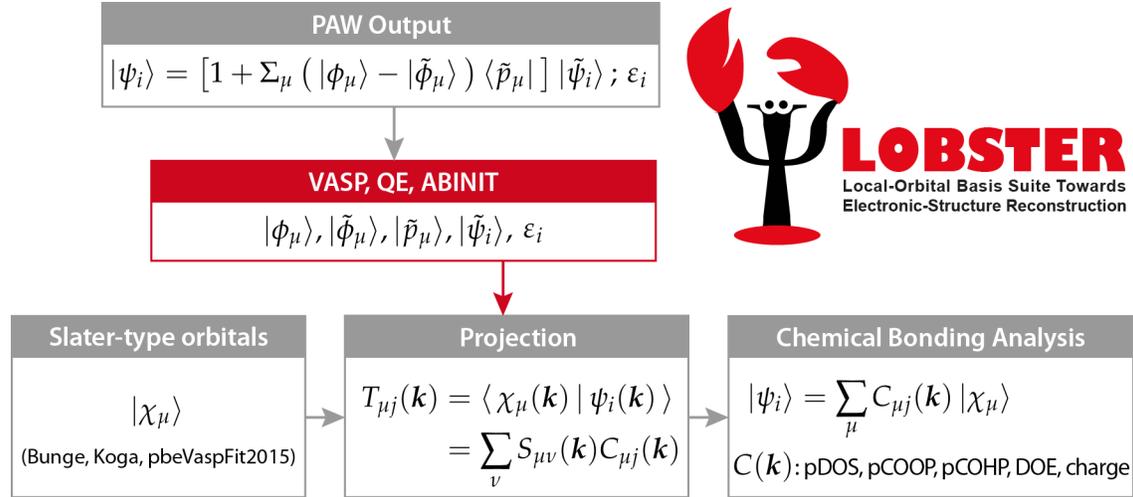
PAW Output

$$|\psi_i\rangle = \left[1 + \Sigma_\mu \left(|\phi_\mu\rangle - |\tilde{\phi}_\mu\rangle\right)\langle\tilde{p}_\mu|\right]|\tilde{\psi}_i\rangle \,;\, \varepsilon_i$$

VASP, QE, ABINIT

$$|\phi_\mu\rangle, |\tilde{\phi}_\mu\rangle, |\tilde{p}_\mu\rangle, |\tilde{\psi}_i\rangle,\, \varepsilon_i$$

LOBSTER
Local-Orbital Basis Suite Towards
Electronic-Structure Reconstruction

Slater-type orbitals

$$|\chi_\mu\rangle$$

(Bunge, Koga, pbeVaspFit2015)

Projection

$$T_{\mu j}(\boldsymbol{k}) = \langle\chi_\mu(\boldsymbol{k})\,|\,\psi_i(\boldsymbol{k})\rangle$$
$$= \sum_\nu S_{\mu\nu}(\boldsymbol{k})C_{\mu j}(\boldsymbol{k})$$

Chemical Bonding Analysis

$$|\psi_i\rangle = \sum_\mu C_{\mu j}(\boldsymbol{k})\,|\chi_\mu\rangle$$

$C(\boldsymbol{k})$: pDOS, pCOOP, pCOHP, DOE, charge

Figure 1. Overview of LOBSTER's data flowchart. It starts from one-electron (Bloch) wavefunctions $|\psi_i\rangle$ and the according eigenvalues ε_i of a quantum-chemical system of interest, brought to self-consistency using a DFT program which employs the projector-augmented wave (PAW) method. After wavefunction-related data are extracted through I/O interfaces (red block) specially tailored to three DFT programs (VASP, ABINIT, and QE), a basis set of contracted Slater-type orbitals $\left(|\chi_\mu\rangle\right)$ is semi-automatically selected from a built-in database. Wavefunctions and basis sets are then brought into the projection routine to determine the overlap matrix $S_{\mu\nu}$ and the transfer matrix $T_{\mu j}$ between the delocalized and localized representations. From those, the projected coefficient and Hamiltonian matrices $C_{\mu j}$ and $H_{\mu\nu}$, respectively, are accessible and enable various bond-analytic tools. The LOBSTER logo is copyrighted by the Chair of Solid-State and Quantum Chemistry at RWTH Aachen University.

Methods

The principal flow of data and procedures is illustrated in Fig. 1. LOBSTER's primary idea is to reconstruct the PAW wavefunctions by projecting them onto a local basis set, hence reconstructing the entire electronic structure in another representation. This very projection, furthermore, yields the coefficients of linear combinations of atomic orbitals (LCAO) that, together with the overlap and Hamiltonian matrices, allow the calculations of various chemical-bonding quantities.

Furthermore, although originally LOBSTER was designed with interfaces to handle only wavefunctions from VASP[13] and ABINIT,[14] another interface has been recently developed to process wavefunctions from another widely used and open-source DFT package, Quantum ESPRESSO (QE),[15] which has made LOBSTER more broadly accessible. Using QE and LOBSTER, the chemical bonding analysis of the electronic and magnetic properties of manganese carbodiimide (MnNCN) and manganese oxide (MnO) was firstly demonstrated and reported in Ref. [16].

