Quasi-Four-Component Method with Numeric Atom-Centered

Orbitals for Relativistic Density Functional Simulations of Molecules

and Solids

Rundong Zhao^{1,2}, Victor Wen-zhe Yu¹, Kimberly Zhang³, Yunlong Xiao⁴, Yong Zhang⁵ and Volker

Blum^{1*}

¹Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC

27707, U.S.A.

²School of Physics, Beihang University, Beijing 100191, China.

³Department of Chemistry, University of California, Irvine, Irvine, CA 92697, U.S.A.

⁴College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

⁵Qingdao Institute for Theoretical and Computational Sciences, Shandong

University, Qingdao 266237, China

* Corresponding author; E-mail: volker.blum@duke.edu

Abstract

We describe and benchmark the quasi-four-component (Q4C) approach to relativistic density functional simulations of molecules and solids, using precise, numerically tabulated atom-centered orbital (NAO) basis sets to discretize Dirac's Equation. The Q4C approach initially projects the atomic solution to (electron-only) positive-energy states and eventually deals with only two components but retains the precision of traditional four-component relativistic methods. While Q4C inherently reduces the dimension of the Hamiltonian matrix in diagonalization, the adoption of localized NAO basis functions in solids further limits the computational demand in real space operations, promising a pathway to investigate large and complex systems containing heavy elements with the precision of a four-component method. Here, we first perform validation and benchmark calculations for cohesive properties of a set of diatomic molecules and of previously established periodic model systems (i.e., silver halides). Then, we report Q4C relativistic energy band structure benchmarks for a series of 103 periodic materials, including chemical elements up to Bi, and providing quantitative comparisons to more approximate scalar-relativistic and spin-orbit coupled treatments. Finally, we demonstrate the applicability of the method to band structure calculations of simple and complex hybrid organic-inorganic (HOIP) perovskites containing Pb and Bi, i.e. Cs₂AgBiCl₆ and a larger system (containing 94 atoms per unit cell), (4-FPEA)₂PbI₄. The effect of full Q4C, compared to scalar relativity, on binding energies can be significant even for relatively light p-orbital bonded main group elements such as Br and I – i.e., 0.3 eV and 0.6 eV for Br_2 and I_2 binding energies, respectively.

I. Introduction

Introductory quantum mechanics is often rationalized by the non-relativistic (NR) Schrödinger equation, but relativistic effects occur in any material or molecule, albeit at different scales. In the light-element material graphene (i.e., carbon, Z=6), high-precision measurements are needed to reveal a spin-orbit coupling gap of 42.2 $\mu eV.[1]$ Towards heavier elements, the importance of relativistic effects increases drastically with the atomic number Z. In the energy band structures of intermediate-Z semiconductors such as Ge, GaAs, or InSb, relativistic effects can no longer be neglected for a qualitatively correct understanding [2, 3] of their electronic properties. In compounds including heavy main-group elements such as Bi and Pb, relativistic effects on energy band structures can reach the order of several eV.[4] A host of other significant properties shaped by relativity – e.g., famously, a $\sim 10\%$ lattice contraction of gold as well as its golden color – is well documented in other materials containing heavy elements.[5-8] The level of theory commonly considered to be mathematically sufficient to capture all relevant relativistic effects in chemistry and materials physics is the Dirac equation, in which the wave function is a vector of four scalar components, which are functions of space and time. The first two scalar component functions are conventionally grouped together as a two-dimensional vector, together called the "large component". Similarly, the third and fourth scalar component functions, grouped together as a second twodimensional vector, are called the "small component". Traditional four-component (4C) schemes including the Dirac-Kohn-Sham (DKS) method[9-11] (i.e., relativistic density functional theory (DFT)) or the 4C coupled cluster [12, 13] method have long been in existence. DFT[14, 15] is today the main available first-principles method for solids at a 4C relativistic level. Relativistic all-electron DFT calculations in solids can be carried out, e.g., using the relativistic Korringa-Kohn-Rostoker (KKR)[16, 17], the relativistic augmented plane wave (RAPW)[18] scheme or the full-potential linear muffin-tin orbital approach[19]. 4C DKS codes for periodic systems have also been reported using numerical atomic basis sets[20, 21] and Gaussian type orbitals[22], respectively, while recently an exact two component (X2C) algorithm based on Slater type orbitals[23] was shown to have comparable accuracy.

From a perspective of accuracy and generality, 4C calculations would be a desirable standard in all electronic structure based simulations in materials science and chemistry. However, the presence of four components adds some mathematical complexity and, depending on the particular implementation, can necessitate additional approximations and/or simply increases the computational cost. Instead, a host of simpler approximations were established in the literature, typically grouped into scalar relativistic (SR) methods[24-26] (effectively treating one-component wave functions or two separate spin components that are only indirectly coupled through the exchange-correlation potential) and methods that include spin-orbit coupling (SOC)[26, 27], i.e., effective two-component methods. While all these alternative approaches are very well understood, their accuracy does depend on the atomic number of the elements involved, as well as on the detailed objectives of a particular simulation. Nevertheless, the vast majority of production simulations avoid outright 4C treatments. Four-component simulations are typically applied to relatively small systems but are much less (if at all) available for larger, more complex systems in production quality. Similarly, much of the functionality that is now well established in scalar-relativistic electronic structure codes is harder to access in its four-component variants.

In this paper, we take a step towards affordable relativistic density functional theory for production simulations of large, complex systems by combining the quasi four-component (O4C) method for relativity and the numeric atom-centered orbital (NAO) basis set approach. The FHIaims code[28, 29] is used for the practical implementation of the method. In the context of SR and SR+SOC electronic structure theory, this NAO-based code already offers high precision[4, 30, 31] and can be applied to large, complex systems at affordable computational cost. Q4C, in analogy to the exact two-component (X2C) method[23, 32, 33], is a two-component formalism that was first developed for molecules [32-37]; the latter was later extended to periodic systems in a previous study[23]. As in other approaches, the objective of Q4C and X2C is to separate the treatments of the large and small components and therefore simplify the four-component representation to a twocomponent one for the majority of the computation. However, Q4C and X2C retain access to the small component through practically exact, atomic-like approximations and can therefore achieve comparable precision to a fully 4C method.[32, 37] We here demonstrate high precision and accuracy of our approach for a broad range of closed-shell molecules and solids. The NAO basis sets used here contain, as a subset, all the basis functions (large and small components) generated by solving the four-component Dirac equation for atoms[38]. With this choice, Q4C turns out to be a simple and effective way to match the precision of the X2C approach.

Alternative and commonly employed, approximate effective one- and two-component schemes for electronic structure theory include the zero-order regular approximation (ZORA)[39-41] method and its more practically usable gauge-invariant versions (e.g., atomic ZORA as implemented in the FHI-aims code[28] used here), the Douglas-Kroll-Hess (DKH)[42-45] method and its generalization to different orders[24, 46-49]. Relativistic versions of the widely used (linearized) augmented plane wave ((L)APW) method[27, 50] also constitute a precise framework to capture the main relativistic effects in solids[51, 52]. In these implementations, SOC is often treated as a correction term[26, 27] added to a preceding, SR calculation: One first diagonalizes a SR Hamiltonian to obtain the eigenvalues and eigenvectors, then truncates them and generates a new basis set for diagonalizing the full Hamiltonian (including the SOC term). This approach typically reduces the size of the full Hamiltonian matrix, e.g., by imposing an energy cutoff for unoccupied SR states to be included. A limitation is that the SR basis set used in the first step could be insufficient to represent the actual shape of SR+SOC orbitals, particularly $p^{1/2}$ derived ones that do not occur in SR theory at all.[53] Thus, one may face convergence problems especially when dealing with systems with *p*-type valence states and/or with extended core states[53, 54]. This situation was later improved by adding $p^{1/2}$ local orbitals as basis functions.[55] For energy band structures, nonself-consistent and self-consistent SOC correction schemes have been benchmarked extensively in reference[4]. The SOC correction scheme can also be combined with the GW method[56] and has been applied to extended systems such as lead chalcogenides[57], mercury chalcogenides[58] and perovskites[59-63]. Finally, the most widely used way of avoiding a direct treatment of the small component as well as of any other strong relativistic effects near a nucleus involves employing pseudopotentials[64-66], effective core potentials[67] or "pseudoized" parameterized effective orbitals (in the context of the projector-augmented wave methods[68]), e.g., in plane-wave based codes[69-72]. All of the latter approaches replace the details of the near-nuclear region, where relativistic effects matter most, with effective parameterized treatments based on different, much smoother orbitals. The parameterization can be performed based on relativistic all-electron calculations[73], leaving only the two scalar components of the "large component" as explicit ingredients of an electronic structure calculation but complicating the ability to directly recover core state properties and other physical quantities that arise near the nucleus when desirable.

One further objective of this paper is to validate and benchmark the present Q4C+NAO approach in comparison to selected approximate SR and SOC treatments. Specifically, we include the SR atomic ZORA treatment, which has already been shown to closely match a wide variety of other SR treatments (both all-electron and pseudoized) for the cohesive properties of simple solids[30]. We also include a NAO based SR plus non-selfconsistent SOC treatment of electronic band structures that has been subject to a previous broad benchmark covering 103 solids from light to heavy (*Pb* and *Bi*) elements in a past paper[4], as well as the self-consistent SR+SOC LAPW+ $p^{1/2}$ treatment mentioned above, as implemented in the *WIEN2k* code, a benchmark method that shows close agreement with the Q4C+NAO approach. The NR and atomic ZORA SR methods are furthermore compared to benchmarks of cohesive properties for molecular dimers, solid silver halides and heavy-element solids containing late 5*d* elements and *Pb*, *Bi* and *Po*. Band structures are finally compared for two larger, more complex systems, i.e. a 96-atom layered hybrid organic-inorganic perovskite (HOIP) unit cell at the NR SR, SR+SOC, and Q4C levels of theory, as well as the *Cs*₂*AgBiCl*₆[74, 75] double perovskite.

The remainder of the paper is organized as follows. Section II introduces relativistic electronic structure theory for the case of DFT and outlines the quasi four-component method. Section III presents important implementation details. In section IV, we first provide benchmarks for diatomic molecules. We then compare benchmark results for *AgCl*, *AgBr*, and *AgI* to previous studies and we report a series of benchmarks for a set of 103 compounds already used to quantify the precision of different approaches to SOC[4]. Section V provides conclusions and additional technical details are reported in an appendix.

II. Theory

A. Four-Component Dirac-Kohn-Sham Theory

We begin with the one-electron Dirac equation

$$\begin{pmatrix} \hat{V} & c\vec{\sigma} \cdot \vec{p} \\ c\vec{\sigma} \cdot \vec{p} & \hat{V} - 2mc^2 \end{pmatrix} \begin{pmatrix} \phi^L \\ \phi^S \end{pmatrix} = E \begin{pmatrix} \phi^L \\ \phi^S \end{pmatrix} \quad (1)$$

where ϕ^L and ϕ^S are the large and small components of the four-component Dirac wave function. In our work, the Dirac-Coulomb Hamiltonian[76] is employed, i.e., the non-relativistic Coulomb interaction is used for electron-nuclear and electron-electron interactions in Eq. (1), and orbital current terms are already neglected [19]. Although we do not pursue them here, corrections to the nonrelativistic Coulomb interaction can be derived from quantum electrodynamics; for example, Chapter 4 of reference [77] provides an extensive discussion of such corrections and how they can be included in atomic and molecular computations. From the second row of Eq. (1) one obtains the relation between the large and small components by an exact rewrite:

$$\phi_i^S(\vec{r}) = \frac{1}{2mc} R_i(\vec{r}) \left(\vec{\sigma} \cdot \vec{p}\right) \phi_i^L(\vec{r}), \quad (2a)$$

$$R_i(\vec{r}) = \left[1 + \frac{\varepsilon_i - V_{\text{eff}}(\vec{r})}{2mc^2}\right]^{-1}.$$
 (2b)

In Dirac-Kohn-Sham theory, V_{eff} stands for the effective potential, composed of the Hartree potential V_{H} , the electron-nuclear attraction V_{N} , and the exchange correlation potential V_{xc} . ε_{i} is the eigenvalue associated with orbital *i*; $\vec{\sigma}$ and \vec{p} are the Pauli matrices and the momentum operator, respectively. The total energy of the many-electron system is then calculated from

$$E_{tot} = \sum_{i=1}^{N_{states}} f_i \varepsilon_i - \int d^3 r \left[n(\vec{r}) V_{xc}(\vec{r}) \right] + E_{xc}[n] - \frac{1}{2} \int d^3 r \left[n(\vec{r}) V_H(\vec{r}) \right] + E_{N-N}, \quad (3)$$

where f_i denotes the occupation number of state *i*, n(r) denotes the electron density, and E_{N-N} denotes the internuclear repulsion energy, as in the nonrelativistic and scalar-relativistic Kohn-Sham formalism. However, the four-component electronic density in relativistic theory should contain the contribution from both large and small components, and is usually calculated as the sum of large component density and small component density:

$$n(\vec{r}) = n^{L}(\vec{r}) + n^{S}(\vec{r}) = \sum_{i=1}^{N_{states}} f_{i} \left(\phi_{i}^{L\dagger} \phi_{i}^{L} + \phi_{i}^{S\dagger} \phi_{i}^{S} \right).$$
(4)

The foregoing equations are valid for both non-periodic and periodic systems. For periodic systems, all matrices and orbitals can be evaluated separately for different crystal momenta k. The electron density can then be generated by integration over the reciprocal space. Unless otherwise stated, the derivations in section II and section III are suitable for both periodic and non-periodic systems, though we are using the non-periodic expression for a briefer presentation. We refer to the nonrelativistic and scalar-relativistic version of *FHI-aims*[28] for more details.

Physically, Eqs. (2a) and (2b) reveal that the small component functions are only of appreciable

size in the spatial regions and/or for states where $R_i(\vec{r})$ is large and in which the gradient of ϕ^L is large (via the momentum operator). For filled states and for reasonably low-energy conduction states, these conditions are met in regions where V_{eff} becomes very negative (attractive) and for states that are very localized, i.e., in regions near the nucleus and/or for core states. In the region near the nucleus, the relatively small potential differences (compared to the electron rest mass energy, mc^2) imposed by chemical environment changes are typically unimportant. This implies that $R_i(\vec{r})$ and thus the relationship between the large and small components are not much affected by the molecular field but, instead, remain rather similar to the relation for the same atom and state in, say, the isolated free atom in vacuum.

B. Four-Component Basis Set Expansion

In relativistic electronic structure theory, the orbitals need to be expanded in a suitable basis set. We here employ a form that is directly inspired by Eq. (2):

$$\phi^L = \sum_{\mu}^{2N} C^L_{\mu n} \chi^L_{\mu}, \qquad (5a)$$

$$\phi^{S} = \sum_{\mu}^{2N} C_{\mu n}^{S} \chi_{\mu}^{S} , \qquad (5b)$$
$$\chi_{\mu}^{S} = \hat{K} \chi_{\mu}^{L} , \qquad (5c)$$

where $\{\chi_{\mu}^{L}(\vec{r})\}\$ is the set of large component basis functions and $\{\chi_{\mu}^{S}(\vec{r})\}\$ is the set of small component basis functions. In Eq. (5), there is precisely one small-component basis function for each large-component basis function. Eq. (5c) confines the choice of the small component basis according to the chosen large component basis. In some implementations, the operator $\hat{K} = \frac{1}{2mc}\vec{\sigma}$.

 \vec{p} , i.e., setting $R_i(\vec{r})$ in Eq. (2b) to its non-relativistic constant limit, i.e., one, a choice known as the kinetically balanced form.[78] This choice, however, usually requires a relatively larger basis set (for convergence) and the large- and small-component coefficients need to be optimized separately in the diagonalization. One can finally obtain the eigenvalues for both positive and negative energy states. In order to retain well-defined precision and to retain limited basis set size, it would be nice if each small component basis function could be mathematically linked directly to each large-component basis function, using a mathematical relationship that is as close to exact as possible for each state, like Eq. (2a). In this case, the large and small components would share the same basis set coefficients $C_{\mu n}^L = C_{\mu n}^S$.

