

The Elephant in the Room of Density Functional Theory Calculations

Stig Rune Jensen,^{*,†} Santanu Saha,[‡] José A. Flores-Livas,[‡] William Huhn,[¶] Volker Blum,[¶] Stefan Goedecker,[‡] and Luca Frediani[†]

[†]*Centre for Theoretical and Computational Chemistry, Department of Chemistry,
UiT - The Arctic University of Norway, N-9037 Tromsø, Norway*

[‡]*Department of Physics, Universität Basel, Klingelbergstr. 82, 4056 Basel, Switzerland*

[¶]*Department of Materials Science and Mechanical Engineering, Duke University,
Durham, NC 27708, USA*

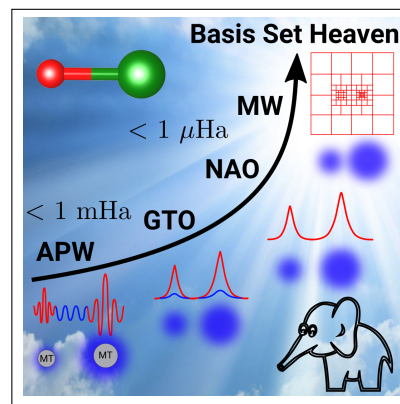
E-mail: stig.r.jensen@uit.no

Abstract

Using multiwavelets, we have obtained total energies and corresponding atomization energies for the GGA-PBE and hybrid-PBE0 density functionals for a test set of 211 molecules with an unprecedented and guaranteed μ Hartree accuracy. These quasi-exact references allow us to quantify the accuracy of standard all-electron basis sets that are believed to be highly accurate for molecules, such as Gaussian-type orbitals (GTOs), all-electron numeric atom-centered orbitals (NAOs) and full-potential augmented plane wave (APW) methods. We show that NAOs are able to achieve the so-called chemical accuracy (1 kcal/mol) for the typical basis set sizes used in applications, for both total and atomization energies. For GTOs, a triple-zeta quality basis has mean errors of ~ 10 kcal/mol in total energies, while chemical accuracy is almost reached for a quintuple-zeta basis. Due to systematic error cancellations, atomization energy errors are reduced by almost an order of magnitude, placing chemical accuracy within reach also for medium to large GTO bases, albeit with significant outliers. In order to check the accuracy of the computed densities, we have also investigated the dipole moments, where in general, only the largest NAO and GTO bases are able to yield errors below 0.01 Debye. The

observed errors are similar across the different functionals considered here.

Graphical TOC Entry



Electronic structure calculations are nowadays employed by a large and steadily growing community, spanning condensed matter physics, physical chemistry, material science, biochemistry and molecular biology, geophysics and astrophysics. Such a popularity is in large part due to the development of Density Functional Theory (DFT) methods,¹ in their Kohn–Sham (KS) formulation.²

Although the exact energy functional of DFT is unknown, many approximate functionals offer an excellent compromise between accuracy and numerical cost, rivaling often the accuracy that can be obtained with correlated methods, such as Coupled-Cluster Singles Doubles (CCSD).^{3–5} During the last decades, extensive efforts have been undertaken to provide ever more accurate approximations to the exact Exchange–Correlation (XC) functional.⁶ This quest for higher accuracy is conceptually captured by John Perdew’s Jacob’s ladder analogy,⁷ leading to the heaven of chemical accuracy: errors of 1 kcal/mol or less in atomization energies and other energy differences that are of primary interest in chemistry and solid state physics. Rungs on this ladder are the Local Density Approximation (LDA), the Generalized Gradient Approximation (GGA),⁸ meta-GGAs,⁹ hybrid and double hybrid functionals.¹⁰ The best modern XC functionals come fairly close to this target, with errors of a few kcal/mol on a wide range of energetic properties relative to experiment, including atomic and molecular energies, bond energies, excitation and isomerization energies and reaction barriers, for main-group elements as well as transition metals and solids.^{11,12}