LOBSTER is a multiplatform tool that is written in object-oriented C++ and parallelized using OpenMP. It employs Boost libraries[17] as well as the highly efficient Eigen library.[18] For further details of the technical aspects, readers may refer to our initial publications.[12]

Time reversal symmetry

As time progresses, (nanoscale) materials at the scientific front can become so complex that their projections, as well as their prior DFT calculations, turn out as being extraordinarily expensive to compute. Parallelism using OpenMP has been implemented in LOBSTER since version 1.1.0 to reduce some of the computational burden but more improvement in the efficiency is still needed. Trivially, the entire projection is computationally costly because it must be done for all $\boldsymbol{k}$-points in the first Brillouin-zone (BZ), which affects a) calculation time and b) storage-space. Fortunately, $\boldsymbol{k}$-space symmetry can be exploited.

Time-reversal (TR) symmetry is always present in systems isolated from external fields; mathematically, this symmetry conforms to an anti-unitary operator $\hat{T}$ whose operation on a given Bloch wavefunction yields the complex conjugate of the wavefunction.[19] However, it is also already well known that $\Psi_{j\boldsymbol{k}}^*(\boldsymbol{r}) = \Psi_{j(-\boldsymbol{k})}(\boldsymbol{r})$. Therefore,

$$\hat{T}\,\Psi_{j\boldsymbol{k}}(\boldsymbol{r}) = \Psi_{j(-\boldsymbol{k})}(\boldsymbol{r}). \tag{1}$$

Furthermore, since $\hat{T}$ commutes with the Kohn–Sham Hamiltonian, $\hat{H}$, we also have $\varepsilon_{jk} = \varepsilon_{j(-k)}$.

In LOBSTER, TR symmetry can be exploited by directly incorporating it into the equations of chemical-bonding indicators, i.e., COOP and COHP. Since COOP and COHP only differ in the use of overlap and Hamiltonian matrices, respectively, in the following we will derive how time reversal symmetry affects COOP, and related conversions can be done for COHP as well.

The first step in implementing TR symmetry in COOP is to split the $\boldsymbol{k}$-point summation into two groups, in which one group is the inversion of the other, and to realize that the each $\boldsymbol{k}$-point and its inversion are energetically degenerate:

$$\mathrm{COOP}_{\mu T \nu T'}(E) =$$

$$S_{\mu T;\nu T'} \sum_{j} \left[\sum_{k \in \left(\frac{BZ}{2}\right)} w_k \, C^*_{\mu T;jk} \, C_{\nu T';jk} \right.$$

$$\left. + \sum_{k'=-k} w_{k'} \, C^*_{\mu T;jk'} \, C_{\nu T';jk'} \right] \delta(\varepsilon_{jk} - E). \quad (2)$$

All the notations are the same as described in Ref. [12]. We also recall that the reconstructed wavefunctions are given as linear combinations of atomic orbitals (LCAO):

$$|\Psi_{jk}\rangle = \sum_{\mu T} C_{\mu T;jk} |\chi_{jT}\rangle. \quad (3)$$

Here, $|\chi_{jT}\rangle$ are, in general, real functions, and thus unaffected by TR symmetry. Accordingly, if we apply eq. (1) to eq. (3), we have

$$C^*_{\mu T;jk} = C_{\mu T;j(-k)}. \quad (4)$$

Then, if eq. (4) is re-inserted into eq. (2), for each $\boldsymbol{k}$-point there is a term that adds with its complex conjugate, which is equivalent to twice the real part of the same term. In the end, we arrive at

$$\mathrm{COOP}_{\mu T \nu T'}(E) =$$

$$S_{\mu T;\nu T'} \sum_{jk \in \left(\frac{BZ}{2}\right)} \widetilde{w}_k \, \Re\{ C^*_{\mu T;jk} \, C_{\nu T';jk} \} \, \delta(\varepsilon_{jk} - E), \quad (5)$$

where $\widetilde{w}_k = 2w_k$. Here, $\Re\{\dots\}$ means the real part of a given complex value. Note that the summation over $\boldsymbol{k}$-points now only covers half of the BZ. Correspondingly, by substituting $\hat{S}$ with $\hat{H}$, the TR-exploited COHP now reads

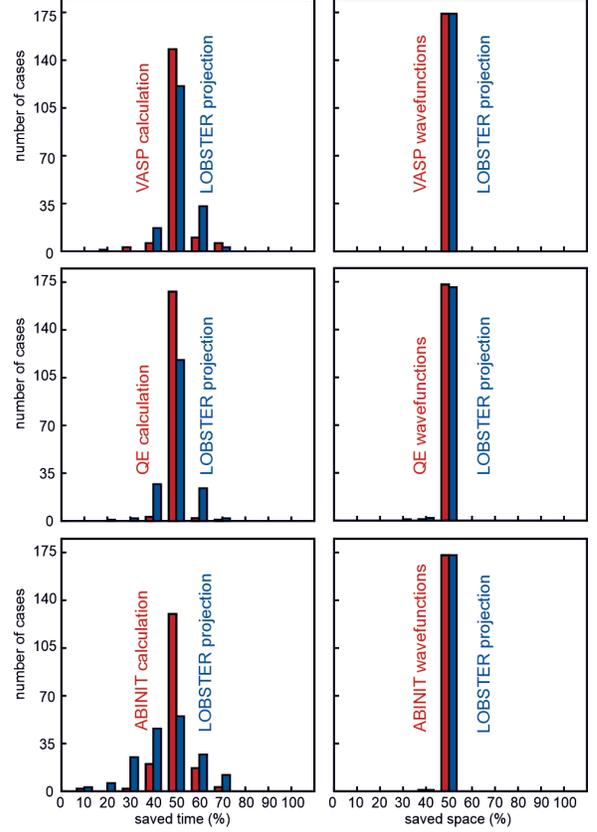$$\mathrm{COHP}_{\mu T \nu T'}(E) =$$

Figure 2. The distributions of saved time (left) and storage-space (right) gained in DFT calculation (red) and LOBSTER projection (blue) after using TR symmetry.