While this exact relation cannot be obtained for all states at once (the factor $R_i(\vec{r})$ in Eq. (2b) is state-dependent through the eigenvalue, whereas the operator \hat{K} in Eq. (5c) is state-independent), it is nevertheless interesting to look at the dependence of $R_i(\vec{r})$ in Eq (2) on the distance from the position to the nucleus. In a previous study[32], the $R_i(\vec{r})$ of several atomic orbitals (1s, 2s, 2p, and 6s) of a Rn atom was discussed and shows an interesting conclusion: the values of $R_i(\vec{r})$ vary rapidly in the close vicinity of the nucleus, but become constants beyond some critical points (say, 0.01 Bohr – roughly the radial expectation values of 2s and 2p of Rn). This is mathematically unsurprising if one looks into Eq. (2b): the radial potential is quite steep near the nucleus but tends to be flat at a larger r, and more importantly, the difference (the numerator) between the eigenvalue (a fixed value for each atomic orbital) and the flat radial potential is nearly a constant compared to the large value of $2mc^2$. This implies that $R_i(\vec{r})$ and thus the relationship between the large and small components are not much affected by the molecular field. As described in more detail below, we here therefore rely on two different choices for the operator \widehat{K} in our approach, depending on the basis functions. (1) For basis functions derived directly from solving the free-atom problem, we will use the exact, state-dependent form $R_i(\vec{r})$ of Eq. (2) directly. Core orbitals in chemically bonded structures remain practically unchanged from their shape in the free atom and so for these largest energy contributions, the choice of a state-dependent $R_i(\vec{r})$ will translate into a practically exact small-component representation. We call this choice of \hat{K} "atomic balance"[79]. (2) Basis functions other than free-atom eigenfunctions largely contribute to the description of the valence electrons. For these basis functions, we use a state-independent form of $R_i(\vec{r})$ in \hat{K} , by using the $V_{\text{eff}}(\mathbf{r})$ for the free atom in the construction of \widehat{K} and setting the eigenvalue $\varepsilon_i = 0$. This choice is similar in spirit to the successful atomic ZORA for scalar relativity [28] but here restricted to only the construction of the small component and, for valence electrons, practically exact near the nucleus. We will come back to (1) and (2) below.

C. Quasi Four-Component (Q4C) Algorithm with Numerical Atomic Basis

As noted in Sec. IIA, Eqs. (2a) and (2b) show that the small component is only of appreciable magnitude very close to a nucleus, the only region in which $V_{\text{eff}}(\mathbf{r})$ becomes numerically relevant compared to the term $2mc^2$ and where the shape of the overall wave function is practically the same as that of a free atom, essentially unaffected by small changes to the chemical environment. Thus, in practice, multi-atom calculations of electron-like states can be carried out with high precision[33, 34, 80-82] while fixing the ratio between the small and large components to the known and precomputed relation between the small and large components in spherical free atoms. This logic is behind the use of atomic-like transformations between the small and large components at the exact two-component (X2C) level, showing comparable precision as traditional four-component methods, for both molecules[32, 83] and solids[23].

The Q4C approach adopts this logic in cases where free-atom-like basis functions (practically exact near the nucleus) are available in the basis set used to expand the wave function in a multiatom electronic structure calculation. Specifically, the relevant molecular four-component spinors ϕ_i (in the Dirac equation) can then be expanded with high precision in a set of atomic fourcomponent spinors that share the same coefficients for small- and large-component basis functions:

$$\phi_i(\vec{r}) = \sum_{\mu}^{N} (\mathbf{C}^+)_{\mu i} \begin{pmatrix} \boldsymbol{\mathcal{X}}_{\mu}^{L} \\ \boldsymbol{\mathcal{X}}_{\mu}^{S} \end{pmatrix}.$$
(6)

Here, the large and small components are linked by the relation in Eq. (2). If the atomic-like basis functions are chosen to be electron-like states associated with free atoms, this constraint physically amounts to projecting the four-component Hamiltonian onto the relativistic positive energy states only.

In the *FHI-aims* code[28] which we here use for our implementation, Eq. (6) is convenient since the code employs NAO basis functions obtained from free atoms (more details will be given in section III A) as the so-called "minimal basis" part of its default basis sets. We use the *dftatom* code[38] to solve the radial equation for spherical free atoms on a dense, high-precision one-dimensional grid, prior to a density-functional calculation for a polyatomic system in a basis set. *dftatom* yields output radial functions for the minimal basis that are already projected onto the positive energy states and thus only describes electrons. Table 1 gives an example of the basis set used for *Si*, with functions included up to the "second tier" (or "level") of basis functions in the terminology used in *FHI-aims*. The notation of basis functions are defined by the same basis-defining potentials as the scalar relativistic basis functions tabulated in *FHI-aims*' standard basis sets, except the Dirac versions are used here.

Basis set level	Atomic orbitals used
Minimal	[Ne] + 2s2p
First tier	$H(2p, 1.4), H(3d, 4.2), H(4f, 6.2), Si^{2+}(3s)$
Second tier	<i>H</i> (1 <i>s</i> , 0.65), <i>H</i> (3 <i>d</i> , 9), <i>H</i> (4 <i>p</i> , 4), <i>H</i> (5 <i>g</i> , 9.4)

Table 1. The "second tier" *FHI-aims* basis set for *Si*. In the Q4C case, the minimal basis corresponds to the free-atom Dirac radial functions (noble gas configuration of the core and quantum numbers of the additional valence radial functions). "H(nl, z)" denotes a hydrogen-like basis function for the bare Coulomb potential z/r, including its radial and angular momentum quantum numbers, *n* and *l*.

 $Si^{2+}(3s)$ denotes a 3s Dirac radial function of a free ion Si^{2+} . The hydrogen-like and ionic basis sets are also generated by solving a radial Dirac equation (more details will be given in section III A).

Substituting Eq. (6) into Eq. (1) leads to a two-component matrix that describes electrons only: $H^+C^+ = S^+C^+E^+$. (7)

The corresponding matrix elements are:

$$\begin{aligned} H_{\mu\nu}^{+} &= \left\langle \chi_{\mu}^{L} \middle| \begin{array}{c} 0 & c\vec{\sigma}\cdot\vec{p} \\ \chi_{\nu}^{S} \middle| c\vec{\sigma}\cdot\vec{p} & -2mc^{2} \middle| \chi_{\nu}^{L} \\ \chi_{\nu}^{S} \right\rangle + \left\langle \chi_{\mu}^{L} \middle| \hat{V} \middle| \chi_{\nu}^{L} \right\rangle + \left\langle \chi_{\mu}^{S} \middle| \hat{V} \middle| \chi_{\nu}^{S} \right\rangle \\ &= \left\langle \chi_{\mu}^{L} \middle| c\vec{\sigma}\cdot\vec{p} \middle| \chi_{\nu}^{S} \right\rangle + \left\langle \chi_{\mu}^{S} \middle| c\vec{\sigma}\cdot\vec{p} \middle| \chi_{\nu}^{L} \right\rangle + \left\langle \chi_{\mu}^{L} \middle| \hat{V} \middle| \chi_{\nu}^{L} \right\rangle + \left\langle \chi_{\mu}^{S} \middle| \hat{V} \middle| \chi_{\nu}^{S} \right\rangle - \left\langle \chi_{\mu}^{S} \middle| 2mc^{2} \middle| \chi_{\nu}^{S} \right\rangle \\ S_{\mu\nu}^{+} &= \left\langle \chi_{\mu}^{L} \middle| \chi_{\nu}^{L} \right\rangle + \left\langle \chi_{\mu}^{S} \middle| \chi_{\nu}^{S} \right\rangle, \qquad (8b) \\ n &= n^{L} + n^{S} = \sum_{i} f_{i} \left(\phi_{i}^{L\dagger} \phi_{i}^{L} + \phi_{i}^{S\dagger} \phi_{i}^{S} \right). \qquad (8c) \end{aligned}$$

 E^+ is the diagonal matrix of eigenvalues. The eigenvalue problem can be solved routinely in parallel, e.g., using the solvers available in the ELSI infrastructure[84, 85].

In Eq. (8c), the four-component electron density is the sum of the large- and small-component densities. However, in practical calculations, one can apply the following strategy to obtain the four-component density instead of updating the small-component density at each self-consistent field (SCF) iteration:

$$n^{SCF} = n^{L,SCF} + \sum_{A} \left(n_A - n_A^L \right). \tag{9}$$

 $n^{L,SCF}$ denotes the density contribution due to the large components only. The sum over *A* runs over all atoms and can be obtained from a superposition of free atom densities n_A and the analogous large-component free-atom densities n^L_A that can be pre-computed. The free-atom density components do not need to be updated in the SCF iterations of a multi-atom computation in a basis set, where only the first term (the large-component density) is updated. As mentioned above, this approach is practically exact in calculations for chemical problems because the small-component density is always highly localized, shaped almost exclusively by the nearby nucleus, and changes very little in the SCF iterations.

D. Reference Methods: X2C Approach in *BDF* and Self-Consistent Spin-Orbit Coupled Augmented Plane Wave Approach in *WIEN2k*

As a comparison based on a completely independent code and implementation, we provide some reference values from the *BDF* code[23, 86, 87], using Slater-type orbital (STO) basis functions and the X2C method [23, 33] for the reference calculations. Since the X2C approach has been described extensively elsewhere, we only summarize the essential differences to the Q4C approach below:

The X2C method defines an X matrix that links the large- and small-component wave functions (viz. $C^S = XC^L$) in the space of basis functions and decouples the positive and negative energy states through a unitary transformation (the corresponding U matrix can be generated using X) that is applied to the Hamiltonian matrix. The key difference between X2C and Q4C is that there is no minimal basis of exact atomic radial functions in X2C. On the positive side, this allows for the use of general basis sets such as STOs, for which appropriate X matrices can still be constructed from exact free-atom solutions. However, unlike the atomic radial solutions in Q4C, the basis functions

in X2C do not allow one to construct a practically exact, known atomic balance (derived from Eq. (2) when it describes an atomic system)[79] for each individual core function from the free-atom limit. In practice, general-shape basis functions such as STOs do not necessarily contribute only to a single core function, so that the eigenvalue in the atomic balance condition to be used would not be unique. Thus, in X2C, one needs the aforementioned matrix transformation, whereas in Q4C we can basically write down the atomic balance directly for the core radial functions.

In the *BDF* reference calculations, uncontracted STO QZ4P basis sets are used. The basis sets in question are summarized in Table S1.

As a second point of reference, widely employed in the solid-state physics community, we utilize the self-consistent spin-orbit coupling approach based on augmented plane wave (APW) basis functions and including $p^{1/2}$ basis functions (called APW+ $p^{1/2}$) as implemented in the *WIEN2k* code[88, 89]. In *WIEN2k*, the atomic core is treated by the full Dirac equation, ensuring that the core states and density, where relativistic effects are strongest, are essentially treated exactly from a relativistic point of view. We note that our present Q4C approach technically covers the same physics for the core states but should provide somewhat higher accuracy for the semicore and valence states, since radial basis functions with the asymptotically correct *l*+1/2 and *l*-1/2 behavior near the nucleus are included for all *l* channels, not just *l*=1 (i.e., *p*-type functions).

III. Implementation Details

In this section, we define specifics for two technical issues: the integration of relativistic matrix elements and the atomic basis set it needs. Other technicalities that are not that pertinent to relativity have been described earlier for the nonrelativistic and scalar relativistic version of *FHI-aims*[28].

A. Relativistic Numeric Atom-Centered Basis

The nonrelativistic and scalar relativistic NAO basis functions can be written in the following form[28, 90-94]:

$$\chi_i(\vec{r}) = \frac{u_i(r)}{r} Y_{lm}(\Omega) . \qquad (10)$$

 $Y_{lm}(\Omega)$ here denotes the real parts (m = 0, ..., l) and imaginary parts (m = -l, ..., -1) of complex spherical harmonics, with l, m being the azimuthal quantum number and magnetic quantum number, respectively. The radial function $u_i(r)$ is numerically tabulated on a logarithmically spaced radial grid as a cubic spline function and is therefore fully flexible in its shape. A logarithmic grid is generally defined as $r(i + 1) = \alpha \cdot r(i)$ for grid points *i*, with r(0) close to the nucleus and α an increasing factor slightly greater than 1. We use $r(0) = 10^{-4}/Z$ Bohr radii, α =1.0123 and a grid that extends out to 100 Bohr radii. As an example, this choice leads to just below 1,500 logarithmic grid points for the heavy element Bi. This dense grid is only used for one-dimensionally constructed and tabulated radial functions, but does not influence production calculations of three-dimensional molecules or solids, in which the basis functions are merely spline-evaluated on a much sparser grid for three-dimensional integrations.[28, 29]

For the relativistic case, we adopt the open source, four-component Dirac-Kohn-Sham atom solver *dftatom*[38] for spherical atoms, on a one-dimensional logarithmic grid. According to the Dirac equation, the large and small components have different radial functions, and the separation of the radial and angular part should be written as

$$\chi(\vec{r}) = \frac{1}{r} \begin{pmatrix} P(r)Y^{L}(\Omega) \\ iQ(r)Y^{S}(\Omega) \end{pmatrix}, \quad (11)$$

in which P and Q are the radial large and small components of the orbital and Y^L and Y^S are the angular large- and small-component parts. The imaginary number *i* keeps the radial component Q real-valued. Substituting this expression into the Dirac equation for a spherical potential, one obtains a pair of coupled radial equations[76]

$$(V_{basis} - \varepsilon)P_{n\kappa} + c \left[\frac{dQ_{n\kappa}}{dr} - \frac{\kappa Q_{n\kappa}}{r}\right] = 0, \quad (12a)$$
$$-c \left[\frac{dP_{n\kappa}}{dr} + \frac{\kappa P_{n\kappa}}{r}\right] + (V_{basis} - \varepsilon - 2mc^2)Q_{n\kappa} = 0. \quad (12b)$$

 V_{basis} here denotes the radial potential of the free atom or atom-like system. *n* is the principal quantum number, and κ is determined by the angular momentum *l* and spin state *s*:

$$\kappa = \begin{cases} -l - 1, \text{ for } j = l + \frac{1}{2}, \text{ i.e. } s = +\frac{1}{2} \\ l, \text{ for } j = l - \frac{1}{2}, \text{ i.e. } s = -\frac{1}{2} \end{cases}$$
(13)

The normalization of $P_{n\kappa}$ and $Q_{n\kappa}$ obeys the condition

$$\int_{0}^{\infty} \left[P_{n\kappa}^{2}(r) + Q_{n\kappa}^{2}(r) \right] dr = 1.$$
 (14)

After calling *dftatom* as a subroutine to solve the Dirac equation on the logarithmic radial grid, the numerically tabulated large- and small-component radial functions of the minimal basis, as well as of the additional basis functions used in this work, are available within *FHI-aims*.

All basis functions are generated using the *dftatom* solver with a confining potential designed to eliminate extended but numerically irrelevant near-zero components far away from the atom to which they correspond. Specifically, V_{basis} in Eq. (12) includes an additional confining potential V_{cut} , following Eq. (9) in reference [28]. This confinement potential is extended enough to avoid a strong shape influence on the resulting radial functions, aiming only to suppress any spatially extended, slow-decaying radial function tails. To ensure a smooth decay of all basis functions and their derivatives to zero, V_{cut} comprises three smooth analytical parts: for $0 \le r \le r_{onset}$, no confining potential is imposed; for $r_{onset} \le r \le r_{cut}$, a smoothly gradually increasing confining potential is applied until r reaches r_{cut} , where the confining potential becomes infinite and the wave functions are exactly zero. We currently adopt the same V_{cut} for both nonrelativistic and relativistic calculations and have found good performance in all tested systems. For most chemical elements, the default "light" settings of *FHI-aims* employ $r_{onset} = 3.5$ Å and a standard width ($r_{cut} - r_{onset} = 1.5$ Å). Tight settings employ ($r_{onset} = 4.0$ Å) and ($r_{cut} - r_{onset} = 2.0$ Å) for all but a few largest atoms.

As noted in Table 1, in addition to the minimal basis set (obtained by solving free atoms), two further types of radial functions are used to create numerically precise basis sets in this work: the radial functions of (1) free ions and of (2) hydrogen-like atoms with a variable nuclear potential – Z/r (with Z in the range of 0.1-20 for light atoms before the 21st element, and up to 60 for the *f*, *g*,

and *h* functions of heavy atoms). For basis functions beyond the minimal basis, the atomic balance condition between large and small components requires some attention. In the construction of each radial basis function, Eq. (12) can be directly applied, but for hydrogen-like functions, V_{basis} now denotes the radial potential of the given hydrogen-like atom (simply -Z/r in this case) instead of the potential V_{eff} of the actual nucleus at this site as defined in Eq. (2). This raises a question regarding the atomic balance condition to be adopted to define the small-component part of the four-component basis functions from a given set of large-component part for a later, self-consistent calculation using the basis set defined for a given redial function, the atomic balance for a given hydrogen-like atom with effective nuclear charge Z. In contrast, in a self-consistent calculation, the same radial function pair would be attached to a given chemical element with a potentially very different nuclear charge Z_{atom} , leading to a very different actual atomic balance as defined in Eq. (2). We will come back to the choice of atomic balance below.