The closer we get to chemical accuracy, the more important becomes the identification of errors due to various other, algorithmic approximations – basis sets, integration grids and pseudopotentials,^{13–15} to cite a few – which can lead to comparable or even larger errors, but their influence is hard to quantify. The importance of this issue has recently been highlighted within the solid state community, with a substantial effort to assess the influence of such approximations on the accuracy. Lejaeghere *et al.*¹⁶ compared the GGA-PBE⁶ calculated equa-

tions of state for 71 elemental crystals from 15 different widely used DFT codes, employing both all-electron methods as well as 40 different pseudopotential sets. For the equation of state, most DFT codes agree within error bars that are comparable to those of experiment, irrespective of the basis-set choice: all-electron numeric atom-centered orbitals (NAOs), augmented plane wave (APW) methods or plane waves with pseudopotentials. The basis set issue has also been highlighted by Mardirossian and Head-Gordon,^{17,18} in connection with the development of the B97M-V and ω B97M-V functionals. Although the functionals have been designed and optimized by making use of a large basis set (aug-cc-pVQZ for B97M-W and def2-QZVPD for ω B97M-V), the authors have extensively explored the effect of smaller bases on the functional performance.

APW methods¹³ are widely believed to be highly accurate, but contain several parameters which are difficult to adjust and which can influence the results in a more or less erratic way. Hence, the magnitude of the error cannot be rigorously quantified without an external reference. Similar limitations exist for atomization energies of molecules obtained with Gaussian-type orbital (GTO) and NAO basis sets: both bases cannot be systematically enlarged to achieve completeness in the L^2 sense, and within standardized basis sets, the convergence to the exact result cannot be achieved. Additionally, for larger systems, linear dependency issues can limit the ability of these basis sets to achieve complete convergence.¹⁹

The basic mathematical formalism for KS DFT calculations leads to a self-consistent three-dimensional partial differential equation. What makes the solution of this equation so challenging are the accuracy requirements for the physically and chemically relevant energy differences. For instance, the atomization energy of the largest molecule (SiCl_4) in our data set is less than 1 Ha, but it is computed as a difference of energies in the order of ~ 2000 Ha. Hence we need at least 7 correct decimal places in the total energy of the molecule to get the atomization energy within chemical accuracy of 1 kcal/mol \sim milli-Hartree (mHa), and even 10 decimal places

for micro-Hartree (μHa) accuracy.

For isolated atoms, and using the appropriate numerical techniques, the associated many-particle problem can be solved with essentially arbitrary numerical precision. Virtually converged LDA energies for spherical atoms are available in the NIST data base.²⁰ For a few dimers, highly accurate energies have been calculated,²¹ and an attempt to obtain total energies free of basis set error was made also for general molecular systems,²² but the accuracy of this approach seems to be limited to around 1 kcal/mol. Similar accuracies were achieved for solids with semicardinal wavelets.²³

In spite of all the progress that has been made in numerical techniques to harness the power of quantum mechanical theory and simulations, none of the traditional techniques is able to furnish, unambiguously, atomization energies for molecules with arbitrary numerical precision. A straightforward, uniform grid- or Fourier transform-based approach is ruled out since it is impossible to provide sufficient resolution for the rapidly varying wave functions near the nucleus. Other basis set techniques are critically hampered by non-orthogonality, which leads to inevitable algebraic ill-conditioning problems at small but finite residual precision.¹⁹ Because of these problems a large part of the community resorts to pseudopotentials methods,¹³ where the $Z/|r - R|$ potential is replaced by a smoother potential that retains approximately the same physical properties of the all-electron atom. The smooth pseudopotential then allows to obtain arbitrarily high accuracy with systematic basis sets such as plane waves. The limitation is that the pseudopotential introduces an approximation error, the magnitude of which is hard to quantify.

The current *de facto* standard technique to assess errors of different methods does not rely on an absolute reference: errors are instead estimated by comparing results obtained with increasingly large bases.^{24,25} The development of Multiwavelet (MW) methods²⁶⁻²⁹ has fundamentally changed the situation. MWs are systematic, adaptive, and can be employed in all-electron calculations. With this approach, it is now possible to achieve all-electron energies

with arbitrarily small errors.