$$H_{\mu T;\nu T'} \sum_{jk \in \left(\frac{BZ}{2}\right)} \widetilde{w}_k \, \Re\{ C^*_{\mu T;jk} \, C_{\nu T';jk} \} \, \delta(\varepsilon_{jk} - E). \quad (6)$$

With these modified equations of chemical-bonding indicators, there is a huge gain in efficiency not only in LOBSTER but also in pre-projection DFT calculations. To demonstrate the benefits of TR symmetry in chemical-bonding analysis, the new LOBSTER version was tested by checking both the DFT and the projection parts for a total of 174 crystalline cases (systems) of various kinds, including insulators, metals, and magnetic materials. The systems were selected from the well-known material database materialsproject.org.[20] This database, moreover, has a large collection of materials that can be downloaded manually or automatically through the material analysis code pymatgen.[21]

The DFT part produced PAW-based DFT data from three different DFT programs, namely VASP, ABINIT, and QE. To achieve DFT convergence throughout, the number of $\boldsymbol{k}$-points targeted a $\boldsymbol{k}$-point density ranging from 6000 to 12000 $\boldsymbol{k}$-points×atoms. The cutoff

energy values used to converge all the DFT calculations are listed in Table 1 of the SI.

Furthermore, we have recently adapted the pymatgen code such that it provides users with VASP and LOBSTER input files specifically for LOBSTER calculations. That is to say that a local basis set for the projection is automatically suggested, and the VASP calculation is adapted accordingly. The VASP input files have then been translated into similar input files for QE and ABINIT using independent scripts. The pymatgen code was also used and partially adapted to compare all computations with and without TR symmetry. These developments can be found in pymatgen version 2019.12.22 or newer. All raw data for our comparison can be accessed via https://zenodo.org/record/3676916.

Once completed, all information about the runtime and storage-space from the DFT and projection calculations with and without TR symmetry was compiled for analysis and benchmarked against its counterpart using the current release of LOBSTER (3.2.0). In the following paragraphs, we report the highlighted summary of the time and storage-space efficiency gained due to this new functionality. The complete numerical information can be found in Table 2 of the SI.

Fig. 2 shows the percentage distributions as regards saved time (left) and storage-space (right) with respect to prior DFT calculations (red) and LOBSTER projections (blue) after having incorporated TR symmetry. Regarding time efficiency (Fig. 2 left), all distributions are normally centered at the 50% value such that all the three DFT programs saved 50% runtime in most cases, ABINIT showing the largest variance. Likewise, the time saved by LOBSTER projection using TR symmetry also follows a normal distribution, LOBSTER typically running twice as fast.
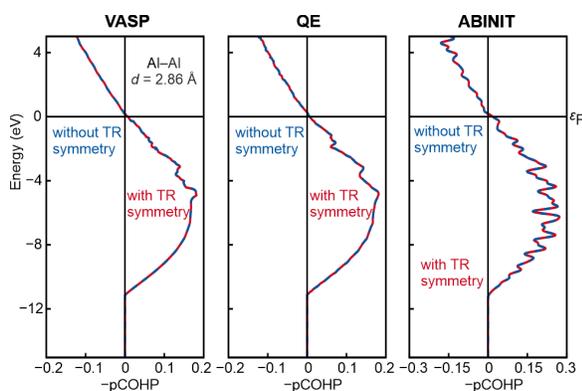

Figure 3. COHP of Al calculated after the LOBSTER projection of PAW data without (blue) and with (red) TR symmetry from three different DFT programs. The curves are shown as dashed lines and are indistinguishable to the naked eye.

The storage-space efficiency gained by TR symmetry is even better (Fig. 2 right) because the distributions have zero or nearly zero variance. In other words, DFT and LOBSTER calculations including TR symmetry almost always use 50% less storage-space.

We have also checked the quality of the total density of states (TDOS) produced by this LOBSTER version compared to standard LOBSTER by measuring the TDOS overlap mismatch which is defined as

$$\% \text{ of TDOS overlap mismatch} = \frac{\int_{-\infty}^{\infty} d\varepsilon \, |\text{TDOS}_w(\varepsilon) - \text{TDOS}_{wo}(\varepsilon)|}{\int_{-\infty}^{\infty} d\varepsilon \, \text{TDOS}_{wo}(\varepsilon)} \times 100\%, \quad (7)$$

thereby quantifying the match between the TDOS with (TDOS_w) and without (TDOS_{wo}) TR symmetry. The numerical results as detailed in Table 3 of the SI evidence a typical overlap mismatch of less than 0.1% and a maximum mismatch still smaller than 10%. In most cases, the DOS with and without TR symmetry are indistinguishable. Moreover, out of 174 cases, there are only 3 cases in VASP, 0 cases in ABINIT, and 2 cases in QE with mismatches larger than 5%; they are likely to go back to a still insufficient number of k-points. It is also worth mentioning that the results for spin-polarized and non-spin-polarized cases do not at all differ in quality, a reassuring sign for magnetic systems.