B. Relativistic Integrals

According to Eq. (8) in Section II C and in molecular (non-periodic) systems, the following set of matrix elements is needed for constructing the final secular equation:

$$\left\langle \chi^{L}_{\mu} \middle| \vec{c\sigma} \cdot \vec{p} \middle| \chi^{S}_{\nu} \right\rangle, \left\langle \chi^{S}_{\mu} \middle| \vec{c\sigma} \cdot \vec{p} \middle| \chi^{L}_{\nu} \right\rangle, \left\langle \chi^{L}_{\mu} \middle| \hat{V} \middle| \chi^{L}_{\nu} \right\rangle, \left\langle \chi^{S}_{\mu} \middle| \hat{V} \middle| \chi^{S}_{\nu} \right\rangle, \left\langle \chi^{L}_{\mu} \middle| \chi^{L}_{\nu} \right\rangle, \left\langle \chi^{S}_{\mu} \middle| \chi^{S}_{\nu} \right\rangle.$$

In periodic boundary conditions, the notation changes since the matrix elements are defined for infinitely extended Bloch basis functions, a sum over basis functions localized in each unit cell and with appropriate phase factors (defined by way of reciprocal-space vectors or crystal momenta k in the first Brillouin zone):

$$\varphi_i^X(\vec{r}, \vec{k}) = \sum_N e^{i\vec{k}\cdot\vec{T}_N} \chi_i^X(\vec{r} - \vec{T}_N), X = L, S_{\perp}$$
(15)

As written in Eq. (15), the sum over localized basis functions runs over the entire crystal and is therefore not yet normalizable. However, expressed in matrix elements, all terms can be rephrased in entities that are normalized over the volume of a single unit cell, as outlined in detail in Eqs. (27-31) in reference [95] and in Fig. 1 of that reference. The resulting expressions (analogous to Eq. (31) in reference [95]) for the matrix elements in periodic boundary conditions needed in this work are:

$$T_{\mu\nu}^{L}(\vec{k}) = \sum_{N,N'}^{N_{cell}} \exp\left[i\vec{k}\cdot\left(\vec{T}_{N'}-\vec{T}_{N}\right)\right] \left\langle\chi_{\mu,N}^{L}\left|c\vec{\sigma}\cdot\vec{p}\right|\chi_{\nu,N'}^{S}\right\rangle_{\text{unit cell}}, \quad (16a)$$

$$T_{\mu\nu}^{S}(\vec{k}) = \sum_{N,N'}^{N_{cell}} \exp\left[i\vec{k}\cdot\left(\vec{T}_{N'}-\vec{T}_{N}\right)\right] \left\langle\chi_{\mu,N}^{S}\left|c\vec{\sigma}\cdot\vec{p}\right|\chi_{\nu,N'}^{L}\right\rangle_{\text{unit cell}}, \quad (16b)$$

$$V_{\mu\nu}(\vec{k}) = \sum_{N,N'}^{N} \exp\left[i\vec{k}\cdot\left(\vec{T}_{N'}-\vec{T}_{N}\right)\right] \left\langle \chi_{\mu,N}^{L} \left|\hat{V}\right| \chi_{\nu,N'}^{L} \right\rangle_{\text{unit cell}}, \quad (16c)$$

$$W_{\mu\nu}(\vec{k}) = \sum_{N,N'}^{N_{cell}} \exp\left[i\vec{k}\cdot\left(\vec{T}_{N'}-\vec{T}_{N}\right)\right] \left\langle \chi_{\mu,N}^{S} \left|\hat{V}\right| \chi_{\nu,N'}^{S} \right\rangle_{\text{unit cell}}, \quad (16d)$$

$$S_{\mu\nu}^{L}(\vec{k}) = \sum_{N,N'}^{N_{cell}} \exp\left[i\vec{k}\cdot\left(\vec{T}_{N'}-\vec{T}_{N}\right)\right] \left\langle\chi_{\mu,N}^{L}\left|\chi_{\nu,N'}^{L}\right\rangle_{\text{unit cell}}, \quad (16e)$$
$$S_{\mu\nu}^{S}(\vec{k}) = \sum_{N,N'}^{N_{cell}} \exp\left[i\vec{k}\cdot\left(\vec{T}_{N'}-\vec{T}_{N}\right)\right] \left\langle\chi_{\mu,N}^{S}\left|\chi_{\nu,N'}^{S}\right\rangle_{\text{unit cell}}. \quad (16f)$$

Here, the real-space basis functions $\chi^X_{\mu,N}$ and $\chi^X_{\nu,N'}$ (*X*=*L*,*S*) are centered in different real-space unit cells associated with translation vectors \vec{T}_N and $\vec{T}_{N'}$. The integrals denoted <...>_{unit cell} are performed only over the volume of only a single unit cell (e.g., the unit cell labeled "0" in real space). The sums over unit cells *N* and *N* are now restricted only to those *N*_{cells} unit cells in which localized basis functions are centered that touch the specific unit cell selected as the integration volume. In mathematical terms and for a given operator \hat{F} , the respective integrals read[28, 96]:

$$\left\langle \chi_{\mu,N}^{X} \left| \hat{F} \right| \chi_{\nu,N'}^{X} \right\rangle_{\text{unit cell}} = \int_{\text{unit cell}} \chi_{\mu,N}^{X}(\vec{r}) \hat{F} \chi_{\nu,N'}^{X}(\vec{r}) d^{3}r, \quad X = L, S_{\perp}$$
(17)

This is the same procedure as used in the scalar-relativistic formalism and visualized in Fig. 1 of reference [95]. For further implementation details, see also reference [28] and reference [96]. The only difference here is that the basis functions are relativistic large- and small-component spinors.

Technically, the expressions for the overlap and potential matrices seem straightforward, but in practice and in a naive approach to integration, significant precision issues can arise due to the Coulomb singularity and associated singularity of Dirac radial functions[38, 76] at the nucleus of each atom. The singularity of the radial functions could be avoided by introducing higher-order quantum-electrodynamic terms[76], an approach not pursued in the present work. Another way of alleviating this issue is to introduce a finite nucleus model (see reference [97] and Chapter 5.4 of reference [77]) and therefore smoothen the radial potential near to the nucleus, although we did not pursue this avenue in the present work.

Instead, our approach is based upon exploiting the specific shape of the free-atom like NAO radial functions that account for the bulk of the electron density in the relativistic regions near the nuclei also in chemically bonded structures, and that are already accurately reflected in the single free-atom solutions available to us from the *dftatom* code. For free-atom like orbitals and following the Dirac equation Eq. (1), we have:

$$\vec{c\sigma} \cdot \vec{p} |\chi_{\mu}^{S}\rangle = (\varepsilon_{\mu} - V_{atom}) |\chi_{\mu}^{L}\rangle, \quad (18a)$$
$$\vec{c\sigma} \cdot \vec{p} |\chi_{\mu}^{L}\rangle = (\varepsilon_{\mu} - V_{atom} + 2mc^{2}) |\chi_{\mu}^{S}\rangle. \quad (18b)$$

Here, μ denotes the atomic orbitals. For an SCF converged atom, the radial potential and eigenvalues are fixed values and therefore can be obtained precisely on a set of dense logarithmic grids. Therefore, the right parts of the two equations can be generated precisely and can then be saved as the large- and small-component kinetic energy basis functions on the left side (no need for explicit differentiation). When performing the real space integrations on the sparser grids for three-dimensional integration (for molecules/solids), the functions in Eq. (18) can be directly used for the kinetic energy matrix integrations in Eqs. (16a) and (16b). For hydrogen-like functions, the hydrogen like eigenvalues and potential are instead used in Eq. (18), and analogous for ionic radial functions.

Next, if we were to directly integrate Eqs. (16c) and (16d) on an affordable, three-dimensional integration grid appropriate for a multi-atom calculation, we would incur a large numerical inaccuracy due to the finite integration grid density and the very steep potential near the nucleus. In the literature, this problem is sometimes alleviated by the generalized transition state (GTS) method[98], in which the energy difference (instead of the total energy) between the real system and a model system (e.g., the superposition of neutral, spherical and unpolarized atoms) is calculated.[86, 87] Instead, we here adopt a much simpler strategy. If one combines the local term of the matrix integration Eq. (16a) with that of Eq. (16c) and uses Eq. (18a), the following sum of integrals can be obtained:

$$T_{\mu\nu}^{L} + V_{\mu\nu} = \left\langle \chi_{\mu,N}^{L} \left| \varepsilon_{\nu} - V_{atom} + V_{eff} \left| \chi_{\nu,N'}^{L} \right\rangle \right\rangle.$$
(19a)

Similarly, we may combine Eqs. (16b) and (16d) and use Eq. (18b):

$$T_{\mu\nu}^{S} + W_{\mu\nu} = \left\langle \chi_{\mu,N}^{S} \middle| \varepsilon_{\nu} - V_{atom} + 2mc^{2} + V_{eff} \middle| \chi_{\nu,N'}^{S} \right\rangle.$$
(19b)

Here, V_{atom} denotes the radial potential of the corresponding free atom, and V_{eff} is the potential in the molecular/solid potential field. Both of them are steep functions near the nuclei, with essentially the same shape. Their steep, peaked parts therefore cancel, yielding high precision for the final total energy of the unit cell or molecule, compared to a direct integration of the heavily peaked potential and kinetic energy contributions on their own. Note that ε_i is calculated on the dense logarithmic grid and therefore numerically precise.

Lastly, although not needed if Eqs. (19a) and (19b) are used, we might still want to calculate the two kinetic energy matrices T^L and T^S , which are defined as the large- and small-component kinetic energy matrices, respectively. Because the direct integration over the $\vec{\sigma} \cdot \vec{p}$ operator is usually a thorny issue, we provide a direct integration method for the $\vec{\sigma} \cdot \vec{p}$ in Appendix (B).

C. Numerical Construction of Radial Functions and Atomic Balance Condition

In our implementation of the above real-space integrals, we typically need four types of precomputed functions: the large component part χ^L_{μ} (of the four-component basis), the small component part χ^S_{μ} (of the four-component basis), $c\vec{\sigma}\cdot\vec{p}\chi^S_{\mu}$, and $c\vec{\sigma}\cdot\vec{p}\chi^L_{\mu}$ (see Eqs. (18) and (19)). The procedure to generate them is as follows:

1. Technically, $\{\chi_{\mu}^{L}\}$ can be defined at the outset. This set can in principle accommodate any type of radial function, including: The minimal basis functions associated with free atoms, Slater type orbitals (STOs), Gaussian type orbitals (GTOs), or (in the standard NAO basis sets associated with the *FHI-aims* code) the hydrogen-like and free-ionic radial functions obtained by solving

separate radial Dirac equations as described in Section IIIA. 2. To generate the corresponding small component part χ^{S}_{μ} , we use Eq. (2) but we replace the

basis-generating potential V_{eff} with the radial potential of the free atom with the same nucleus to which each basis function is attached in the multi-atom calculation. For radial functions other than the minimal basis, ε_i is set to 0 since, in principle, such radial functions could contribute to more than one eigenstate of the multi-atom calculation. This approach assumes that the radial functions beyond the minimal basis set do not contribute significantly to core states in the actual, extended multi-atom calculation, for which the eigenvalues could be appreciable in magnitude compared to mc^2 . In contrast, for valence and semicore-like states, ε_i is negligibly small compared to mc^2 in the denominator of Eq. (2b). Eventually, for any basis function other than the free-atom like minimal basis, we use the following atomic balance condition to define the small component part (of the four-component basis function):

$$\chi^{s}_{\mu}(\vec{r}) = \frac{1}{2mc^{2} - V_{atom}(r)} c\vec{\sigma} \cdot \vec{p} \chi^{L}_{\mu}(\vec{r}). \qquad (20)$$

3. Since χ_{μ}^{L} is exact, the corresponding $c\vec{\sigma}\cdot\vec{p}\chi_{\mu}^{L}$ can be easily generated by applying $c\vec{\sigma}\cdot\vec{p}$ onto χ_{μ}^{L} . See Eqs. (18a) and (18b) for radial functions derived from their own Dirac-like form. Appendix (B) shows how to deal with the $\vec{\sigma}\cdot\vec{p}$ operator for fixed-form radial functions such as STOs and GTOs.

4. Similarly, with Eq. (20) at hand, we can generate the $c\vec{\sigma}\cdot\vec{p}\chi^{s}_{\mu}$ by operating $c\vec{\sigma}\cdot\vec{p}$ onto χ^{s}_{μ} . The mathematical details are provided in Appendix (C).



Fig. 1. Comparison of the LDA-VWN 1s radial functions associated with a free Xe atom (Z=54) and a hydrogen-like atom with Z=54, for two different choices of the effective potential used to

define the atomic balance (AB) condition that links the small component to the large component. The black curve denotes the radial 1s orbital of the free Xe atom; the red curve denotes the radial 1s orbital of the hydrogen-like atom, with the atomic balance condition in Eq. (2), i.e., including the 1s eigenvalue of the radial function. For comparison, the blue curve denotes the radial 1s orbital of the hydrogen-like atom but with the atomic balance condition in Eq. (20), in which the radial potential V comes from the free Xe atom instead of the hydrogen-like atom itself and in which the 1s eigenvalue is not included in the denominator. For (A) χ^L_{μ} and (D) $c\vec{\sigma} \cdot \vec{p}\chi^L_{\mu}$, there is no difference between the two hydrogen-like functions since no atomic balance condition is needed in generating these functions. For (B) χ^S_{μ} and (C) $c\vec{\sigma} \cdot \vec{p}\chi^S_{\mu}$, the results for the two different

atomic balance conditions agree closely.

To test the impact of the different atomic balance conditions in Eq. (2) vs. Eq. (20), we compare the basis function components associated with the 1s orbital of Xe (Z=54) and the 1s orbital of a hydrogen-like atom with Z=54. Since these are very deep core orbitals, largely unaffected by screening effects of the other electrons, the radial potentials generating these basis functions are very similar. However, for the hydrogen-like function, we know the analytically exact shape and 1s

eigenvalue (-1519.4736 *Ha*). In short, the corresponding χ^L_{μ} , χ^S_{μ} , $\vec{c\sigma} \cdot \vec{p} \chi^S_{\mu}$, and $\vec{c\sigma} \cdot \vec{p} \chi^L_{\mu}$

functions of the 1*s* orbitals are plotted and compared in Fig. 1, showing close agreement between all parts and thus indicate a negligible impact of the approximate atomic balance condition used in Eq. (20).

	Hydrogen-like KB	Free atom KB
Energy for Au atom (Ha)	-19037.5712418	-19037.5712897
Energy for Au ₂ (Ha)	-38075.2580525	-38075.2581532
Binding energy (Ha)	0.1155689	0.1155738

Table 2. LDA-VWN total energies and binding energies (unit: *Hartree*) for Au atom and Au_2 with the two different atomic balance conditions shown in Eqs. (2) and (20). The binding distance of the Au_2 dimer is 2.447Å.

We also see that for a gold dimer (Au_2) testcase, the Vosko-Wilk-Nusair[99] (VWN) localdensity approximation (LDA) binding energy difference (using a tier 2 basis set, see Table S2 in the Appendix) between the two atomic balance conditions is ~0.00013 eV (see Table 2 for the energy values). This confirms that the two different atomic balance conditions impact the valence orbitals very little. An exemplary hydrogen-like basis function H(4f, Z=7.4) used as a part of the Au basis set, is plotted in Fig. 2, showing negligible difference between the radial wavefunctions (black and red curves) generated with the two atomic balance conditions. In fact, from our tests (on all the compounds reported in section IV), the difference in the SCF converged eigenvalues is less than ~0.00001 eV for valence orbitals and is less than ~0.01 eV for the innermost orbital (viz. 1s).



Fig. 2. The corresponding (A) χ_{μ}^{L} , (B) χ_{μ}^{S} , (C) $c\vec{\sigma} \cdot \vec{p}\chi_{\mu}^{S}$, and (D) $c\vec{\sigma} \cdot \vec{p}\chi_{\mu}^{L}$ functions of the hydrogen-like basis function H(4f, Z=7.4) for Au, which is a valence-type basis function found in *FHI-aims*' standard NAO basis sets for Au. The black curve denotes the radial orbital of the hydrogen-like atom, with the atomic balance (AB) condition in Eq. (2); the red curve denotes the radial 1s orbital of the hydrogen-like atom, with the atom instead of the hydrogen-like atom itself. The curves are visually identical. One key point to note is that the valence basis functions shown here are non-zero only much further away from the nucleus than the core functions shown in Fig, 1. They are therefore not numerically affected by the region in which the atomic-like potential is deep enough to have a significant numerical effect via the exact atomic balance condition.

IV. Results

A. Diatomic Molecules

We first investigate several typical diatomic molecules and report their relativistic molecular properties, viz. the optimized bond lengths and vibrational frequencies. These dimers (i.e. F_2 , Cl_2 , Br_2 , I_2 , At_2 , AgH, AuH, Ag_2 , Au_2 , Pb_2 , Bi_2 , and Po_2) have been studies in the past at various levels of theory such as relativistic Hartree-Fock, coupled cluster theory, Møller-Plesset perturbation theory, pseudopotential, Dirac Kohn Sham theory.[10-12, 83, 100-108] Here, we report the NR, SR, and Q4C results from *FHI-aims*[4, 28], and the X2C results from *BDF*[86, 87]. The basis sets used are

the tier 2 NAO basis functions in *FHI-aims* with really tight setting and the STO QZ4P basis functions in *BDF*. All calculations are done using the VWN LDA functional[99] so as to be comparable to previous studies.