In the present work, we use MWs to obtain error bars of less than a μHa in the atomization energies for a large test set of 211 molecules with standard DFT functionals. We focus on three widely used and well established functionals, LDA-SVWN5,³⁰ GGA-PBE,⁶ as well as hybrid-PBE0.³¹ PBE and PBE0 are both relatively accurate for atomization energies,³² and have stood the test of time.³³ Our MW results provide quasi-exact reference values that can be employed to quantify the accuracy of standard basis sets, such as GTO, NAO and APW methods, as well as of novel approaches based for instance on finite element methods³⁴⁻³⁷ or discontinuous Galerkin methods.³⁸

Real-space methods have a long history in computational chemistry and have been used for benchmarking purposes for decades.³⁹ However, because of the so-called *curse of dimensionality*, the naïve numerical treatment of molecular systems is prohibitively expensive, and its applicability relies on high symmetry to reduce the dimensionality.⁴⁰ The multi-scale nature of the problem renders the traditional uniform grid discretization highly inefficient, unless the problematic nuclear region is treated separately, e.g. by means of pseudopotentials. The mathematical theory to solve these issues was developed in the '90s, when Alpert introduced the MW basis, allowing for non-uniform grids with strict control of the discretization error, as well as sparse representations of a range of physically important operators, with high and controllable precision.^{41,42}

Alpert's construction starts from a standard polynomial basis of order k , such as the Legendre or the Interpolating polynomials, re-scaled and orthonormalized on the unit interval $[0, 1]$. Then, an orthonormal scaling basis at refinement level 2^{-n} is constructed by dilation and translation of the original basis functions $\phi_{i,l}^n(x) = 2^{n/2}\phi_i(2^n x - l)$, where $\phi_{i,l}^n$ is the i -th polynomial in the interval $[l/2^n, (l+1)/2^n]$ at scale n . The set of scaling functions on all 2^n translations at scale n defines the scaling space V_k^n , and in this way a ladder of spaces is constructed such that the complete L^2 limit can be approached in a

systematic manner:

$$V_k^0 \subset V_k^1 \subset \dots \subset V_k^n \subset \dots \subset L^2. \quad (1)$$

The wavelet spaces W_k^n are simply the orthogonal complement of two subsequent scaling spaces V_k^n and V_k^{n+1} :

$$W_k^n \oplus V_k^n = V_k^{n+1}, \quad W_k^n \perp V_k^n. \quad (2)$$

Completeness in the L^2 sense can be achieved both by increasing the polynomial order (larger k) of the basis and by increasing the refinement in the ladder of spaces (larger n).

Two additional properties are essential in achieving fast and accurate algorithms: the vanishing moments of the wavelet functions and the disjoint support of the scaling and wavelet functions. The former leads to fast convergence in the representation of smooth functions and narrow-banded operators, whereas the latter enables simple algorithms for adaptive refinement of the underlying numerical grid, which is essential to limit storage requirements. The extension to several dimensions is achieved by standard tensor-product methods; to minimize the impact of the curse of dimensionality, it is necessary to apply operators in a separated form,^{43,44} by rewriting the full multi-dimensional operator as a product of one-dimensional contributions. For many important operators such a separation is not exact, but Beylkin and coworkers have shown that it can be achieved to any predefined precision as an expansion in Gaussian functions.^{43,45,46} To apply such operators in multiple dimensions, and simultaneously retain the local adaptivity in the representation of functions, it is essential to employ the non-standard form of operators,^{47,48} which, in contrast to the standard one, allows to decouple length scales.