Regarding COHP, the results in general agree well with the TDOS although there are few cases (not shown here) where COHP mismatch values are larger than for the TDOS counterparts such that k-dependency is more pronounced here, so these cases might still need additional k-points (since covalent bonding may become highly directional and then requires extreme k-convergence). As a more typical example, Fig. 3 shows a COHP plot of one particular test case, i.e. aluminum (mp-134), in which the COHP calculated without TR symmetry using PAW data from different DFT programs can be perfectly reproduced by the much quicker TR approach. Note that the plot from ABINIT is wigglier than those from VASP and QE because the tetrahedron table necessary for the tetrahedron method is not output by ABINIT, so LOBSTER must use the default method, i.e., Gaussian smearing. Nevertheless, as shown in Table 3 of the SI, for all three programs the TDOS overlap mismatches of Al (mp-134) are zero in full agreement with COHP.

New Analysis Tools

Once the LCAO coefficients have been reconstructed (Fig. 1), LOBSTER can readily calculate DOS and COOP

by k-space integration.[12] Furthermore, reconstructing the Hamiltonian matrix in a second step facilitates COHP analysis.[2],[12] Additionally, LOBSTER writes their energy-integrated counterparts IDOS (which yields the total number of electrons of the respective atoms, i.e. Mulliken's gross population), ICOOP (Mulliken's overlap population) and ICOHP.

Population analyses and charge

Computing the gross populations and charges is a feature which was recently added in LOBSTER[22],[23] using the techniques of population analyses as introduced by Mulliken and Löwdin.[24]-[27] In Mulliken's approach, the gross (orbital) population GP_μ is determined as follows by using the density-matrix formalism

$$GP_\mu = \sum_k \sum_\nu P_{\mu\nu;k} S_{\mu\nu;k} w_k, \qquad (8)$$

with $P_{\mu\nu;k}$ and $S_{\mu\nu;k}$ as k-dependent density and overlap matrix elements for orbitals μ and ν, respectively,

$$P_{\mu\nu;k} = \sum_j f_{jk} C^*_{\mu;jk} C_{\nu;jk}, \qquad (9)$$

$$S_{\mu\nu;k} = \langle \chi_{\mu;k} | \chi_{\nu;k} \rangle. \qquad (10)$$

with the occupation number f_j of band j and the LCAO-CO coefficients $C_{\mu j;k}$ in reciprocal space.

For Löwdin's approach, a new basis set χ' and a new coefficient matrix C' are attained from the original ones by applying Löwdin's symmetric orthogonalization (LSO)[25],[27]

$$\chi'_k = S_k^{-1/2} \chi_k, \qquad (11)$$
$$C'_k = S_k^{1/2} C_k. \qquad (12)$$

The new basis functions χ', orthogonalized via LSO, are closest in Hilbert space to the original functions.[27] A new density matrix $P' = S^{1/2} P S^{1/2}$ is also obtained, so that the Löwdin gross orbital population GP_μ has the following form

$$GP_\mu = \sum_k P'_{\mu\mu;k} w_k. \qquad (13)$$

The Mulliken or Löwdin charge q_A for any atom A is computed from the difference of the number of the atom's valence electrons N (when using pseudopotentials) and the gross orbital population GP_μ including all orbitals μ belonging to atom A,

$$q_A = N - \sum_{\mu \in A} GP_\mu. \qquad (14)$$

We have recently demonstrated that the orbital-based Mulliken and Löwdin population analyses, as implemented in LOBSTER by combining the advantages of PW and all-electron local basis sets, no longer suffer from the typical basis-set dependency as known from MO calculations. Additionally, both methods have been proven as being highly competitive approaches compared to the traditional practice of using the density-based Bader charge analysis.[22],[28] A comparison of both methods for one application will be given later in the text.

k-dependent COHP

Another feature recently added in LOBSTER is the k-dependent COHP. While COHP is useful in describing chemical bonding in energy space, the chemical bonding often is direction-dependent, so a particular k-point may be of special interest or at least more significant than bonding at any other k-points. Whenever topology comes into play, knowing this information may be crucial in studying a given material. We have therefore developed a new routine, based on the original definition of COHP, that resolves COHP into individual contributions from each k-point and band. Mathematically, the k-dependent COHP is defined as:

$$COHP_{\mu\nu;jk} = C^*_{\mu;jk} H_{\mu\nu;k} C_{\nu;jk}. \qquad (15)$$

The first k-dependent COHP application was demonstrated in an optimization study of thermoelectric materials eventually yielding n-type Mg_3Sb_2.[29] Using the k-dependent COHP, the authors revealed the main covalent interactions existing at the conduction-band minimum which plays a decisive role in improving the thermoelectric property of the material. The new feature was also employed in a high-throughput computational identification of a high Curie-temperature ferromagnetic semiconductor, $In_2Mn_2O_7$,[30] and it also rationalized the distinct semiconducting and ferromagnetic properties. Likewise, k-dependent COHP contributed to high-throughput computational screening of diamond-like ABX_2 compounds[31] such as to help predicting exceptional and novel thermoelectric materials. Additionally, we demonstrate the use of the feature in one of the applications below.

Applications

To conclude this paper, we present three representative applications. The first one demonstrates the advantage of the TR functionality, whereas the second
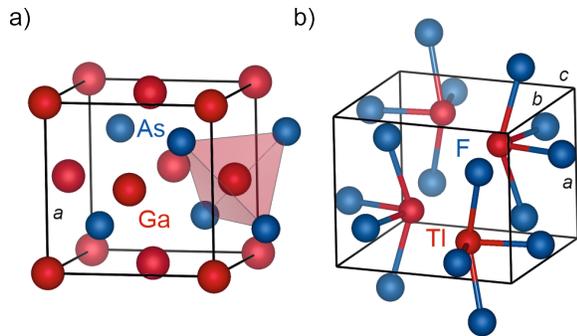
Figure 4. The crystal structure of a) GaAs and b) TlF.

and the third applications describe, respectively, the population analyses-and-charge feature and the use of the k-dependent COHP technique.