To optimize the bond length of each molecule, we use seven points around the equilibrium with an equidistant separation of 0.05 Å. We perform a 6th order polynomial fitting (the fitted curve is checked visually to rule out any unexpected oscillations due to the polynomial fitting technique) for these chosen bond lengths and the corresponding energies and obtain the optimized bond length at the energy minimum. As for the frequency calculations for the molecular ground state, the molecular oscillation can be taken as a harmonic oscillation and is thus subjected to Hooke's law. Specifically, a parabola is fitted using all available energy points.

The Q4C (X2C) relativistic LDA bond lengths, vibrational frequencies and binding energies calculated by FHI-aims (BDF) are plotted in Fig. 3, with more detailed data listed in Table S3. Good agreement between the Q4C results of FHI-aims and the X2C results of BDF can be observed for F₂, Cl₂, Br₂, I₂, At₂, AgH, AuH, Ag₂, and Au₂. Fig. 4 and Tables S4-S15 compare the same quantities for different levels of relativity available in FHI-aims. For the scalar-relativistic (SR) calculations, atomic ZORA as defined in reference [28] is used. For the light-element dihalogens F_2 , Cl_2 , Br_2 , and I_2 , relativity only has a minor influence on bond lengths (< 0.01 Å) and frequencies (< 4%). For these two observables, even the nonrelativistic calculations can give good predictions, which is in accordance with Visscher and Dyall's HF, MP2, CISD, CCSD, and CCSD(T) results.[105] For binding energies, the difference between non- or scalar-relativistic values on the one hand and Q4C relativistic values on the other hand is still low for F_2 and Cl_2 , but becomes more pronounced for Br_2 and I_2 . This is not unexpected since the primary difference between scalar and Q4C relativity is the self-consistent inclusion of spin-orbit coupling terms, which are especially pronounced for the p orbitals that serve as the primary bonding orbitals in the dihalogens. Here, the SOC splittings between 4(Br) and 5(I) $p^{1/2}$ and $p^{3/2}$ valence orbitals are quite significant and, correspondingly, the bonding orbitals in the free atoms and dimers inherit different orderings in the scalar- vs. fullyrelativistic cases. The magnitude of the effect of Q4C, compared to scalar relativity, on dimer binding energies as light as Br_2 and I_2 is noteworthy, given that many practical computations of total energies of main-group element containing molecules and solids in this atomic number range neglect effects of spin-orbit coupling in practice. For the much heavier halide dimer At_2 , both scalar relativistic effects and SOC effect become significant.

Turning next to Ag_2 , Au_2 , AgH, and AuH, there is a much more pronounced difference between the non-relativistic results on the one hand and scalar- or Q4C-relativistic results on the other. In contrast, for these four systems, we see that the SOC effect is in fact not that pronounced; therefore, the SR calculations already provide reasonable descriptions. This observation is attributed to the fact that the primary bonding orbitals have *s* character, which are particularly affected by scalarrelativistic orbital contraction effects but which are not themselves much affected by spin-orbit splitting.[40]



Fig. 3. Q4C and X2C (**A**) relativistic optimized LDA bond lengths (in \mathring{A}), (**B**) predicted vibrational frequencies (in cm^{-1}) and (**C**) binding energies (in eV) of several diatomic molecules. *FHI-aims* calculations (really tight setting and tier 2 basis set) used the Q4C approach whereas the X2C approach was used in *BDF*.



Fig. 4. (A) optimized bond lengths (in \mathring{A}), **(B)** vibrational frequencies (in cm^{-1}), and **(C)** binding energies (eV) of several diatomic molecules at the nonrelativistic, scalar relativistic, and Q4C levels of theory, respectively. The *FHI-aims* really tight setting and tier 2 basis set were used.

Finally, the heavy 6p main group metal atom dimers Pb_2 , Bi_2 , and Po_2 show very pronounced effects of full relativity compared to non- or scalar-relativistic treatments, especially on the binding energies. Again, this is due to the fact that the primary bonding orbitals are *p*-derived, showing the qualitatively strongest spin-orbit splittings. Additionally, we also note that this group of dimers is known to be heavily affected by multireference effects that are not well captured by the simple density functional applied in the present work. More sophisticated correlation treatments should be adopted for dimers in this row of the periodic table (e.g., reference [107] for the Tl₂ dimer, which is not included in Figs. 3 and 4 and for which the *BDF* and *FHI-aims* implementations compared here lead to different self-consistent single-reference states). Here, our focus rests on elucidating the effects of relativity in the kinetic energy operator in the presence of an otherwise consistent, but not necessarily exact correlation term. Thus, we simply provide a qualitative NR, SR, and FR outcome within the DFT framework for reference, which is expected to still reflect the effects of relativity in a qualitatively correct fashion. This is confirmed by comparing to a recent CAS-SCF study of heavy dimers including *Pb*₂ and *PbPo* by Martens *et al*[109], which shows the same significant binding energy reduction upon including the SOC effects as that observed by us for *Pb*₂, *Bi*₂, and *Po*₂.

B. Benchmark Results for Silver Halide Systems

To test the validity of the implemented code for extended systems, we perform calculations on a series of model systems (i.e. AgCl, AgBr, and AgI in the rocksalt structure) and compare the results to previous benchmark studies by some of us[23]. Nonrelativistic, scalar relativistic, and Q4C calculations were conducted with both LDA[99] and PBE[110] functionals. The reported properties include: equilibrium lattice constant, bulk modulus, and predicted Kohn-Sham band gap. The "really tight" setting for precision and the "second tier" setting for the basis set of *FHI-aims* were adopted for the basis set and real space integrations; as for the reciprocal space integration, $7 \times 7 \times 7 k$ points (with respect to the 2-atom Bravais cell) are used for each system. The equilibrium lattice constants and bulk moduli are obtained from the Birch-Murnaghan equation of state[111], fitted to seven points around the minimum-energy lattice parameter with an equidistant separation of 0.05 Å.

Silver halides are systems containing elements of intermediate atomic number. They are chosen as test systems here and in past work[23]) since literature data on relativistic effects are already available for comparison.[23, 112-117] The predicted structural properties of AgCl, AgBr, and AgI are listed in Tables 3-5 and are compared with X2C results from BDF[33] and with APW + $p^{1/2}$ self-consistent SOC results from the WIEN2k code[88, 89].

As has long been known, these systems show indirect band gaps with the valence band maximum (VBM) locates at L point (of the reciprocal space) and the conduction band minimum (CBM) locates at the Γ point for AgCl and AgBr but at the X point for AgI. This is mainly due to the strong p-d hybridization between the 4d orbital of Ag and the p-valence orbitals of the halogens: for all these silver halides, the Ag 4d orbitals in the energy region of the top of the valence bands are nearly degenerate with the p-valence orbitals of the halogens, and they are therefore intimately mixed, leading to some complexity and large band widths in the valence bands.[118] The indirect band gap character can be seen from Tables 3-5, which are in agreement with other theoretical and experimental results.

Functional	Method	а	B_0	L-L	Г-Г	X-X	L-F
LDA	NR	5.419	62.7	4.07	3.74	5.16	1.39

		(0.000)	(-0.8)	(-0.02)	(-0.02)	(0.00)	(-0.02)
		[-0.001]	[-0.3]	[-0.01]	[-0.01]	[0.00]	[0.00]
	SR	5.354	69.2	3.66	3.39	4.09	0.64
		(0.001)	(-1.3)	(-0.03)	(-0.02)	(-0.01)	(-0.02)
		[0.002]	[-0.1]	[0.00]	[-0.02]	[0.00]	[-0.01]
	Q4C	5.351	70.0	3.59	3.33	3.91	0.57
		(0.003)	(-2.0)	(-0.01)	(-0.01)	(0.00)	(0.01)
		[0.006]	[-0.6]	[0.03]	[-0.01]	[0.01]	[0.03]
PBE	NR	5.691	37.4	4.74	3.45	5.29	1.68
		(0.001)	(-0.4)	(-0.02)	(-0.01)	(0.00)	(-0.01)
		[0.009]	[-2.5]	[0.02]	[-0.01]	[0.00]	[0.01]
	SR	5.613	41.8	4.34	3.10	4.24	0.93
		(0.000)	(-0.4)	(-0.03)	(-0.01)	(-0.01)	(-0.01)
		[-0.004]	[-2.5]	[0.00]	[-0.01]	[0.00]	[0.00]
	Q4C	5.606	43.5	4.27	2.99	4.03	0.86
		(0.006)	(-2.0)	(0.00)	(0.00)	(0.00)	(0.02)
		[0.002]	[-3.8]	[0.03]	[0.03]	[0.01]	[0.03]
Expt.		5.550	51.4		5.2		3.0

Table 3. Minimum-energy lattice parameters a (in \mathring{A}), bulk moduli B_0 (in GPa), direct energy gaps $(L-L, \Gamma-\Gamma, \text{ and } X-X, \text{ in } eV)$, and indirect energy gaps $(L-\Gamma, \text{ in } eV)$ of AgCl. All the results are obtained using the nonrelativistic (NR), scalar relativistic (SR), and Q4C methods with the *FHI-aims* code and basis sets. The differences between the results from BDF and the present ones are listed in parentheses; the differences between the results of WIEN2k (as obtained from reference [23]) and the present ones are given in square brackets. The experiment values of equilibrium lattice constant[119], bulk modulus[120], Γ - Γ direct energy gap[118], and L- Γ indirect energy gap[121] are listed for reference only.

Functional	Method	а	B_0	L-L	Г-Г	X-X	L-F
LDA	NR	5.659	56.5	3.73	3.21	4.74	1.32
		(0.002)	(-0.4)	(-0.03)	(-0.01)	(0.00)	(-0.01)
		[0.001]	[-0.2]	[-0.02]	[-0.01]	[0.00]	[0.00]
	SR	5.584	63.1	3.28	2.63	3.74	0.39
		(0.002)	(-0.5)	(-0.04)	(-0.01)	(-0.01)	(-0.02)
		[0.000]	[-0.2]	[-0.02]	[-0.01]	[0.00]	[-0.01]
	Q4C	5.583	63.1	3.14	2.43	3.56	0.26
		(0.003)	(-0.6)	(0.01)	(0.00)	(0.00)	(0.03)
		[0.004]	[0.3]	[0.02]	[-0.01]	[0.01]	[0.04]
PBE	NR	5.934	33.9	4.34	2.97	4.81	1.58
		(0.003)	(-0.2)	(-0.03)	(0.00)	(0.00)	(-0.01)
		[0.010]	[1.5]	[0.01]	[-0.01]	[-0.02]	[0.00]
	SR	5.842	38.6	3.90	2.44	3.87	0.69
		(0.001)	(-0.2)	(-0.03)	(-0.01)	(0.00)	(-0.01)
		[0.007]	[-0.9]	[0.00]	[-0.02]	[0.00]	[0.00]
	Q4C	5.845	37.7	3.76	2.26	3.67	0.57

	(-0.002)	(1.0)	(0.01)	(-0.01)	(0.00)	(0.03)	
	[0.002]	(-0.6]	[0.04]	[-0.02]	[0.01]	[0.04]	
Expt.	5.774			4.3		2.5	

Table 4. Minimum-energy lattice parameters a (in \mathring{A}), bulk moduli B_0 (in GPa), direct energy gaps $(L-L, \Gamma-\Gamma, \text{ and } X-X, \text{ in } eV)$, and indirect energy gaps $(L-\Gamma, \text{ in } eV)$ of AgBr. All the results are obtained using the nonrelativistic (NR), scalar relativistic (SR), and Q4C methods with the *FHI-aims* code and basis sets. The differences between the results from BDF and the present ones are listed in parentheses; the differences between the results of *WIEN2k* (as obtained from reference [23]) and the present ones are given in square brackets. The experiment values of equilibrium lattice constant[122], Γ - Γ direct energy gap[123], and L- Γ indirect energy gap[121] are listed for reference only.

Functional	Method	а	B_0	L-L	Г-Г	X-X	L-X
LDA	NR	5.993	49.2	3.41	3.53	3.66	1.06
		(-0.005)	(-0.3)	(-0.03)	(-0.01)	(0.00)	(0.00)
		[-0.001]	[-0.2]	[-0.02]	[0.01]	[0.01]	[-0.01]
	SR	5.901	55.5	2.90	2.56	3.14	0.16
		(0.005)	(-1.3)	(-0.02)	(-0.03)	(-0.01)	(0.01)
		[-0.001]	[-0.2]	[-0.01]	[-0.02]	[0.00]	[-0.01]
	Q4C	5.911	56.4	2.61	2.16	2.83	-0.13
		(-0.010)	(-2.8)	(0.01)	(0.01)	(0.01)	(0.05)
		[-0.006]	[-1.1]	[0.03]	[-0.01]	[-0.01]	[0.04]
PBE	NR	6.275	29.6	3.94	3.15	3.56	1.60
		(0.005)	(-0.2)	(-0.03)	(-0.01)	(0.00)	(0.00)
		[0.006]	[-1.3]	[-0.02]	[-0.02]	[-0.02]	[-0.02]
	SR	6.161	34.6	3.45	2.26	3.05	0.74
		(0.004)	(-0.2)	(-0.03)	(0.01)	(0.02)	(0.00)
		[-0.001]	[-0.8]	[-0.02]	[0.00]	[-0.01]	[-0.03]
	Q4C	6.179	32.5	3.16	1.91	2.75	0.46
		(-0.010)	(1.4)	(0.01)	(-0.01)	(0.01)	(0.03)
		[-0.018]	[1.3]	[0.02]	[0.00]	[0.00]	[0.01]
Expt.		6.067	50.0				

Table 5. Minimum-energy lattice parameters a (in \mathring{A}), bulk moduli B_0 (in GPa), direct energy gaps $(L-L, \Gamma-\Gamma, \text{ and } X-X, \text{ in } eV)$, and indirect energy gaps (L-X, in eV) of AgI. All the results are obtained using the nonrelativistic (NR), scalar relativistic (SR), and Q4C methods with the *FHI-aims* code and basis sets. The differences between the results from BDF and the present ones are listed in parentheses; the differences between the results of *WIEN2k* (as obtained from reference [23]) and the present ones are given in square brackets. The experiment values of equilibrium lattice constant[124] and bulk modulus[125] are listed for reference only.

As can be seen from Tables 3-5, all properties reported here match well with *BDF* and *WIEN2k*'s results at all relativistic levels: for scalar relativistic calculations, *FHI-aims* uses the specific version of the atomic ZORA approximation as defined in reference [28], which is in principle at the same level with the *BDF* X1C[23] treatment; the Q4C algorithm is considered to be

equivalent to the *BDF* X2C algorithm. Likewise, the "scalar relativity plus self-consistent SOC, with additional $p^{1/2}$ basis functions" approach available in *WIEN2k* shows good performance for the present fifth row elements.

Reference [23] has already provided detailed analyses for the test systems used here. One issue to comment on is that the *BDF* results are based on a series of specially designed QZ4P Slater type orbital (STO) basis set denoted as rQZ4P. The reason for the modifications is that such large basis sets (QZ4P), designed originally for atoms and molecules, can result in severe linear dependence for condensed-phase periodic systems. To avoid this, the most diffuse valence s and p functions were deleted from the QZ4P basis functions used in BDF. Analogous findings were also reported for Gaussian type orbitals (GTOs) by Kadek et al. [22] The linear scaling performance of these codes also benefits from removing the diffuse functions, as the real space integration (Eq. (17)) in this case will involve fewer cells: the numbers N and N' are significantly reduced. Still, the extent of the remaining STO-rQZ4P functions can be ~15 Å for medium heavy elements like Ag and I; as for GTOs, this value is reported to be ~ 12 Å for the DZ basis if not significantly truncated[22], and a higher-quality basis will bring in a much larger radius. In this respect, the present Q4C theory benefits from the confined NAOs with an extension of only 6 Å for Ag, Cl, Br, and I while showing comparable precision. We note that the basis set parameters used in FHI-aims, which correspond to high precision settings, did not need any modification to conduct the present work. It is also important to re-emphasize that the confinement applied to the NAO radial functions (see Sec. IIIA, with overall radial extents to 6 Å or larger) is chosen large enough to only remove mathematically irrelevant tails, ensuring that the numerical convergence of total energies and other quantities is not impacted in any significant way (see reference [28] for a more detailed discussion, including a visualization of radial function tails and their magnitude).

C. Energy Band Structure Benchmarks for 103 Materials: Comparison Between the Nonrelativistic (NR), Scalar Relativistic (SR), Scalar Relativistic with Spin-Orbit Correction (SR+SOC) and Quasi Four-Component (Q4C) Methods

In addition to the Q4C method described in the present paper, *FHI-aims* also offers atomic ZORA[28] as a scalar relativistic approach and a non-self-consistent SOC correction method for energy band structures based on an initial self-consistent scalar-relativistic set of eigenstates (called SR+SOC in the following). This non-self-consistent approach avoids computationally expensive operations such as additional density and (for hybrid density functionals) exchange matrix evaluations that would be required in a self-consistent treatment. In a previous study of energy band structures[4] the accuracy of the non-self-consistent SOC scheme for band structures was shown to be essentially exact for elements up to Cd (Z=48), very reasonable for 5p and 5d elements (deviations to the APW+ $p^{1/2}$ scheme within about 0.1 eV) and qualitatively correct even for 6p elements including Pb, Bi, and Po.