These combined efforts (MW representation of functions and operators, separable operator representations, non-standard form of operators) made the accurate application of several important convolution operators efficient in three dimensions:

$$g(r) = \hat{G}_\mu f(r) = \int G_\mu(r - r') f(r') dr'. \quad (3)$$

Among such operators are the Poisson ($\mu = 0$) and the Bound-State Helmholtz (BSH) ($\mu^2 > 0$) kernels:

$$G_\mu(r - r') = \frac{e^{-\mu|r-r'|}}{4\pi|r-r'|}. \quad (4)$$

This mathematical framework was introduced to the computational chemistry community in the mid 2000's by Harrison and co-workers.²⁶⁻²⁹ They demonstrated that MWs could be employed to solve the KS equations in their integral reformulation:⁴⁹

$$\varphi_i = -2\hat{G}_{\mu_i} \hat{V} \varphi_i, \quad (5)$$

where the φ_i 's are the KS orbitals and the potential operator \hat{V} includes external (nuclear), Hartree, exchange and correlation contributions, while the kinetic operator and the orbital energies are included in the BSH operator as $2\hat{G}_{\mu_i} = (\hat{T} - \epsilon_i)^{-1}$, with $\mu_i^2 = -2\epsilon_i$. The ordinary KS equations $\hat{H}\varphi_i = \epsilon_i\varphi_i$, where $\hat{H} = \hat{T} + \hat{V}$ is the KS Hamiltonian, can be recovered by recalling that $(\nabla^2 - \mu^2)G_\mu(r - r') = -\delta(r - r')$. Such an integral formalism, when combined with a Krylov subspace accelerator,⁵⁰ leads to fast and robust convergence of the fix-point iteration of Eq. (5).

The integral formulation, in combination with MWs, provides unprecedented accuracy for all-electron calculations, without relying on molecular symmetry,²⁷⁻²⁹ and has also been extended to excited states⁵¹⁻⁵³ and electric^{54,55} and magnetic⁵⁶ linear response properties. In this approach all functions and operators, such as orbitals, densities and potential energy contributions to the KS Hamiltonian, are represented using MWs. Concerning potential energy terms, the external potential is obtained by projection* onto the MW basis, the Hartree potential is computed through Poisson's equation:

$$V_{\text{Hartree}}(r) = \int \frac{\rho(r')}{4\pi|r-r'|} dr', \quad (6)$$

and the XC potential is computed explicitly in

*For the singularity of the nuclear potential, a simple smoothing is employed, but its effect on the accuracy is easily controlled by a single parameter.²⁷

the MW representation from the following expression:[†]

$$V_{\text{XC}}(r) = \frac{\partial f_{\text{XC}}}{\partial \rho} - 2\nabla \frac{\partial f_{\text{XC}}}{\partial |\nabla \rho|^2} \cdot \nabla \rho. \quad (7)$$

The partial derivatives of the XC kernel f_{XC} can be mapped point-wise through external XC libraries,^{57,58} and the gradients are computed by the approach of Alpert *et al.*⁵⁹ For hybrid functionals, a fraction of the exact Hartree-Fock exchange contribution is included in the KS Hamiltonian (25% for PBE0). In our work we follow the method proposed by Yanai and coworkers,²⁸ where the exchange operator is defined as:

$$\hat{K}f(r) = \sum_j \varphi_j(r) \int \frac{\varphi_j(r')f(r')}{4\pi|r-r'|} dr'. \quad (8)$$

The above expression is again computed directly within the MW framework through repeated application of the Poisson operator to different orbital products.

While MWs are able to provide high-accuracy solution of integral equations in the form of Eq. (3), the same is not true for differential operators. In particular, high-order derivatives should be avoided in order to maintain accuracy in numerical algorithms. For this reason, we have found that the direct evaluation of the kinetic energy as a 2nd derivative of the wave function does not give the desired accuracy. Instead, we avoid the kinetic operator by computing the update to the eigenvalue directly:[‡]

$$\Delta \epsilon^n = \frac{\langle \varphi^{n+1} | \hat{V}^n | \Delta \varphi^n \rangle}{\langle \varphi^{n+1} | \varphi^{n+1} \rangle} + \frac{\langle \varphi^{n+1} | \Delta \hat{V}^n | \varphi^{n+1} \rangle}{\langle \varphi^{n+1} | \varphi^{n+1} \rangle}, \quad (9)$$

where the Δ 's refer to differences between iterations n and $n+1$. In contrast, the gradients in the expression for the GGA potential in Eq. (7) have not been found to affect the accuracy, partly because of a slightly conserva-

tive over representation of the density grid,[§] and partly because the XC energy is (by construction) only a small part of the total energy, thus reducing its relative accuracy requirement.