Gallium Arsenide (mp-2534)

The III-V semiconductor GaAs, our first application, is part of the test cases and also included in the LOBSTER package (Fig. 4a). It is a wide-bandgap semiconductor crystallizing in the zinc blende structure ($F\bar{4}3m$) with two atoms per primitive unit-cell. The PAW data of GaAs presented here were obtained by QE using a $14 \times 14 \times 14$ k-point mesh.

Without TR symmetry, the DFT calculation of GaAs used more than 4 CPU hours but by exploiting TR symmetry (see Table 2) the runtime was reduced by more than 50%. Likewise, the LOBSTER projection calculation without TR symmetry took about 17 CPU minutes but TR symmetry reduced its runtime by 51%. Taking DFT and LOBSTER projection together, TR symmetry
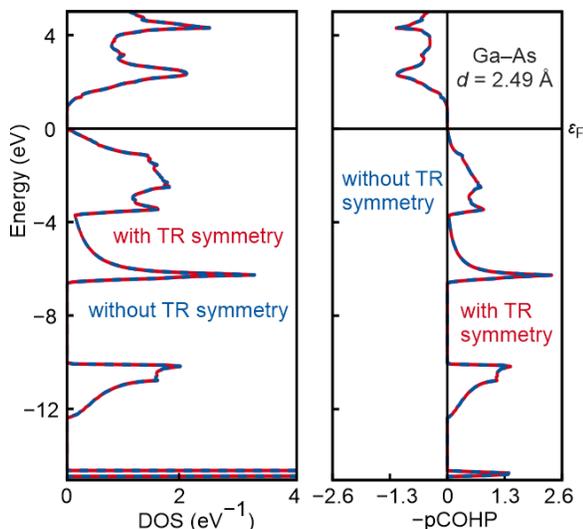

Figure 5. Total DOS (left) and the shortest Ga–As bond's COHP (right) calculated after the LOBSTER projection of QE PAW data without (blue) and with (red) the time reversal (TR) symmetry.

reduces the total runtime from 296 min to a mere 145 min, a speedup of 2.04.

Fig. 5 shows TDOS (left) and COHP (right) of GaAs obtained by projecting QE PAW wavefunctions onto local orbitals without (blue) and with (red) TR symmetry. Clearly, both TDOS and COHP plots cannot be distinguished. The entire valence band is Ga–As bonding, just like for the isoelectronic silicon phase. Furthermore, as given in Table 3 of the SI, the DOS overlap mismatches are virtually zero for all three different DFT programs.

Lithium intercalation in carbonaceous anode materials

Carbon nanomaterials including "hard" and "porous" carbons with substantial degrees of structural disorder are promising candidates for the application as anode materials in Li-ion and Na-ion batteries. Such structures were generated by Huang et al. via a combination of a machine-learning-based interatomic potential to generate the carbonaceous host framework and DFT to then describe the Li intercalation, using a protocol detailed in Ref. [28]. One, Li-rich, structural model is directly taken from that study and used here
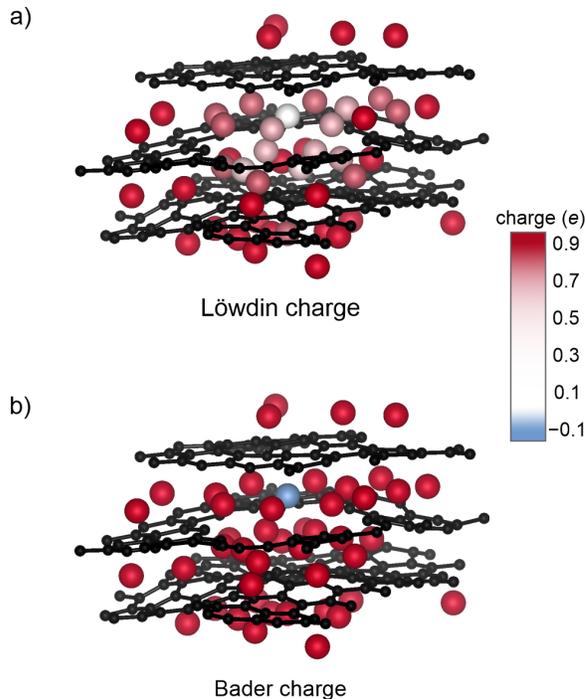

Figure 6. a) Löwdin and b) Bader charges of a Li-rich structural model of intercalation in disordered carbon. The figure shows only part of the simulation cell to better focus on the central region, where a pore in one of the graphite-like sheets is found; see Ref. [28].