For the Q4C approach, we here provide a thorough comparison to the NR, SR, and SR+SOC treatments of the band structures of the same 103 materials considered in reference [4], combining chemical elements from lithium to polonium in the periodic table. This set of materials covers 10 typical crystal structures: face-centered cubic (FCC), body-centered cubic (BCC), simple cubic (SCC), hexagonal close-packed (HCP), graphite (GRA), diamond (DIA), cubic zincblende (ZB), wurtzite (WUR), rocksalt (RS), and cesium chloride (CSCL). They are divided into three subsets: compound semiconductors (37 materials), elemental materials (45 materials), and alkali halides (21

materials), as listed below in Table 6.

Family	Materials
Compound	C (DIA), MgO (RS), AlN (WUR), AlN (ZB), SiC (ZB), BP (ZB), AlP
semiconductors	(ZB), MgS (RS), ZnO (WUR), ZnS (WUR), ZnS (ZB), GaN (WUR),
	GaN (ZB), GaP (ZB), BAs (ZB), AlAs (ZB), GaAs (ZB), MgSe (RS),
	ZnSe (ZB), CdS (WUR), CdS (ZB), CdSe (WUR), CdSe (ZB), InN
	(WUR), InP (ZB), InAs (ZB), AlSb (ZB), GaSb (ZB), InSb (ZB), ZnTe
	(ZB), CdTe (ZB), HgS (ZB), HgSe (ZB), HgTe (ZB), PbS (RS), PbSe
	(RS), and $PbTe$ (RS)
Elemental	Be (HCP), C (GRA), Ne (FCC), Mg (HCP), Al (FCC), Si (DIA), Ca
materials	(FCC), Sc (HCP), Ti (HCP), V (BCC), Cr (BCC), Mn (FCC), Fe (BCC),
	Co (HCP), Ni (FCC), Cu (FCC), Zn (HCP), Ge (DIA), Sr (FCC), Y
	(HCP), Zr (HCP), Nb (BCC), Mo (BCC), Tc (HCP), Ru (HCP), Rh
	(FCC), Pd (FCC), Ag (FCC), Cd (HCP), Sn (DIA), Xe (FCC), Ba
	(BCC), Lu (HCP), Hf (HCP), Ta (BCC), W (BCC), Re (HCP), Os
	(HCP), Ir (FCC), Pt (FCC), Au (FCC), Tl (HCP), Pb (FCC), Bi (BCC),
	and Po (SCC)
Alkali halides	LiF (RS), NaF (RS), LiCl (RS), NaCl (RS), KF (RS), KCl (RS), LiBr
	(RS), NaBr (RS), KBr (RS), RbF (RS), RbCl (RS), RbBr (RS), Lil (RS),
	NaI (RS), KI (RS), RbI (RS), CsF (RS), CsCl (CSCL), CsCl (RS), CsBr
	(CSCL), and <i>CsI</i> (CSCL)

Table 6. Materials used in this study, grouped by type.

For the band structure benchmark, we use the same computational settings for tightly converged computations using PBE functional[110] as in reference [4]. The supplementary material of reference [4] includes detailed computational settings regarding the lattice parameters[126-128], k grids, basis set, etc.

For the 37 compound semiconductors, 21 alkali halides, and 4 nonmetal elemental materials (*Ne*, *Si*, *Ge*, *Xe*) which show observable band gaps, we summarize their NR, SR, SR+SOC, and Q4C predicted PBE band gaps in Tables S16, S17, and S18, respectively. For the 41 metal systems without band gaps, we present in Table S19 the amplitude of the same SOC splittings as reference [4] at selected *k* points, so as to assess the gradually increasing SOC effect.

The band gap is one of the most widely studied properties of materials. Therefore, we first display in Fig. 5 the summary of the band gap results calculated at different levels of theory with PBE functional:

- 1. Nonrelativistic (NR) results calculated by FHI-aims.
- 2. Scalar relativistic (SR) results calculated by FHI-aims.
- 3. The non-self-consistent SOC corrected results (SR+SOC n.s.c) calculated by FHI-aims.
- 4. The self-consistent SOC corrected results with (L)APW+lo basis set including the Dirac $p^{1/2}$ orbitals (APW+ $p^{1/2}$ s.c.), calculated by *WIEN2k*.
- The Q4C results calculated by *FHI-aims*.
 From the band gaps presented in Fig. 5, we see that the self-consistent SOC corrected results

(APW+ $p^{1/2}$, *WIEN2k*, yellow dots) match the Q4C results very well, with deviations less than 0.03 eV. The non-self-consistent SOC correction method (the red dots) can actually provide qualitatively excellent results even for very heavy elements, in line with the previous study.[4] The SOC effect typically reduces band gaps due to the splitting of valence orbitals. But this is not always true for scalar relativistic effects: for certain systems (e.g. *PbTe* (RS) and *CsCl* (CSCL)), it is interesting to see that scalar relativistic effects enlarge the band gaps while SOC reduces them. Therefore, the nonrelativistic calculations sometimes provide seemingly acceptable band gap results in a few cases due to error cancellation.



Fig. 5. Deviations of the NR, SR, SR+SOC n.s.c., and APW+ $p^{1/2}$ s.c. band gaps from the Q4C band gaps. The **(A)** 37 semiconductor compounds and **(B)** 21 alkali halide compounds listed on the horizontal axis are ordered according to their maximum atomic number Z_{max} .

To quantify the band structure differences in somewhat more detail, we follow the "band delta" (or \triangle_{band}) concept proposed previously[4], which is defined as a root-mean-square deviation defined on the energy levels of two band structures $\varepsilon_{1,n}[\vec{k}_i]$ and $\varepsilon_{2,n}[\vec{k}_i]$ within a given energy window $[-\varepsilon_l, \varepsilon_u]$:

$$\Delta_{band}\left[-\varepsilon_{l},\varepsilon_{u}\right] = \sqrt{\frac{1}{N_{E}}\sum_{i=1}^{N_{k}}\sum_{\substack{\varepsilon_{1,n} \leq \varepsilon_{u} \\ \varepsilon_{2,n} \leq \varepsilon_{u} \\ \varepsilon_{1,n} \geq -\varepsilon_{l} \\ \varepsilon_{1,n} \geq -\varepsilon_{l}}} \left(\varepsilon_{1,n}\left[\vec{k}_{i}\right] - \varepsilon_{2,n}\left[\vec{k}_{i}\right]\right)^{2} \right]$$
(21)

 N_k is the number of unique k points calculated along the k path; N_E is the total number of energy eigenvalues counted across all \vec{k}_i that lie within the energy window $[-\varepsilon_l, \varepsilon_u]$; $\varepsilon_{1,n}[\vec{k}_i]$ and

 $\varepsilon_{2,n}[\vec{k}_i]$ are the energy eigenvalues of the two band structures being compared at the *k* point \vec{k}_i . In practice, we choose the range of valence bands [VBM-10 eV, VBM] and conduction bands [CBM, CBM+5 eV] as the energy windows to quantify the band delta between different relativistic approaches (for metals, the Fermi level is taken as the VBM and CBM). The results are plotted in Fig. 6.

It is not surprising that the scalar relativistic results show significant deviations from Q4C benchmarks especially for materials with $Z_{max} > 31$ (Z_{max} denotes the maximum atomic number in the compound), for which SOC effects start to become notable. However, there are still several metals (i.e. *Sr*, *Y*, *Zr*, *Nb*, *Mo*, *Tc*, *Ru*, *Rh*, *Pd*, *Ba*, and *Hf*) and semiconductors (i.e. *InN* and *InP*) with larger Z_{max} , for which the SR calculations in fact give a relatively good description ($\Delta_{band} < 0.07 \text{ eV}$), which indicates that the SOC effect in these materials are not that strong. From Fig. 6(B) we see that the deviations of the conduction bands are often (but not always) smaller than that of the valence bands in the set of compounds considered here.

As for the non-self-consistent SR+SOC calculations, the Q4C benchmarks show that this computationally cheap non-self-consistent approach can already provide very accurate band structures for materials with $Z_{max} \le 50$, for which the valence $\triangle_{band} < 0.015 \ eV$ and the conduction $\triangle_{band} < 0.037 \ eV$. From $Z_{max}=51$ (*Sb*) to $Z_{max}=75$ (*Re*), the non-self-consistent SR+SOC treatment can still safely serve to capture any qualitatively relevant band-structure effects, with elevated valence \triangle_{band} values up to $0.082 \ eV$ ($Z_{max}=71$, Lu) and conduction \triangle_{band} values up to $0.136 \ eV$ ($Z_{max}=75$, *Re*). For the heaviest materials with $Z_{max} \ge 76$, in which the conduction and/or valence bands probed here include 6p shells (which are strongly affected by SOC[129, 130]), \triangle_{band} can be as large as $0.206 \ eV$ ($Z_{max}=83$, Bi) for valence bands and $0.320 \ eV$ ($Z_{max}=85$, Po) for conduction bands. Nevertheless, as noted previously[4], the relative error of the non-self-consistent SOC correction is still within only 11% of the overall magnitude of SOC effect, justifying the use of non-self-consistent SR+SOC treatment for qualitative analyses of band structures even for very heavy elements.

Finally, from Figs. 6(E) and 6(F) we see that the self-consistent SOC corrected results from *WIEN2k* (known as a benchmark-quality code[30]) show yet better agreement with the Q4C band structures: the deviations are roughly within 0.030 *eV* for valence bands and 0.05 *eV* for conduction bands. Note that, with a few exceptions, the deviations for light elements and heavy elements are generally at the same level, implying that the associated slight uncertainty should come from technical aspects like basis set effect, numerical errors, etc. However, there are notable outliers in the heavy-element part, especially metals, such as: *Os* (Z_{max} =76, valence $\triangle_{band} = 0.026 \ eV$, conduction $\triangle_{band} = 0.154 \ eV$), *Ir* (Z_{max} =77, valence $\triangle_{band} = 0.046 \ eV$, $\triangle_{band} = 0.122 \ eV$), and *Pt*

 $(Z_{\text{max}}=78, \text{ valence } \triangle_{\text{band}} = 0.027 \ eV$, conduction $\triangle_{\text{band}} = 0.094 \ eV$). A possible reason is that *WIEN2k* adopts additional $p^{1/2}$ orbitals only for 6*p* orbitals but misses the $5d^{3/2}$ orbitals, which might contribute somewhat to the quantitative result. However, this does not lead to fundamental differences (between *FHI-aims* and *WIEN2k*) in the overall band structures, as can be seen from the band structure comparison in Fig. S1.



Fig. 6. Band delta between **(A)** SR and Q4C valence bands, **(B)** SR and Q4C conduction bands, **(C)** non-self-consistent SR+SOC and Q4C valence bands, **(D)** non-self-consistent SR+SOC and Q4C conduction bands, **(E)** self-consistent APW+ $p^{1/2}$ and Q4C valence bands, **(F)** self-consistent APW+ $p^{1/2}$ and Q4C conduction bands. The 103 materials are ordered according to their maximum atomic number Z_{max} .

As a final point of this section, we chose twelve heavy-element based materials for a comparison of predicted NR, SR and FR equilibrium lattice constants and bulk moduli, as plotted in Fig. 7 and listed in Table S20. It can be seen that the relativistic effects (both SR and SOC) typically reduce the bond lengths of compounds (this is true for both the silver halides and semiconductors reported in Tables 3-5 and Table S20). However, for *Pb*, SOC in fact enlarges the lattice parameter slightly.



Fig. 7. The optimized lattice constants (in \hat{A}) and bulk moduli (in *GPa*) of several heavy-element based compounds at NR, SR, and Q4C levels, respectively.

D. Larger Perovskite Systems Containing Up to 94 Atoms

Hybrid organic-inorganic perovskites (HOIPs)[131, 132] have been found to be promising for solar energy and other semiconductor applications and have received significant attention both in experiments and in computations. The heavy elements (e.g. *I*, *Pb*, *Bi*) often involved in these

structures feature strong relativistic effects that mainly manifest in the form of SOC. While selfconsistent scalar relativity followed by hybrid density functional theory and non-self-consistent SOC has yielded remarkably good qualitative results in past work by some of us[75, 133-135], the capability to handle these large systems in a four-component way is expected to contribute to further improved computational predictions for these materials. While support for total-energy gradients ("forces") and hybrid density functionals is not yet available in the present work, the qualitative impact of non-self-consistent SOC vs. self-consistent Q4C can already be studied at the level of DFT-PBE. We choose a typical perovskite system 4-fluorophenethylammonium (4-FPEA) lead iodide, (4-FPEA)₂PbI₄[136, 137], containing 94 atoms. The crystal structure is shown in Fig. 8 (figures are made by the open source software VESTA[138]).



Fig. 8. The structure of $(4\text{-}FPEA)_2PbI_4$ which consists of 2 lead atoms, 8 iodine atoms, 4 fluorine atoms, 4 nitrogen atoms, 32 carbon atoms, and 44 hydrogen atoms in a unit cell. Each unit cell contains 4 organic molecules that is shown in (B). (C), (D), and (E) show the sideview of the unit cell along different axes. The unit cell lengths are: a = 8.513 Å, b = 8.767 Å, c = 16.684 Å. The angle between *a* and *b* is $\alpha = 89.978^{\circ}$, the angle between *b* and *c* is $\beta = 100.256^{\circ}$, the angle between *a* and *c* is $\gamma = 90.005^{\circ}$.

Band structure calculations using nonrelativistic (NR), atomic ZORA[28] scalar relativistic (SR), non-self-consistent scalar relativistic with SOC correction (SR+SOC), and Q4C methods, respectively, were performed using the PBE functional. Tight setting of *FHI-aims* and a k-space grid of $3 \times 3 \times 2$ k-points for reciprocal space integrals during the self-consistent field cycle were used. The plotted band structures are shown in Fig. 9.

For the present system with a large band gap, it can be seen that while NR and SR calculations give a wrong description, the non-self-consistent SR+SOC recipe can actually provide a quite reliable qualitative account of the band structure. Strong SOC splitting effect can be observed in the conduction bands (significantly impacted by the *Pb* 6*p* states): taking the energy levels at the Gamma point as an example, the CBM is lowered while the CBM+1 is significantly moved upwards. On the other hand, the valence bands (mainly contributed by the *I* atoms) change relatively little due to their much weaker SOC. The deviation of SR+SOC from Q4C in the present case is mainly reflected in the band gap, which is narrowed by $0.21 \ eV$.



Fig. 9. Band structures of (*4-FPEA*)₂*PbI*₄ calculated by nonrelativistic (NR), atomic ZORA scalar relativistic (SR), scalar relativistic with non-self-consistent SOC correction (SR+SOC), and Q4C methods. The Brillouin zone is also depicted for reference.

We finally consider the band structure of another exemplary (double) perovskite system containing bismuth, i.e. $Cs_2AgBiCl_6$. The lattice constant of the conventional unit cell used is 10.957Å; tight setting of *FHI-aims* and $5 \times 5 \times 5$ k points are adopted. We plot the SR+SOC bands and the Q4C bands in one figure for comparison, as shown in Fig. S2. It can be seen that the non-self-consistent SR+SOC recipe provides a good description for the valence band, and therefore the main difference comes from the conduction band: the (indirect) band gap is reduced from 1.81 *eV* (SR+SOC) to 1.51 *eV* (Q4C). This matches a similar observation for the same system, compared to *WIEN2k*'s APW+ $p^{1/2}$ self-consistent approach, shown in Figure S5 of reference [75].

V. Conclusions and Perspectives

In this paper, an efficient quasi-four-component (Q4C) relativistic method (maintaining the exact two-component regime) based on numerically tabulated atom-centered basis sets is developed. In our implementation, the following choices are made:

- We rely on the self-consistent Dirac-Coulomb Operator Eq. (1)[76, 77] in a form that neglects orbital current terms[19] but which still improves significantly over simpler scalar-relativistic or non-self-consistent spin-orbit coupled schemes by (1) including spin-orbit coupling self-consistently for all orbitals, (2) including basis functions with the correct numerical shape for the Dirac-Coulomb operator towards each nucleus, and (3) including an explicit representation of the small-component density and orbitals (i.e., considering all four components).
- The key point of the Q4C approach is that, for each large-component radial basis function in a multi-atom structure, one includes a specific small-component radial basis function, the shape of which is mathematically tied exactly to the large-component radial basis function from which it is derived. For each Dirac-Kohn-Sham orbital, the basis set coefficients for each small-component basis function are constrained to be exactly the same as for the corresponding large-component basis function (see Eq. (6)), i.e., the small component has no variational flexibility of its own. The net result is an effective twocomponent method, i.e., the two scalar component functions of the large component can vary independently, but the shape of the corresponding small component then follows from the large component with no further flexibility.
- Crucially, the Q4C prescription includes numerically exact radial solutions to spherically symmetric free atoms, obtained independently by the four-component Dirac-Coulomb method, as "minimal" basis functions. This choice captures the radial behavior of all basis functions towards the nucleus (where relativistic effects matter most and where the small component is numerically significant) practically exactly also in multi-atom systems. For the small component radial basis functions of the "minimal" basis, the atomic balance condition in Eq. (2) can be used to relate them to the large-component radial basis functions, a condition that is expected to yield practically exact Dirac-Coulomb results for the near-nuclear behavior in multi-atom systems as well.
- For all other radial basis functions associated with a particular atom, the approximate atomic balance condition Eq. (20), which is appropriate for the valence electron regime, can be used, building on the expectation that the core electrons of a multi-atom structure will be represented almost exclusively by the core radial functions of the "minimal" basis

of Dirac-Kohn-Sham radial functions of free atoms.