In this work, the MW calculations are performed with MRChem,⁶⁰ the GTO^{61,62} calculations with NWChem⁶³ and the NAO calculations with FHI-AIMS.^{64,65} APW+local orbital (APW+lo) calculations are performed with ELK.⁶⁶ The exchange-correlation functionals are calculated using the LIBXC⁵⁷ library in case of NWChem and FHI-AIMS, and the XCFUN⁵⁸ library for MRChem. We underline that the basis sets chosen are decoupled from their implementation in a particular code, with a host of other codes providing access to fundamentally the same numerical discretization schemes: See, for example, WIEN2K⁶⁷ or EXCITING⁶⁸ for APWs; DMOL3,^{69,70} FPLO,⁷¹ ADF,⁷² or PLATO⁷³ for NAOs; GAUSSIAN09,⁷⁴ GAMESS⁷⁵, DALTON⁷⁶, or MOLCAS⁷⁷ for GTOs; and MADNESS⁷⁸ for MWs.

The raw data of our study, as well as instructions for its reproducibility is available in the Supporting Information (SI).⁷⁹ Our test set comprises 211 molecules. In addition to the 147 systems from the G2/97 test set⁸⁰ containing light elements up to the third row, it contains molecules with chemical elements that are underrepresented in the G2/97 test set (Be, Li, Mg, Al, F, Na, S and Cl) as well as 6 non-bonded systems. For most of the systems, the experimental structure obtained from the NIST Computational Chemistry Comparison and Benchmark Database²⁴ was employed. In the remaining cases, geometries have been optimized at the MP2 level of theory, using the largest Gaussian basis set (see SI⁷⁹ for details).

[§]The grid is constructed such that it holds both the density and its gradient within the requested accuracy.

[†]Displayed as spin-unpolarized for clarity. Extension to spin-DFT is fairly straightforward.

[‡]This update is exact, provided that the orbital update comes directly from the application of the BSH operator defined by the previous (not necessarily exact) eigenvalue: $\Delta \varphi^n = -2\hat{G}_\mu^n [\hat{V}^n \varphi^n] - \varphi^n$. Generalizations can be made for multiple orbitals.