without further modification; albeit such charge analyses have been reported in [28], we repeat them here to exemplify the approach (and to directly compare two separate charge-analysis methods as applied to the same structure). Because the intercalation behavior of anode materials is important for the functionality of batteries (i.e., the charging process), Löwdin charges on the Li atoms in this structure are examined and compared to those from the widely-used method of Bader charge analysis.[32] Recently strong differences in consuming computing resources between Bader's method compared to the Mulliken and Löwdin population analyses were demonstrated.[22] Similar results are also found now, as shown in Fig. S1 in the SI. Another problem working with the electron density is that this method sometimes yields chemically non-intuitive charges, e.g. negative charges on alkali metals, like Na and Li, if combined with the more electronegative carbon, as reported in a previous study of such systems[33] and also seen in our present calculations in Fig. 6b. According to Bader charge analysis, there is one Li ion with a small negative charge occupying the center of the pore in the carbon backbone, while the Li ions around it and further distributed in the structure exhibit charges close to +1. This kind of charge distribution is at variance with the chemically intuitive assumption of an alkali metal being either positively charged or neutral (metallic), but not negative. In the case of Löwdin charges (Fig. 6a) directly and far more quickly calculated from the wavefunction, a qualitatively different situation is observed. The center of the pore is occupied by an almost metallic (uncharged) Li ion, from which other Li ions with increasing charges are clustering around it, so that a smooth charge distribution from the pore to the rest of the structure can be observed.[28] Details on the charges can be found in Fig. S2 in the SI. Whilst the intercalation of isolated Li atoms into carbonaceous materials is expected to lead to a charge stage close to Li⁺, it has been argued that pores in disordered carbons can be filled with metallic-like clusters, and evidence of a (partial) charge transfer has been obtained, for example, from operando NMR measurements on Li[34] and Na[35] anodes. For example, approximate charges of +0.66 e and +0.1 e have been deduced for two classes of Li atoms in such environments,[34] in qualitative agreement with the findings of Huang et al.[28] and with the data exemplified in Fig. 6.

Thallium Fluoride (mp-720)

LOBSTER can also be employed to analyze orbital characters of electronic bands through fatbands and orbital interactions in those bands through the k-dependent COHP. To demonstrate such features, we use the simple, yet low-symmetry fluoride TlF, which is also part of our test cases. TlF is an insulator which has an orthorhombic structure ($Pbcm$) with 6 atoms per unit cell (Fig. 4b).

TlF is also known as one of the systems showing a "lone-pair" crystal structure associated with the so-called inert-pair effect. In a previous study, Häussermann et al.[36] have proposed a simple, yet very reasonable, bonding picture to understand why TlF adopts the aforementioned orthorhombic instead of the more symmetric CsCl structure. Their qualitative picture, based on the DOS and a schematic diagram of orbital interactions of Tl and F in the hypothetical
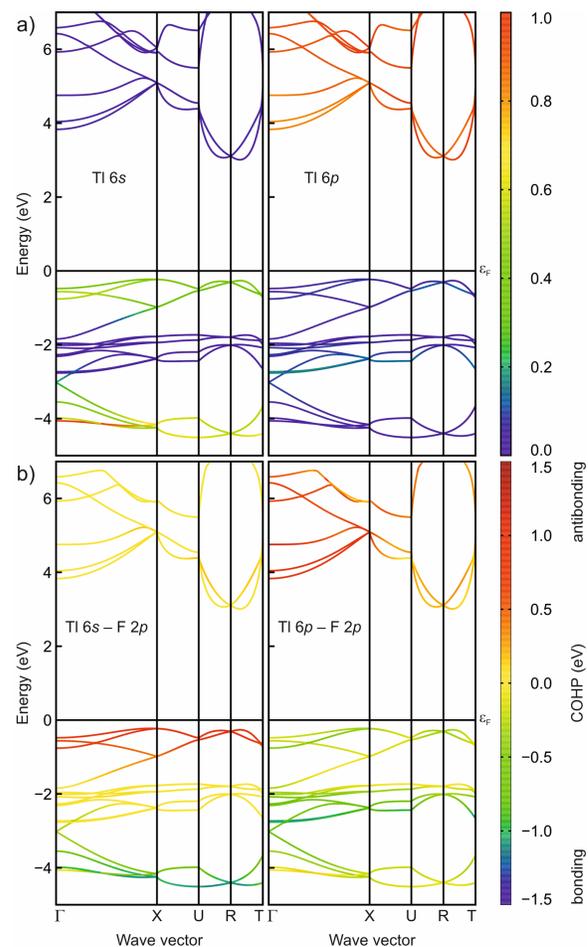

Figure 7. Contributions to TlF electronic structure from a) fatbands of Tl 6s (left) and Tl 6p (right) as well as b) k-dependent Tl–F COHP of 6s–2p (left) and 6p–2p (right).

CsCl structure, yields that the orthorhombic structure is a natural way to circumvent orbital interactions which would be quite repulsive in the CsCl structure. Here we will demonstrate for the first time a more quantitative and spatially resolved picture of the orbital interactions of TlF supporting the earlier deductive conclusion by Häussermann et al.[36]

PAW wavefunctions of TlF needed for the projection were calculated by VASP using 763 k-points, also including TR symmetry saving runtime by about 50% (Table 2 of SI). After the projection, the $6s$ and $6p$ fatbands of Tl were calculated and are depicted in Fig. 7a. As expected, the valence bands just below the Fermi level are moderately composed by Tl $6s$ (left, the inert pair showing up here), whereas the (unoccupied) conduction bands are almost fully dominated by Tl $6p$ (right). In addition, F $2p$ (not shown) also moderately contributes to the top of the valence bands.

While the fatband feature reveals the orbital characters of each band, it does not provide any insight into the mechanism of the electronic structure formation through chemical bonding. To directly gain that information, Fig. 7b depicts the k-dependent COHP of Tl $6s$ — F $2p$ (left) and Tl $6p$ — F $2p$ (right) interactions. As usual, negative (positive) values represent bonding (antibonding) interactions.