• The same underlying approximation of a fixed atomic-like reference solution for the small component can be applied in the construction of the electron density, in which a superposition of free-atom small-component densities is used to represent the multi-atom small-component density.

The present paper reports benchmark results for a set of systems including dimers and over 100 different solids, including complex organic-inorganic perovskites with up to 96 atoms/unit cell. We show that the Q4C strategy summarized above provides essentially the same precision as the "exact two-component" (X2C) approach as implemented in the Beijing Density Functional (*BDF*) code[23, 86, 87] and as the "Dirac core plus self-consistent spin-orbit coupled valence" all-electron implementation in the benchmark-quality *WIEN2k* code[88, 89], when including $p^{1/2}$ orbitals in the basis set; in fact, the precision of the present Q4C approach may be slightly better, when used with general high-quality basis sets, since it includes radial basis functions with the correct Dirac-Coulomb behavior for all other angular momenta as well. Importantly, the Q4C band structure of a large perovskite system demonstrates the ability of the method to treat complex systems with many heavy atoms in a unit cell.

One of the anonymous reviewers of this work raised the very interesting question of what is the "worst case scenario" for the validity of the Q4C approximation as we employ it here. In fact, for general multiatom structures and assuming standard basis sets in which the core levels of a given atom are represented by nearly exact free-atom derived Dirac-Coulomb radial functions centered at the same atom, we expect uniformly high accuracy throughout. From a numerical perspective, a breakdown of the method could only occur if, e.g., a core orbital centered at one atom were numerically significantly represented by one or more basis functions centered at another atom, a highly unlikely scenario even for large NAO basis sets as long as they remain well conditioned. Likewise, one might imagine that the validity of the atomic balance condition Eq. (2) for free-atom radial functions deteriorates if neutral free-atom basis functions were used to model a constrained electronic system with a core hole; however, this practical deficiency would easily be cured by including the Dirac-Coulomb radial functions of a free atom with a core hole in the set of minimal basis functions instead. In our view, a true worst-case scenario would arise only if the near-nuclear potential of an atom in a multiatom structure were substantially altered away from the free-atomlike near-nuclear potential, e.g., in a high-energy nuclear collision that places two heavy nuclei in very close proximity. In this case, the atomic balance condition of Eq. (2) and, in fact, the reliance on a free-atom derived minimal basis set for multiatom calculations in the first place, would no longer furnish a good numerical foundation for the overall multi-atom problem. However, in the scenario of a nuclear collision, the application of time-independent ground-state density functional theory itself (as done in this work) becomes questionable, i.e., a very different overall approach to the electronic structure of the problem might need to be employed.

Regarding the use of the Dirac-Coulomb Hamiltonian itself, we note that this model is based on several further approximations, which could easily be relaxed in a future extension of this work since corresponding, well established textbook expressions are known[77]. We specifically mention quantum electrodynamic (QED) vacuum fluctuations near the nucleus (e.g., Lamb shift, spontaneous emission). Likewise, the use of a finite-nucleus model (see reference[97] and Chapter 5.4 of reference [77]), more physical and actually numerically simpler than the electron-nuclear Coulomb singularity, is not yet pursued in the present work. Finally, the electron-electron interaction in QED theory is also more complex than the simple Coulomb interaction and wellknown correction terms exist.

We emphasize that the present work focuses on the numerical accuracy of the Q4C method in the NAO framework. This is the core objective of this paper. The present implementation allows one to routinely address the energetics and energy band structures of rather complex molecules and solids, but significant potential remains to extend the reach of the method in future work. For example, it should be straightforward to combine the approach with an efficient hybrid density functional implementation[139] or many-body methods beyond density functional theory, an avenue that we plan to pursue in future work. Likewise, we have not yet included open-shell systems or total energy gradients ("forces") with respect to nuclear positions in this paper, two further important extensions that will be critical for practical applications and that we will pursue. Finally, the present implementation exploits computational parallelism, but unlike the scalar-relativistic parts of the underlying code[28, 84, 85, 96], has not yet been fully optimized for or benchmarked on massively parallel computing platforms or GPU-based accelerator architectures[96, 140] expected to power the next generation of exascale-capable computing platforms. Extending the present code to become a much more capable high-performance computing implementation should be straightforward and is expected to truly unlock 4C treatments even for very large, complex problems in the future.

In summary, with numeric atomic orbital basis sets, the Q4C theory allows one to perform selfconsistent relativistic calculations, reflecting all four components, for systems containing a large number of heavy elements and therefore promises to place some formerly unaffordable application areas within reach. Since the SOC effect can here be treated precisely, we expect the present development to open the possibility to studying large periodic systems containing heavy elements, especially strong SOC systems such as heavy-element containing perovskites, topological insulators, and similar materials.

Acknowledgements

The authors are grateful to Dr. Ondřej Čertík at Los Alamos National Laboratory for his enthusiastic technical support on the open source code *dftatom* he authored, to Dr. Daoling Peng at South China Normal University and Dr. Ingo Opahle at Technische Universität Darmstadt for the discussion on Q4C theory, to Prof. Wenjian Liu at Shandong University for providing suggestions and for offering codes (which assemble the scalar integrals to spinor integrals, i.e., those in the first paragraph of Section IIIB) of the *BDF* program. The authors thank Matthias Scheffler and the Fritz-Haber-Institute, Berlin, for supporting this project. This work was also partially supported by NSF Awards DMR-1729297 and ACI-1450280. Finally, we are grateful to the anonymous reviewers of the paper for their thoughtful suggestions.

Appendix

Element	STO functions used
Н	1s (3.3, 2.0, 1.4, 1.0, 0.71), 2p (2.0, 1.0), 3d (2.5, 1.5)
F	1s (17.3, 10.05, 7.6), 2s (6.6, 4.45, 2.7, 1.8, 0.9), 2p (6.55, 3.65, 1.95, 1.15), 3d
	(2.5, 1.5), 4f(4.0, 2.0)

(A) Basis sets

Cl	1s (64.0, 20.6, 15.35), 2s (15.5, 6.85, 5.05), 3s (4.75, 3.15, 2.1, 1.5, 0.75), 2p
	(12.05, 7.05, 4.7), 5p(5.1, 2.45, 1.05, 1.06), 5a(4.0, 2.0, 1.0), 4f(2.5, 1.25)
Br	<i>Is</i> (520.0, 135.0, 47.2, 33.75), <i>2s</i> (30.15, 15.65, 12.5), <i>3s</i> (12.05, 7.35, 5.3), <i>4s</i>
	(3.95, 2.85, 2.05, 1.55), 2p (110.0, 36.3, 18.0, 12.8), $3p$ (9.8, 6.6, 4.65), $4p$ (4.4,
	2.8, 1.75, 1.13), 3d (18.6, 9.7, 5.55, 3.35), 4d (2.3, 1.15), 4f (4.0, 2.0, 1.0)
Ag	1s (1950.0, 480.0, 160.0, 65.3, 46.5), 2s (40.4, 22.4, 18.3), 3s (13.8, 10.4, 8.2), 4s
	(7.4, 5.05, 3.6), 5s (2.75, 1.8, 1.2, 0.82), 2p (480.0, 160.0, 47.8, 24.85, 18.3), 3p
	(13.9, 10.3, 7.9), 4p (6.35, 4.3, 2.95), 5p (2.5, 1.5, 0.82), 3d (18.3, 10.95, 7.25),
	4 <i>d</i> (5.45, 3.35, 2.0, 1.2), 4 <i>f</i> (4.0, 2.0, 1.0)
Ι	1s (2250.0, 550.0, 185.0, 76.5, 53.25), 2s (46.4, 25.9, 21.3), 3s (17.0, 12.75,
	10.1), 4 <i>s</i> (9.05, 6.35, 4.75), 5 <i>s</i> (3.5, 2.4, 1.7, 1.0), 2 <i>p</i> (550.0, 185.0, 59.6, 29.1,
	21.3), 3 <i>p</i> (15.2, 11.7, 9.2), 4 <i>p</i> (8.0, 5.6, 4.05), 5 <i>p</i> (3.25, 2.2, 1.45, 1.02), 3 <i>d</i> (21.5,
	13.1, 9.0), 4d (7.8, 5.6, 4.0, 2.75), 5d (1.5), 4f (4.0, 2.0, 1.0)
Au	1s (9750.0, 2180.0, 695.0, 270.0, 125.0, 85.45), 2s (73.75, 42.8, 35.85), 3s (30.0,
	22.95, 18.4), 4s (16.25, 11.9, 9.2), 5s (7.15, 5.15, 3.85), 6s (3.35, 2.35, 1.6, 1.07),
	2p (2180.0, 600.0, 190.0, 78.35, 44.05, 34.15), 3p (37.4, 21.85, 17.1), 4p (16.65,
	11.15, 8.3), 5 <i>p</i> (6.5, 4.5, 3.15), 6 <i>p</i> (2.5, 1.7, 1.1), 3 <i>d</i> (67.5, 31.3, 20.9, 15.6), 4 <i>d</i>
	(14.3, 9.65, 6.85), 5d (4.9, 3.05, 1.9, 1.2), 4f (20.45, 12.3, 7.3, 4.35), 5f (2.7, 1.35)
Tl	1s (10250.0, 2300.0, 730.0, 280.0, 131.0, 88.3), 2s (76.4, 44.25, 37.05), 3s (31.0,
	23.8, 19.1), 4s (17.9, 12.7, 9.8), 5s (7.65, 5.55, 4.15), 6s (3.3, 2.3, 1.65, 1.2), 2p
	(2300.0, 620.0, 200.0, 81.8, 45.6, 35.25), 3 <i>p</i> (37.0, 23.0, 17.65), 4 <i>p</i> (17.5, 11.7,
	8.75), 5 <i>p</i> (6.8, 4.75, 3.4), 6 <i>p</i> (2.5, 1.6, 1.05, 0.67), 3 <i>d</i> (67.75, 31.8, 21.4, 16.05),
	4d (15.05, 10.15, 7.25), 5d (6.15, 4.15, 2.7, 1.75), 6d (1.25), 4f (22.45, 13.45,
	8.05, 4.95), 5 <i>f</i> (3.0, 1.5)
Pb	1s (10500.0, 2350.0, 745.0, 285.0, 133.0, 89.75), 2s (77.8, 45.0, 37.7), 3s (32.0,
	24.2, 19.4), 4s (17.8, 12.9, 9.95), 5s (7.8, 5.7, 4.3), 6s (3.35, 2.35, 1.7, 1.25), 2p
	(2350.0, 630.0, 200.0, 83.3, 46.3, 35.8), 3p (39.5, 22.9, 18.0), 4p (18.0, 11.95,
	8.95), 5p (7.05, 5.0, 3.6), 6p (2.55, 1.65, 1.1, 0.73), 3d (69.25, 32.35, 21.75,
	16.35), 4d (15.45, 10.45, 7.45), 5d (6.3, 4.3, 2.85, 1.95), 6d (1.25), 4f (22.8,
	13.75, 8.35, 5.2), 5 <i>f</i> (3.2, 1.6)
Bi	1s (10750.0, 2410.0, 765.0, 295.0, 136.0, 91.2), 2s (79.2, 45.75, 38.3), 3s (33.0,
	24.6, 19.8), 4s (17.75, 13.1, 10.2), 5s (8.05, 5.95, 4.5), 6s (3.5, 2.5, 1.75, 1.25), 2p
	(2410.0, 640.0, 205.0, 85.0, 47.1, 36.35), 3 <i>p</i> (40.25, 23.3, 18.3), 4 <i>p</i> (18.45, 12.2,
	9.2), 5 <i>p</i> (7.25, 5.2, 3.75), 6 <i>p</i> (2.9, 2.0, 1.35, 0.91), 3 <i>d</i> (71.9, 33.15, 22.25, 16.7),
	4d (15.85, 10.7, 7.7), 5d (6.6, 4.6, 3.15, 2.15), 6d (1.3), 4f (23.6, 14.25, 8.7, 5.45),
	5 <i>f</i> (3.4, 1.7)
Ро	1s (11000.0, 2460.0, 780.0, 300.0, 139.0, 92.7), 2s (80.7, 46.45, 38.95), 3s (34.0,
	24.95, 20.1), 4s (20.25, 13.65, 10.55), 5s (8.4, 6.2, 4.7), 6s (4.6, 3.3, 2.3, 1.65,
	1.0), 2p (2460.0, 650.0, 210.0, 86.75, 47.9, 36.9), 3p (40.9, 23.65, 18.6), 4p
	(18.85, 12.5, 9.4), 5 <i>p</i> (7.5, 5.4, 3.9), 6 <i>p</i> (3.6, 2.45, 1.6, 1.05), 3 <i>d</i> (74.0, 33.85,
	22.65, 17.05), 4d (16.25, 10.95, 7.9), 5d (6.6, 4.65, 3.25, 2.3), 6d (1.4), 4f (24.4,
	14.7, 9.0, 5.75), 5f (3.6, 1.8)
At	1s (11250.0, 2520.0, 800.0, 305.0, 142.0, 94.2), 2s (82.1, 47.25, 39.6), 3s (34.0,
	25.4, 20.45), 4s (20.2, 13.85, 10.75), 5s (8.55, 6.35, 4.85), 6s (4.35, 3.25, 2.35,

1.75, 1.0), 2p (2520.0, 660.0, 210.0, 88.15, 48.6, 37.45), 3p (41.6, 24.0, 18.95),
4p (18.85, 12.65, 9.6), 5p (7.7, 5.55, 4.1), 6p (3.45, 2.4, 1.65, 1.12), 3d (76.4,
34.65, 23.1, 17.4), 4d (16.65, 11.25, 8.15), 5d (6.6, 4.65, 3.25, 2.35), 6d (1.5), 4f
(25.15, 15.15, 9.35, 6.0), 5 <i>f</i> (4.0, 2.0, 1.0)

Table S1. The QZ4P STO basis sets parameters used in *BDF*. The Zeta value for each STO basis function is listed in the parenthesis.

Basis set level	Atomic orbitals used
Minimal	[Xe] + 6s5d4f
First tier	$Au^{2+}(6p, 1.4), H(4f, 7.4), Au^{2+}(6s), H(5g, 10), H(6h, 12.8), H(3d, 2.5)$
Second tier	H(5f, 14.8), H(4d, 3.9), H(3p, 3.3), H(1s, 0.45), H(5g, 16.4), H(6h, 13.6)

Table S2. The "second tier" *FHI-aims* basis set for *Au*. Also see the caption of Table 1 for the description of the parameters.

(B) Integrations Involving the $\vec{\sigma} \cdot \vec{p}$ Operator

 $\vec{\sigma} \cdot \vec{p}$ is a Hermitian operator that is commonly involved in integrals related to the relativistic electronic structure problems. Here, we analyze the effect of applying $\vec{\sigma} \cdot \vec{p}$ to an atomic orbital, since this could be useful to many of the aforementioned integrations. Some fundamental relations used here can be found from Kenneth G. Dyall and Knut Fægri, Jr.'s book.[76]

Since the atomic basis is practically defined as a product of a radial part and an angular part, it is necessary to also separate the $\vec{\sigma} \cdot \vec{p}$ operator into two terms acting on the radial and angular basis functions, respectively. We start from Dyall and Fægri's derivation:

$$\vec{\sigma} \cdot \vec{p} = \sigma_r \hat{p}_r + i\sigma_r r^{-1} (\vec{\sigma} \cdot \vec{l} + 1) \equiv \sigma_r \hat{p}_r - i\sigma_r r^{-1} \hat{K}, \quad (S1)$$

where the radial momentum operator has the form

$$\hat{p}_r = -i\left(\frac{d}{dr} + \frac{1}{r}\right), \qquad (S2)$$

which can operate on the radial large or small components. σ_r is a pure angular operator with an important feature:

$$\sigma_r Y^L = Y^S, \ \sigma_r Y^S = Y^L.$$
 (S3)

Eq. (S3) means that this operator transforms the large-component angular function to the smallcomponent angular function, and vice versa.

 \widehat{K} is also a pure angular momentum operator,

$$\hat{K} = -1 - \vec{\sigma} \cdot \vec{l} = \hat{l}^2 + \hat{s}^2 - \hat{j}^2 - 1,$$
 (S4)

which can be applied to the angular momentum basis function and gives:

$$\hat{K}Y_{l}^{m_{l}}\eta(m_{s}) = \kappa Y_{l}^{m_{l}}\eta(m_{s}) = \left[l(l+1) - j(j+1) - \frac{1}{4}\right]Y_{l}^{m_{l}}\eta(m_{s})^{.}$$
(S5)

With all the conditions summarized above, we now observe the consequence of applying $\vec{\sigma} \cdot \vec{p}$

to a large/small component basis function:

$$\vec{\sigma} \cdot \vec{p} \left| \chi_{\nu}^{L} \right\rangle = -i \left(\frac{d}{dr} + \frac{\hat{\kappa} + 1}{r} \right) \hat{\sigma}_{r} R^{L}(r) Y_{lm}^{L} \eta(m_{s})$$

$$= -i \left[\frac{dR^{L}(r)}{dr} + \frac{R^{L}(r)}{r} \right] Y_{lm}^{S} \eta(m_{s}) - i \frac{R^{L}(r)}{r} \left[l(l+1) - j(j+1) - \frac{1}{4} \right] Y_{lm}^{S} \eta(m_{s})$$

(S6)

$$\vec{\sigma} \cdot \vec{p} \left| \chi_{\nu}^{s} \right\rangle = -i \left(\frac{d}{dr} + \frac{\hat{\kappa} + 1}{r} \right) \hat{\sigma}_{r} R^{s}(r) Y_{lm}^{s} \eta(m_{s})$$

$$= -i \left[\frac{dR^{s}(r)}{dr} + \frac{R^{s}(r)}{r} \right] Y_{lm}^{L} \eta(m_{s}) - i \frac{R^{s}(r)}{r} \left[l(l+1) - j(j+1) - \frac{1}{4} \right] Y_{lm}^{L} \eta(m_{s})$$
(S7)

(S7)

It is interesting to see that this operator transforms the large-component spherical harmonic function to a small-component one (this is due to the feature of the σ_r as can be seen from Eq. (S3)). And eventually, $\vec{\sigma} \cdot \vec{p}$ can be separated into a radial part that operates on the radial basis function and an angular part that is expressed as an eigenvalue comprising *l* and *j* quantum numbers.

(C) Generating $\vec{c\sigma} \cdot \vec{p} | \chi^s_{\mu} \rangle$

With the $|\chi_{\mu}^{s}\rangle$ generated from Eq. (20), we see that the $c\vec{\sigma}\cdot\vec{p}|\chi_{\mu}^{s}\rangle$ can be devided into two terms according to the chain rule:

$$\left. c\vec{\sigma} \cdot \vec{p} \right| \chi^{S}_{\mu} \rangle = \left[\vec{\sigma} \cdot \vec{p} \frac{c^{2}}{2c^{2} - V_{atom}(r)} \right] \vec{\sigma} \cdot \vec{p} \left| \chi^{L}_{\mu} \right\rangle + \frac{c^{2}}{2c^{2} - V_{atom}(r)} \left(\vec{\sigma} \cdot \vec{p} \right) \left(\vec{\sigma} \cdot \vec{p} \right) \left| \chi^{L}_{\mu} \right\rangle$$
(S8)

(S8)

As
$$(\vec{\sigma} \cdot \vec{p})(\vec{\sigma} \cdot \vec{p}) = \vec{p}^2 = -\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{l^2}{r^2}$$
, the second term can be obtained accordingly.

As for the first term, the $\vec{\sigma} \cdot \vec{p} |\chi_{\mu}^{L}\rangle$ part has been analyzed in Eq. (S6); and the remaining part

regarding $\vec{\sigma} \cdot \vec{p} \frac{c^2}{2c^2 - V_{\text{reg}}(r)}$ can be further expanded with the help of Eq. (S1):

$$\vec{\sigma} \cdot \vec{p} \frac{c^2}{2c^2 - V_{atom}(r)} = -i \left(\frac{d}{dr} + \frac{\hat{K} + 1}{r} \right) \frac{c^2}{2c^2 - V_{atom}(r)} = \frac{-ic^2}{\left[2c^2 - V_{atom}(r) \right]^2} \frac{dV_{atom}(r)}{dr}.$$
(S9)

Note that the \hat{K} operator equals -1 for the spherical symmetric potential of free atoms.

(D) Dimers

Stratom	FHI-	aims	BD	F	Liter	ature
System	r (Å)	ω (cm ⁻¹)	r (Å)	ω (cm ⁻¹)	r (Å)	ω (cm ⁻¹)

\mathbf{F}_{2}	1.384	1075.1	1.385	1062.6	1.416 ^a	920 ^a
Cl ₂	1.978	563.1	1.981	559.9	2.019 ^a	544 ^a
Br ₂	2.272	338.4	2.278	326.0	2.315 ^a	312 ^a
I ₂	2.643	218.3	2.669	211.7	2.717 ^a	206 ^a
At ₂	2.927	123.7	2.973	118.2	3.046 ^a	108 ^a
AgH	1.594	1863.6	1.590	1870.9		
AuH	1.529	2241.9	1.524	2320.1	1.525 ^b	2327.9 ^b
Ag_2	2.478	207.9	2.483	207.6	2.53°	207°
Au ₂	2.447	193.9	2.450	197.2	2.454 ^d	196 ^d
Pb ₂	2.844	130.6	2.891	120.0		
Bi ₂	2.569	204.5	2.637	184.5		
Po ₂	2.707	176.4	2.770	162.2	2.797^{f}	151.8 ^f

Table S3. Optimized LDA bond lengths r (in \mathring{A}) and vibrational frequencies ω (in cm^{-1}) of several diatomic molecules.

- a. 4C CCSD (T) by Visscher *et al.*[105]
- b. 4C DKS (LDA) by Saue *et al.*[10]
- c. 4C DKS (LDA) by Varga et al.[104]
- d. 4C DKS (LDA) by Liu *et al*.[141]
- e. Multireference Kramers' restricted CI by Han et al.[107]
- f. CAS-SCF by Mertens et al.[109]

Property	NR	SR	Q4C
Bond length	1.384	1.384	1.384
Vibrational frequency	1064.0	1063.8	1064.1
Binding Energy	4.21	4.21	4.17

Table S4. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (*eV*) of the F_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	1.979	1.979	1.978
Vibrational frequency	563.2	562.2	563.1
Binding Energy	3.95	3.94	3.86

Table S5. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (*eV*) of the Cl_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	2.280	2.279	2.272
Vibrational frequency	352.1	350.2	338.4
Binding Energy	3.41	3.39	3.06

Table S6. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (*eV*) of the Br_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C

Bond length	2.652	2.645	2.643
Vibrational frequency	226.9	225.6	218.3
Binding Energy	3.03	2.97	2.38

Table S7. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (*eV*) of the I_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	2.858	2.818	2.927
Vibrational frequency	163.8	162.0	123.7
Binding Energy	2.75	2.63	1.63

Table S8. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (*eV*) of the At_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	1.652	1.588	1.594
Vibrational frequency	1702.1	1873.3	1863.6
Binding Energy	3.76	4.01	3.97

Table S9. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (*eV*) of the *AgH* molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	1.699	1.523	1.529
Vibrational frequency	1705.3	2329.6	2241.9
Binding Energy	3.72	4.87	4.62

Table S10. The optimized bond lengths (in \hat{A}), vibrational frequencies (in cm^{-1}), and binding energies (eV) of the AuH molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	2.554	2.480	2.478
Vibrational frequency	183.1	207.3	207.9
Binding Energy	2.34	2.63	2.69

Table S11. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (eV) of the Ag_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	2.675	2.447	2.447
Vibrational frequency	137.9	196.0	193.9
Binding Energy	2.38	3.25	3.19

Table S12. The optimized bond lengths (in \hat{A}), vibrational frequencies (in cm^{-1}), and binding energies (*eV*) of the Au_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	2.928	2.878	2.844

Vibrational frequency	112.0	127.9	130.6
Binding Energy	3.56	3.58	1.58

Table S13. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (eV) of the Pb_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	2.676	2.612	2.569
Vibrational frequency	202.0	208.8	204.5
Binding Energy	5.99	6.01	3.40

Table S14. The optimized bond lengths (in \hat{A}), vibrational frequencies (in cm^{-1}), and binding energies (eV) of the Bi_2 molecule at NR, SR, and Q4C levels, respectively.

Property	NR	SR	Q4C
Bond length	2.763	2.716	2.707
Vibrational frequency	188.8	188.9	176.4
Binding Energy	4.44	4.38	3.50

Table S15. The optimized bond lengths (in \mathring{A}), vibrational frequencies (in cm^{-1}), and binding energies (eV) of the Po_2 molecule at NR, SR, and Q4C levels, respectively.

(E) 103 Compound Benchmark Set

Ductoture	Matarial	Band Gap	Band Gap	Band Gap	Band Gap	Band Gap
Prototype	Material	NR	SR	SR+SOC	APW+p1/2	Q4C
DIA	С	4.149	4.141	4.137	4.133	4.135
RS	MgO	4.766	4.714	4.702	4.718	4.699
WUR	AlN	4.244	4.208	4.208	4.216	4.208
ZB	AlN	3.314	3.313	3.307	3.312	3.307
ZB	SiC	1.228	1.371	1.367	1.349	1.366
ZB	BP	1.228	1.242	1.228	1.234	1.228
ZB	AlP	1.594	1.582	1.562	1.571	1.561
RS	MgS	2.840	2.778	2.747	2.774	2.768
WUR	ZnO	1.057	0.810	0.790	0.810	0.788
WUR	ZnS	2.436	2.159	2.138	2.136	2.132
ZB	ZnS	2.382	2.109	2.088	2.076	2.079
WUR	GaN	2.099	1.919	1.917	1.917	1.912
ZB	GaN	1.858	1.676	1.673	1.674	1.670
ZB	GaP	1.671	1.600	1.573	1.572	1.560
ZB	BAs	1.197	1.188	1.121	1.119	1.111
ZB	AlAs	1.487	1.443	1.342	1.354	1.343
ZB	GaAs	1.183	0.530	0.415	0.418	0.416
RS	MgSe	1.863	1.784	1.636	1.679	1.670
ZB	ZnSe	1.862	1.300	1.164	1.166	1.168
WUR	CdS	1.670	1.177	1.158	1.169	1.158
ZB	CdS	1.638	1.152	1.134	1.136	1.132

-							
	WUR	CdSe	1.411	0.675	0.544	0.566	0.553
	ZB	CdSe	1.387	0.656	0.525	0.540	0.531
	WUR	InN	0.252	0.000	0.000	0.002	0.001
	ZB	InP	1.228	0.712	0.676	0.666	0.660
	ZB	InAs	0.632	0.000	0.000	0.000	0.000
	ZB	AlSb	1.256	1.204	0.949	1.016	1.009
	ZB	GaSb	0.881	0.127	0.000	0.000	0.000
	ZB	InSb	1.188	0.000	0.000	0.000	0.000
	ZB	ZnTe	2.284	1.253	0.908	0.956	0.971
	ZB	CdTe	1.962	0.787	0.448	0.506	0.514
	ZB	HgS	1.164	0.000	0.038	0.109	0.108
	ZB	HgSe	0.983	0.000	0.000	0.000	0.000
	ZB	HgTe	1.743	0.000	0.000	0.000	0.000
	RS	PbS	0.382	0.370	0.113	0.062	0.078
	RS	PbSe	0.529	0.307	0.006	0.177	0.193
	RS	PbTe	0.254	0.731	0.148	0.088	0.097

Table S16. For a series of compound semiconductors, the band gaps calculated with nonrelativistic (NR), scalar relativistic (SR), scalar relativistic with SOC correction (SR+SOC), and fully relativistic (Q4C) methods are presented. All values are given in units of *eV*. Here, DIA, RS, WUR, and ZB denote diamond, rocksalt, wurtzite, and zinc blende structures, respectively.

D	Matarial	Band Gap	Band Gap	Band Gap	Band Gap	Band Gap
Prototype	Material	NR	SR	SR+SOC	APW+p1/2	Q4C
RS	LiF	9.195	9.154	9.135	9.137	9.134
RS	NaF	6.389	6.334	6.314	6.349	6.327
RS	LiCl	6.441	6.328	6.287	6.293	6.306
RS	NaCl	5.167	5.088	5.046	5.073	5.048
RS	KF	6.214	6.136	6.116	6.157	6.123
RS	KC1	5.181	5.164	5.124	5.157	5.140
RS	LiBr	5.272	4.930	4.745	4.775	4.755
RS	NaBr	4.412	4.138	3.956	4.002	3.975
RS	KBr	4.647	4.411	4.238	4.292	4.265
RS	RbF	6.010	5.682	5.676	5.723	5.752
RS	RbCl	5.165	4.972	4.928	4.966	4.950
RS	RbBr	4.642	4.348	4.175	4.220	4.193
RS	LiI	4.352	4.276	3.851	3.955	3.939
RS	NaI	4.225	3.627	3.207	3.323	3.298
RS	KI	4.430	3.942	3.543	3.663	3.638
RS	RbI	4.382	3.884	3.492	3.600	3.572
RS	CsF	5.479	5.434	5.388	5.458	5.426
RS	CsCl	5.246	4.877	4.791	4.836	4.828
CSCL	CsCl	5.154	5.270	5.192	5.208	5.202
CSCL	CsBr	4.716	4.529	4.327	4.352	4.343
CSCL	CsI	4.097	3.888	3.562	3.648	3.640

Table S17. For a series of alkali halides in rocksalt (RS) or cesium chloride (CSCL) structure, the band gaps calculated with nonrelativistic (NR), scalar relativistic (SR), scalar relativistic with SOC correction (SR+SOC), and fully relativistic (Q4C) methods are presented. All values are given in units of *eV*.

Prototype	Material	Band Gap NR	Band Gap SR	Band Gap SR+SOC	Band Gap Q4C
FCC	Ne	11.659	11.631	11.597	11.593
DIA	Si	0.585	0.570	0.554	0.553
DIA	Ge	0.623	0.049	0.000	0.000
FCC	Xe	6.736	6.264	5.724	5.839

Table S18. For a series of nonmetal elemental materials in face-centered cubic (FCC) or diamond (DIA) structure, the band gaps calculated with nonrelativistic (NR), scalar relativistic (SR), scalar relativistic with SOC correction (SR+SOC), and fully relativistic (Q4C) methods are presented. All values are given in units of eV.

Prototype	Material	K-Point	SR Energy	SO split	Q4C SO spit
НСР	Be	N/A	-	-	-
GRA	С	N/A	-	-	-
НСР	Mg	N/A	-	-	-
FCC	Al	W	-0.836	0.017	0.017
FCC	Ca	N/A	-	-	-
НСР	Sc	N/A	-	-	-
НСР	Ti	N/A	-	-	-
BCC	V	Г	0.413	0.037	0.039
BCC	Cr	Г	-0.827	0.051	0.053
FCC	Mn	Γ	-0.776	0.068	0.071
BCC	Fe	Н	9.594	0.260	0.257
НСР	Co	Н	-1.024	0.113	0.116
FCC	Ni	Г	-1.888	0.140	0.146
FCC	Cu	Γ	-2.939	0.181	0.187
НСР	Zn	N/A	-	-	-
DIA	Ge	Γ	0.000	0.294	0.294
FCC	Sr	W	0.619	0.051	0.052
НСР	Y	Γ	1.736	0.075	0.068
НСР	Zr	Г	1.522	0.111	0.114
BCC	Nb	Н	9.166	0.635	0.630
BCC	Мо	Н	9.177	0.730	0.724
НСР	Tc	Γ	-0.825	0.239	0.242
НСР	Ru	Г	-1.980	0.296	0.300
FCC	Rh	Γ	-3.101	0.253	0.255
FCC	Pd	Х	0.296	0.288	0.290
FCC	Ag	Г	-4.766	0.390	0.394
НСР	Cd	N/A	-	-	-

	DIA	Sn	Г	2.179	0.473	0.474
	BCC	Ba	Н	4.190	0.480	0.427
	НСР	Lu	Н	0.346	0.323	0.307
l	НСР	Hf	K	-1.797	0.387	0.349
	BCC	Ta	Р	-1.180	0.756	0.689
l	BCC	W	Н	5.269	0.748	0.726
	HCP	Re	А	-2.327	0.347	0.464
	НСР	Os	Н	-1.475	0.667	0.647
	FCC	Ir	Г	-3.917	0.841	0.831
l	FCC	Pt	Г	-3.800	1.002	0.983
	FCC	Au	W	6.574	1.448	1.259
l	НСР	Tl	Г	8.136	1.825	1.212
	FCC	Pb	Г	7.386	3.574	3.321
l	BCC	Bi	Г	6.080	4.581	4.127
	SCC	Ро	R	-2.694	2.803	2.562

Table S19. For a series of metals, the largest spin-orbit split in the energy range [Fermi energy - 10 eV, Fermi energy + 10 eV] is presented. All values are given in units of eV. HCP, GRA, FCC, BCC, and SCC denote hexagonal close-packed, graphite, face-centered cubic, body-centered cubic, simple cubic crystal structures, respectively. "N/A" indicates that either the spin-orbit splittings in the material are too weak to be visually distinguished in the energy range [VBM - 10 eV, VBM + 10 eV] or no unambiguous spin-orbit splitting can be visually identified in the energy range.