To start with, the top of the valence bands (about 1 eV below the Fermi level) is composed of strong Tl $6s$ — F $2p$ antibonding interactions, so the "inert pair" is not at all inert but strongly destabilizing at this particular energy. More below, roughly between −2 and −4 eV, however, there are two other sets, now mainly exhibiting nonbonding and bonding Tl $6s$ — F $2p$ character, such that the bonding arising from Tl $6s$ — F $2p$ interactions in its entirety is hardly significant and will not stabilize the system; the "inert pair" is in charge, once again. Quite to the contrary, one finds moderately strong Tl $6p$ — F $2p$ bonding interactions in the valence bands throughout, as seen from Fig. 7b (right), and they stabilize the system. The latter interactions are possible due to the adopted orthorhombic structure with symmetry lower than the CsCl structure. Note also that such Tl $6p$ — F $2p$ interactions would be strongly antibonding if occupied in the conduction band (say, by electronic excitation), and the strength of the antibonding effects varies within the Brillouin zone, very much direction-dependent, being most destabilizing at Γ but only weakly so at **R**.

These k-dependent COHP pictures are in agreement with the prior description given by Häussermann et al.,[36] but the k-dependent COHP provides a more quantitative picture and also allows for a more straightforward and convenient analysis.

Conclusions

We have presented new developments in the LOBSTER software for chemical-bonding analysis. LOBSTER processes delocalized PAW wavefunctions calculated with VASP, ABINIT, or QE and performs projection onto an auxiliary LCAO basis, which makes bond-analytic tools such as DOS, COOP, and COHP accessible for state-of-the-art PAW-based simulations. Furthermore, LOBSTER's efficiency has been significantly improved such as to save 50% of the entire process (DFT and LOBSTER calculations) as regards time and storage-space, namely by implementing time-reversal symmetry in the chemical bonding routines.

Recent development has also added two major features, namely atomic charges and related populations directly from the wavefunction as well as the k-dependent COHP. The former is useful for examining ionic bonding in suchlike (or polar) compounds, for example simple salts, Zintl phases, intermetallics, thermoelectrics, or even phase-change materials,[22][23][23],[37]-[39] whereas the latter serves useful for detailed band-wise and k-point-wise chemical bonding analysis.

Acknowledgments

References and Notes

[1] T. Hughbanks, R. Hoffmann, *J. Am. Chem. Soc.* **1983**, *105*, 3528.

[2] R. Dronskowski, P. E. Blöchl, *J. Phys. Chem.* **1993**, *97*, 8617.

[3] O. K. Andersen, O. Jepsen, *Phys. Rev. Lett.* **1984**, *53*, 2571.

[4] R. Dovesi, A. Erba, R. Orlando, C. M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rerat, S. Casassa, J. Baima, S. Salustro, B. Kirtman, *WIREs Comput. Mol. Sci.* **2018**, *8*, e1360.

[5] a) D. Sánchez-Portal, E. Artacho, J. M. Soler, *Solid State Commun.* **1995**, *95*, 685; b) D. Sánchez-Portal, E. Artacho, J. M. Soler, *J. Phys. Condens. Matter* **1996**, *8*, 3859.

[6] M. D. Segall, R. Shah, C. J. Pickard, M. C. Payne, *Phys. Rev. B* **1996**, 54, 16317.

[7] L. Foppa, C. Copéret, A. C.-Vives, *J. Am. Chem. Soc.* **2016**, 138, 16655-16668.

[8] J. Lin, S. Zhang, W. Guan, G. Yang, Y. Ma, *J. Am. Chem. Soc.* **2018**, 140, 9545-9550.

[9] C. Liu, I. Tranca, R. A. van Santen, E. J. M. Hensen, E. A. Pidko, *J. Phys. Chem. C* **2017**, *121*, 42, 23520-23530.

[10] C. Zhong, B. Liu, J. Ding, X. Liu, Y. Zhong, Y. Li, C. Sun, X. Han, Y. Deng, N. Zhao, W. Hu, *Nat. Energy* **2020**, https://doi.org/10.1038/s41560-020-0584-y.

[11] P. E. Blöchl, *Phys. Rev. B* **1994**, *50*, 17953.

[12] a) S. Maintz, V. L. Deringer, A. L. Tchougréeff, R. Dronskowski, *J. Comput. Chem.* **2013**, *34*, 2557; b) S. Maintz, V. L. Deringer, A. L. Tchougréeff, R. Dronskowski, *J. Comput. Chem.* **2016**, *37*, 1030–1035.

[13] a) G. Kresse, J. Hafner, *Phys. Rev. B* **1993**, *47*, 558; b) G. Kresse, J. Furthmüller, *Comput. Mater. Sci.* **1996**, *6*, 15; c) G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169; d) G. Kresse, D. Joubert, *Phys. Rev. B* **1999**, *59*, 1758.

[14] a) X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.-Y. Raty, D.C. Allan, *Comput. Mater. Sci.* **2002**, *25*, 478; b) X. Gonze, G.-M. Rignanese, M. Verstraete, J.-M. Beuken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zerah, M. Mikami, Ph. Ghosez, M. Veithen, J.-Y. Raty, V. Olevano, F. Bruneval, L. Reining, R. Godby, G. Onida, D.R. Hamann, D.C. Allan, *Z. Kristallogr.* **2005**, *220*, 558; c) X. Gonze, B. Amadon, P. M. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Cote, T. Deutsch, L. Genovese, Ph. Ghosez, M. Giantomassi, S. Goedecker, D. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M. J. T. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G.-M. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M. J. Verstraete, G. Zérah, J. W. Zwanziger, *Comput. Phys. Commun.* **2009**, *180*, 2582; d) M. Torrent, F. Jollet, F. Bottin, G. Zérah, X. Gonze, *Comput. Mater. Sci.* **2008**, *42*, 337.