Fig. S1. A comparison of the PBE band structures of Os and Ir calculated using Q4C (in *FHI-aims*) and APW+ $p^{1/2}$ (in *WIEN2k*). The blue lines denote Q4C results and the red lines denote APW+ $p^{1/2}$ results.

Material	Property	NR	SR	Q4C
DLC	а	6.055	6.008	5.935
POS —	B_{θ}	49.2	52.4	58.5
	а	6.277	6.217	6.144
Pose —	B_{θ}	44.9	46.4	54.8

	a	6.628	6.570	6.479
PbTe —	Bo	36.0	38.4	50.3
	 a	6.186	5.994	5.962
HgS —	$B_{ heta}$	47.5	50.2	54.0
	а	6.450	6.262	6.227
HgSe —	Bo	40.6	42.5	46.0
II. T.	а	6.872	6.654	6.645
Hgle —	B_{θ}	32.7	34.0	33.8
I	а	4.022	3.870	3.875
Ir —	B_{θ}	229.0	345.0	311.6
D4	а	4.174	3.970	3.956
Pl —	$B_{ heta}$	137.5	247.6	227.5
4	а	4.428	4.156	4.136
Au —	$B_{ heta}$	67.7	136.6	122.5
DL	а	5.098	5.037	5.150
P0 —	$B_{ heta}$	39.4	37.4	33.2
D:	а	4.027	3.987	4.024
DI —	B_{θ}	50.9	53.2	41.1
D ₀ —	a	3.419	3.348	3.353
r0 —	B_{θ}	40.8	45.4	36.0

Table S20. Optimized lattice constants a (in \mathring{A}), bulk moduli B_{θ} (in GPa) at NR, SR, and Q4C levels, respectively, for twelve heavy-element containing solids.



(F) Energy Band Structure of CsAgBiCl₆

Fig. S2. The band structure of $Cs_2AgBiCl_6$ calculated by scalar relativistic with non-self-consistent SOC correction (SR+SOC) and fully relativistic Q4C methods. The lattice constant used for the

FCC structure is 5.478 Å. The red dots denote the SR+SOC bands; the blue dots denote the Q4C bands. See the appendix of reference [75] for a more detailed discussion.

References

- [1] J. Sichau *et al.*, Physical Review Letters **122**, 046403 (2019).
- [2] G. Dresselhaus, Physical Review 100, 580 (1955).
- [3] R. Winkler *et al., Spin-Orbit Coupling in Two-Dimensional Electron and Hole Systems* (Springer, 2003), Vol. 41.
- [4] W. P. Huhn, and V. Blum, Physical Review Materials 1, 033803 (2017).
- [5] P. Pyykkö, Annual Review of Physical Chemistry 63, 45 (2012).
- [6] P. Pyykkö, Chemical Reviews 88, 563 (1988).
- [7] N. E. Christensen, and B. O. Seraphin, Physical Review B 4, 3321 (1971).
- [8] W. Liu, National Science Review 3, 204 (2016).
- [9] L. Belpassi et al., Physical Chemistry Chemical Physics 13, 12368 (2011).
- [10] T. Saue, and T. Helgaker, Journal of Computational Chemistry 23, 814 (2002).
- [11] T. Yanai et al., The Journal of Chemical Physics 115, 8267 (2001).
- [12] L. Visscher, E. Eliav, and U. Kaldor, The Journal of Chemical Physics 115, 9720 (2001).
- [13] T. Fleig, L. K. Sørensen, and J. Olsen, Theoretical Chemistry Accounts 118, 347 (2007).
- [14] P. Hohenberg, and W. Kohn, Physical Review 136, B864 (1964).
- [15] W. Kohn, and L. J. Sham, Physical Review 140, A1133 (1965).
- [16] W. Kohn, and N. Rostoker, Physical Review 94, 1111 (1954).
- [17] S. Bei der Kellen, and A. J. Freeman, Physical Review B 54, 11187 (1996).
- [18] L. F. Mattheiss, Physical Review 151, 450 (1966).
- [19] D. A. Rehn et al., Physical Review B 101, 085114 (2020).
- [20] S. Suzuki, and K. Nakao, Journal of the Physical Society of Japan 68, 1982 (1999).
- [21] H. Eschrig, M. Richter, and I. Opahle, in Relativistic Electronic Structure Theory, Part 2.

Applications, edited by Schwerdtfeger (Elsevier, Theoretical and Computational Chemistry, 2004), pp. 723.

- [22] M. Kadek, M. Repisky, and K. Ruud, Physical Review B 99, 205103 (2019).
- [23] R. Zhao et al., J. Chem. Phys. 144, 044105 (2016).
- [24] W. A. de Jong, R. J. Harrison, and D. A. Dixon, The Journal of Chemical Physics 114, 48 (2001).
- [25] L. Visscher, and E. van Lenthe, Chemical physics letters 306, 357 (1999).
- [26] D. D. Koelling, and B. N. Harmon, Journal of Physics C: Solid State Physics 10, 3107 (1977).
- [27] A. H. MacDonald, W. E. Picket, and D. D. Koelling, Journal of Physics C: Solid State Physics **13**, 2675 (1980).
- [28] V. Blum et al., Computer Physics Communications 180, 2175 (2009).
- [29] V. Havu et al., Journal of Computational Physics 228, 8367 (2009).
- [30] K. Lejaeghere et al., Science 351, aad3000 (2016).
- [31] S. R. Jensen et al., The Journal of Physical Chemistry Letters 8, 1449 (2017).
- [32] D. Peng et al., The Journal of chemical physics 127, 104106 (2007).
- [33] W. Liu, and D. Peng, The Journal of chemical physics 131, 031104 (2009).
- [34] W. Liu, Molecular Physics 108, 1679 (2010).
- [35] W. Liu, Physics Reports 537, 59 (2014).

- [36] W. Liu, SCIENTIA SINICA Chimica 50, 1672 (2020).
- [37] W. Liu, The Journal of Chemical Physics 152, 180901 (2020).
- [38] O. Čertík, J. E. Pask, and J. Vackář, Computer Physics Communications 184, 1777 (2013).
- [39] E. v. Lenthe, E. J. Baerends, and J. G. Snijders, The Journal of Chemical Physics 101, 9783 (1994).
- [40] E. v. Lenthe, J. G. Snijders, and E. J. Baerends, The Journal of Chemical Physics 105, 6505 (1996).
- [41] E. v. Lenthe, E. J. Baerends, and J. G. Snijders, The Journal of Chemical Physics 99, 4597 (1993).
- [42] M. Reiher, Theoretical Chemistry Accounts 116, 241 (2006).
- [43] M. Reiher, WIREs Computational Molecular Science 2, 139 (2012).
- [44] M. Douglas, and N. M. Kroll, Annals of Physics 82, 89 (1974).
- [45] G. Jansen, and B. A. Hess, Physical Review A 39, 6016 (1989).
- [46] T. Nakajima, and K. Hirao, The Journal of Chemical Physics 113, 7786 (2000).
- [47] C. van Wüllen, The Journal of Chemical Physics 120, 7307 (2004).
- [48] A. Wolf, M. Reiher, and B. A. Hess, The Journal of Chemical Physics 117, 9215 (2002).
- [49] M. Reiher, and A. Wolf, The Journal of Chemical Physics 121, 10945 (2004).
- [50] E. Wimmer et al., Physical Review B 24, 864 (1981).
- [51] K. Schwarz, P. Blaha, and G. K. H. Madsen, Computer Physics Communications 147, 71 (2002).
- [52] P. Kurz et al., Physical Review B 69, 024415 (2004).
- [53] L. Nordström et al., Physical Review B 63, 035103 (2000).
- [54] P. Blaha et al., Physical Review B 46, 1321 (1992).
- [55] J. Kuneš et al., Physical Review B 64, 153102 (2001).
- [56] F. Aryasetiawan, and S. Biermann, Physical Review Letters 100, 116402 (2008).
- [57] A. Svane et al., Physical Review B 81, 245120 (2010).
- [58] A. Svane et al., Physical Review B 84, 205205 (2011).
- [59] P. Scherpelz et al., Journal of Chemical Theory and Computation 12, 3523 (2016).
- [60] F. Brivio et al., Physical Review B 89, 155204 (2014).
- [61] J. M. Frost et al., Nano Letters 14, 2584 (2014).
- [62] A. Goyal et al., Chemistry of Materials 30, 3920 (2018).
- [63] S. McKechnie et al., Physical Review B 98, 085108 (2018).
- [64] D. Vanderbilt, Physical Review B 41, 7892 (1990).
- [65] N. Troullier, and J. L. Martins, Physical Review B 43, 1993 (1991).
- [66] W. E. Pickett, Computer Physics Reports 9, 115 (1989).
- [67] M. Dolg, in *Theoretical and Computational Chemistry*, edited by P. Schwerdtfeger (Elsevier, 2002), pp. 793.
- [68] G. Kresse, and D. Joubert, Phys. Rev. B 59, 1758 (1999).
- [69] G. Kresse, and J. Furthmüller, Phys. Rev. B 54, 11169 (1996).
- [70] P. Giannozzi et al., Journal of Physics: Condensed Matter 29, 465901 (2017).
- [71] J. C. Stewart et al., Zeitschrift für Kristallographie Crystalline Materials 220, 567 (2005).
- [72] X. Gonze et al., Computer Physics Communications 248, 107042 (2020).
- [73] T. Hangele et al., The Journal of Chemical Physics 136, 214105 (2012).
- [74] E. T. McClure et al., Chemistry of Materials 28, 1348 (2016).
- [75] M. K. Jana et al., Journal of the American Chemical Society 141, 7955 (2019).
- [76] K. G. Dyall, and K. Fægri Jr, *Introduction to relativistic quantum chemistry* (Oxford University Press, 2007).
- [77] I. P. Grant, Relativistic quantum theory of atoms and molecules: theory and computation

(Springer Science & Business Media, 2007), Vol. 40.

- [78] R. E. Stanton, and S. Havriliak, The Journal of Chemical Physics 81, 1910 (1984).
- [79] L. Visscher et al., International Journal of Quantum Chemistry 40, 131 (1991).
- [80] K. G. Dyall, The Journal of Chemical Physics 106, 9618 (1997).
- [81] K. G. Dyall, The Journal of Chemical Physics 109, 4201 (1998).
- [82] W. Kutzelnigg, and W. Liu, The Journal of Chemical Physics 123, 241102 (2005).
- [83] Y. J. Franzke, N. Middendorf, and F. Weigend, The Journal of Chemical Physics **148**, 104110 (2018).
- [84] V. W.-z. Yu et al., Computer Physics Communications 222, 267 (2018).
- [85] V. W.-z. Yu et al., Computer Physics Communications 256, 107459 (2020).
- [86] W. Liu et al., Theor. Chem. Acc. 96, 75 (1997).
- [87] W. Liu, F. Wang, and L. Li, J. Theor. Comput. Chem. 2, 257 (2003).
- [88] P. Blaha et al., WIEN2k: An Augmented Plane Wave Plus Local Orbitals Program for Calculating
- Crystal Properties (Vienna University of Technology, Austria, 2001).
- [89] K. Schwarz, Journal of Solid State Chemistry 176, 319 (2003).
- [90] J. M. Soler et al., Journal of Physics: Condensed Matter 14, 2745 (2002).
- [91] G. te Velde et al., Journal of Computational Chemistry 22, 931 (2001).
- [92] K. Koepernik, and H. Eschrig, Physical Review B 59, 1743 (1999).
- [93] B. Delley, and D. E. Ellis, The Journal of Chemical Physics 76, 1949 (1982).
- [94] F. W. Averill, and D. E. Ellis, The Journal of Chemical Physics 59, 6412 (1973).
- [95] F. Knuth et al., Computer Physics Communications 190, 33 (2015).
- [96] W. P. Huhn et al., Computer Physics Communications 254, 107314 (2020).
- [97] D. Andrae, Physics Reports 336, 413 (2000).
- [98] T. Ziegler, and A. Rauk, Theoret. Chim. Acta 46, 1 (1977).
- [99] S. H. Vosko, L. Wilk, and M. Nusair, Canadian Journal of Physics 58, 1200 (1980).
- [100] P. Schwerdtfeger, Chemical Physics Letters **183**, 457 (1991).
- [101] W. A. de Jong, L. Visscher, and W. C. Nieuwpoort, The Journal of Chemical Physics **107**, 9046
- (1997).
- [102] W. A. de Jong *et al.*, The Journal of Chemical Physics **108**, 5177 (1998).
- [103] A. K. Mohanty, and F. A. Parpia, Physical Review A 54, 2863 (1996).
- [104] S. Varga *et al.*, Physical Review A **59**, 4288 (1999).
- [105] L. Visscher, and K. G. Dyall, The Journal of Chemical Physics 104, 9040 (1996).
- [106] S. Varga *et al.*, The Journal of Chemical Physics **112**, 3499 (2000).
- [107] Y.-K. Han, and K. Hirao, The Journal of Chemical Physics **112**, 9353 (2000).
- [108] T. Bastug *et al.*, Physical Review A **55**, 1760 (1997).
- [109] M. A. J. Mertens *et al.*, The Journal of Physical Chemistry Letters **10**, 2879 (2019).
- [110] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996).
- [111] F. D. Murnaghan, Proceedings of the National Academy of Sciences **30**, 244 (1944).
- [112] J. E. Peralta, J. Uddin, and G. E. Scuseria, The Journal of Chemical Physics **122**, 084108
- (2005).
- [113] B. N. Onwuagba, Solid State Communications 97, 267 (1996).
- [114] R. H. Victora, Physical Review B 56, 4417 (1997).
- [115] C. M. I. Okoye, physica status solidi (b) **234**, 580 (2002).
- [116] T. Söhnel, H. Hermann, and P. Schwerdtfeger, Angewandte Chemie International Edition 40,

4381 (2001).

[117] T. Söhnel, H. Hermann, and P. Schwerdtfeger, The Journal of Physical Chemistry B **109**, 526 (2005).

[118] D. Vogel, P. Krüger, and J. Pollmann, Physical Review B 58, 3865 (1998).

[119] C. R. Berry, Physical Review 97, 676 (1955).

[120] W. Hidshaw, J. T. Lewis, and C. V. Briscoe, Physical Review 163, 876 (1967).

[121] A. H. Sommer, Photoemissive Materials: Preparation, Properties, and Uses (Wiley, New York,

1968), Vol. Chap. 3.

[122] P. Wang *et al.*, Chemistry – A European Journal **15**, 1821 (2009).

[123] F. Bassani, R. S. Knox, and W. B. Fowler, Physical Review 137, A1217 (1965).

[124] G. Piermarini, and C. Weir, J. Res. Natl. Bur. Stand. A 66, 325 (1962).

[125] S. Ves *et al.*, Physical Review B **24**, 3073 (1981).

[126] P. Villars, and L. Calvert, American Society of Metals, Cleveland, OH (1985).

[127] S. Adachi, *Handbook on physical properties of semiconductors* (Springer Science & Business Media, 2004), Vol. 2.

[128] P. Villars, and J. L. C. Daams, Journal of Alloys and Compounds 197, 177 (1993).

[129] P. Carrier, and S.-H. Wei, Physical Review B **70**, 035212 (2004).

[130] P. Larson, Physical Review B **68**, 155121 (2003).

[131] B. Saparov, and D. B. Mitzi, Chemical Reviews **116**, 4558 (2016).

[132] W. Li et al., Nature Reviews Materials 2, 16099 (2017).

[133] C. Liu *et al.*, Physical Review Letters **121**, 146401 (2018).

[134] W. A. Dunlap-Shohl *et al.*, Materials Horizons **6**, 1707 (2019).

[135] M. K. Jana *et al.*, Nature Communications **11**, 4699 (2020).

[136] D. B. Mitzi, C. D. Dimitrakopoulos, and L. L. Kosbar, Chemistry of Materials 13, 3728 (2001).

[137] A. H. Slavney *et al.*, Inorganic Chemistry **56**, 46 (2017).

[138] K. Momma, and F. Izumi, J. Appl. Crystallogr. 44, 1272 (2011).

[139] S. V. Levchenko *et al.*, Computer Physics Communications **192**, 60 (2015).

[140] V. W.-z. Yu *et al.*, Computer Physics Communications **262**, 107808 (2021).

[141] W. Liu, and C. van Wüllen, The Journal of Chemical Physics **113**, 2506 (2000).