[15] a) P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. Fabris, G. Fratesi, S. de Gironcoli, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, R. M. Wentzcovitch, *J. Phys.: Condens. Matter* **2009**, *21*, 395502; b) P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio Jr., A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la-Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, S. Baroni, *J. Phys.: Condens. Matter* **2017**, *29*, 465901.

[16] R. Nelson, P. M Konze, R. Dronskowski, *J. Phys. Chem. A* **2017**, *121*, 40, 7778-7786.

[17] http://www.boost.org.

[18] G. Guennebaud, B. Jacob, http://eigen.tuxfamily.org, **2010**.

[19] M.S. Dresselhaus, G. Dresselhaus, A. Jorio, Group Theory; Springer: Verlag Berlin Heidelberg, **2008**.

[20] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Materials* **2013**, *1(1)*, 011002.

[21] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. Chevrier, K. A. Persson, G. Ceder, *Comput. Mater. Sci.* **2013**, *68*, 314–319.

[22] C. Ertural, S. Steinberg, R. Dronskowski, *RSC Adv.* **2019**, *9*, 29821–29830.

[23] W.-L. Li, C. Ertural, D. Bogdanovski, J. Li,R. Dronskowski, *Inorg. Chem.* **2018**, *57*, 12999–13008.

[24] a) R. S. Mulliken, *J. Chem. Phys.* **1955**, 23, 1833; b) R. S. Mulliken, *J. Chem. Phys.* **1955**, *23*, 1841; c) R. S. Mulliken, *J. Chem. Phys.* **1955**, *23*, 2338; d) R. S. Mulliken, *J. Chem. Phys.* **1955**, *23*, 2343.

[25] P. O. Löwdin, *J. Chem. Phys.* **1950**, *18*, 365–375.

[26] R. Dronskowski, Computational Chemistry of Solid State Materials - A Guide for Material Scientists, Chemists, Physicists and others, WILEY-VCH Verlag GmbH & Co. KGaA:Weinheim, **2005**.

[27] L. Piela, Ideas of Quantum Chemistry (Appendix J), Elsevier B.V., Amsterdam, First edn., **2007**.

[28] J.-X. Huang, G. Csányi, J.-B. Zhao, J. Cheng, V. L. Deringer, *J. Mater. Chem. A* **2019**, *7*, 19070–19080.

[29] X. Sun, X. Li, J. Yang, J. Xi, R. Nelson, C. Ertural, R. Dronskowski, W. Liu, G. J. Snyder, D. J. Singh, W. Zhang, *J. Comput. Chem.* **2019**, *40*, 1693–1700.

[30] W. Chen, J. George, J. B. Varley, G.-M. Rignanese, G. Hautier, *npj Comput. Mater*, **2019**, *5*, 72.

[31] R. Li, X. Li, L. Xi, J. Yang, D. J. Singh, W. Zhang, *ACS Appl. Mater. Interfaces* **2019**, *11*, 28, 24859-24866.

[32] a) R. F. W. Bader, *Chem. Rev.* **1991**, *91*, 893–928; b) R. F. W. Bader, Atoms in Molecules: A Quantum Theory, Oxford University Press, Oxford, **1994**; c) G. Henkelman, A. Arnaldsson, H. Jónsson, *Comput. Mater. Sci.*, **2006**, *36*, 354–360; d) E. Sanville, S. D. Kenny, R. Smith, G. Henkelman, *J. Comp. Chem.* **2007**, *28*, 899–908; e) W. Tang, E. Sanville, G. Henkelman, *J. Phys.: Condens. Matter* **2009**, *21*, 084204; f) M. Yu, D. R. Trinkle, *J. Chem. Phys.* **2011**, *134*, 064111.

[33] V. L. Deringer, C. Merlet, Y. Hu, T. H. Lee, J. A. Kattirtzi, O. Pecher, G. Csanyi, S. R. Elliott, C. P. Grey, *Chem. Commun.* **2018**, *54*, 5988–5991.

[34] M. Letellier, F. Chevallier, C. Clinard, E. Frackowiak, J.-N. Rouzaud, F. Béguin, M. Morcrette, J.-M. Tarascon, *J. Chem. Phys.* **2003**, *118*, 6038.

[35] J. M. Stratford, P. K. Allan, C. P. Grey, *Chem. Commun.* **2016**, *52*, 12430.

[36] U. Häussermann, P. Berastegui, S. Carlson, J. Haines, J.-M. Léger, *Angew. Chem.* **2001**, *113*, 4760.

[37] P. M. Konze, R. Dronskowski, V. L. Deringer, *Phys. Status Solidi RRL* **2019**, *13*, 1800579.

[38] P. Bielec, R. Nelson, R. P. Stoffel, L. Eisenburger, D. Günther, A.-K. Henß, J. P. Wright, O. Oeckler, R. Dronskowski, W. Schnick, *Angew. Chem. Int. Ed.* **2019**, *58*, 1432.

[39] K. Chen, M. Fehse, A. Laurita, J. J. Arayamparambil, M. T. Sougrati, L. Stievano, R. Dronskowski, *Angew. Chem. Int. Ed.* **2020**, *59*, 3718 – 3723.