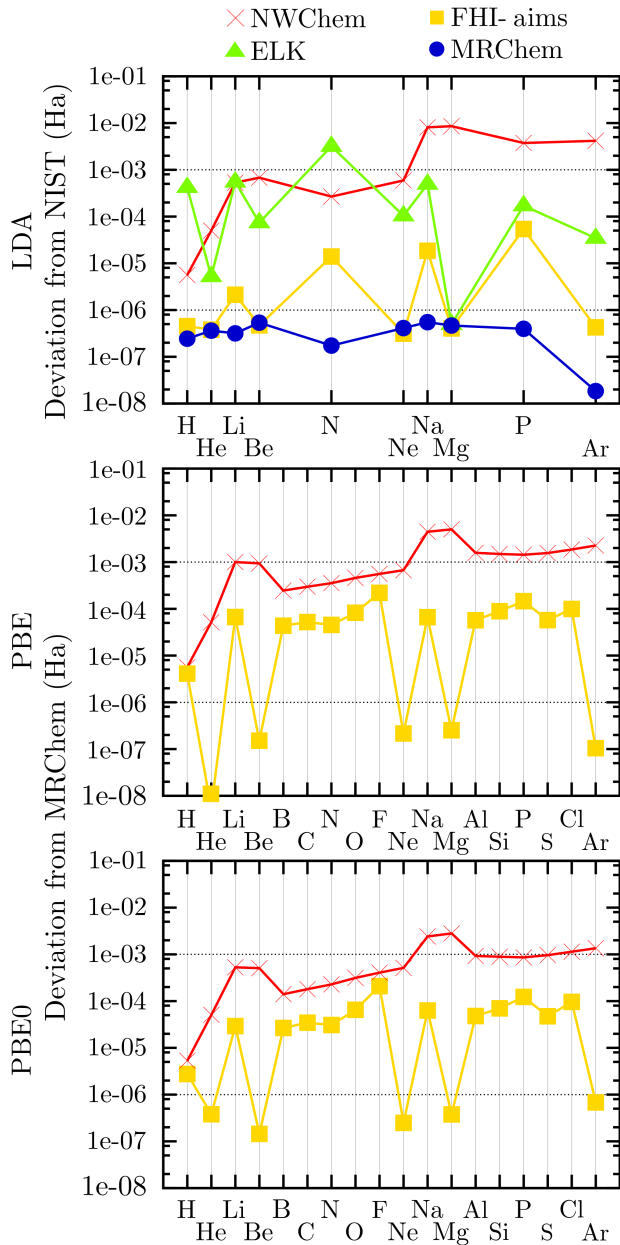


Figure 1: Absolute deviations in total energy found for different functionals for selected atoms. For LDA-SVWN5, energy differences are w.r.t. NIST all-electron values.²⁰ For GGA-PBE and hybrid-PBE0, the energy differences are w.r.t. MRChem. In all codes the largest basis set and tighter parameters were used. In all plots the reference values (NIST for LDA and MRChem for PBE/PBE0) are given with 6 decimal precision; a displayed error below 1e-06 Ha means that no discrepancy is detectable.

In the main part of our results we have considered four different basis sets each within NWChem and FHI-AIMS (APW+lo results with ELK could be obtained only for a small

subset), ranging from *small ones* intended for prerelaxations and energy differences between bonded structures (“light”, aug-cc-pVDZ), *production* basis sets considered in most publications (“tight”, “tier2” for FHI-AIMS and aug-cc-pVTZ, aug-cc-pVQZ for Gaussian codes), as well the *largest available* basis sets: “tier4” for FHI-AIMS and aug-cc-pV5Z for NWChem.[¶]

The construction and philosophy behind GTO basis sets is well documented in the quantum chemical literature, e.g. in the recent review by Jensen.⁸¹ In particular for the aug-cc-pVXZ bases the reader is referred to the original works from Dunning and coworkers.^{82–85}

The construction of the NAO basis sets used here is documented in detail in Ref.⁶⁴, and precise basis set definitions can be found in the SI.⁷⁹ The “tier” radial functions form a fixed basis set library, established for elements 1-102, and the choice of the exact radial functions for each element was carried out by an automated, computer guided strategy as described in Ref.⁶⁴. The sequence of successive “tier” basis sets is strictly hierarchical, beginning from a minimal basis of radial functions for the occupied core and valence states of free atoms. Additional “tiers” or groups of radial functions (constructed for free ions or hydrogen-like potentials) can then be added to increase the basis set size towards numerical convergence for DFT. An accurate, global resolution-of-identity approach (“RI-V” in Ref.⁶⁵) is employed to evaluate the four-center Coulomb operator in hybrid-PBE0 in FHI-AIMS. It is important to note that the NAO basis sets include a “minimal basis” of atomic radial functions determined for the same XC functional as used later in the three-dimensional SCF calculations. This is standard practice in FHI-AIMS for semi-local density functionals. For hybrid-PBE0, these radial functions are provided by linking FHI-AIMS to the “atom_sphere” atomic solver code for spherically symmetric free atoms developed in the

[¶]For Li, Be, Na and Mg the largest basis set is aug-cc-pVQZ and has therefore been employed. For the other elements, the corresponding 6Z basis is also available, but attempts to employ such a basis led to overcompleteness problems, often resulting in energies higher than the 5Z results.

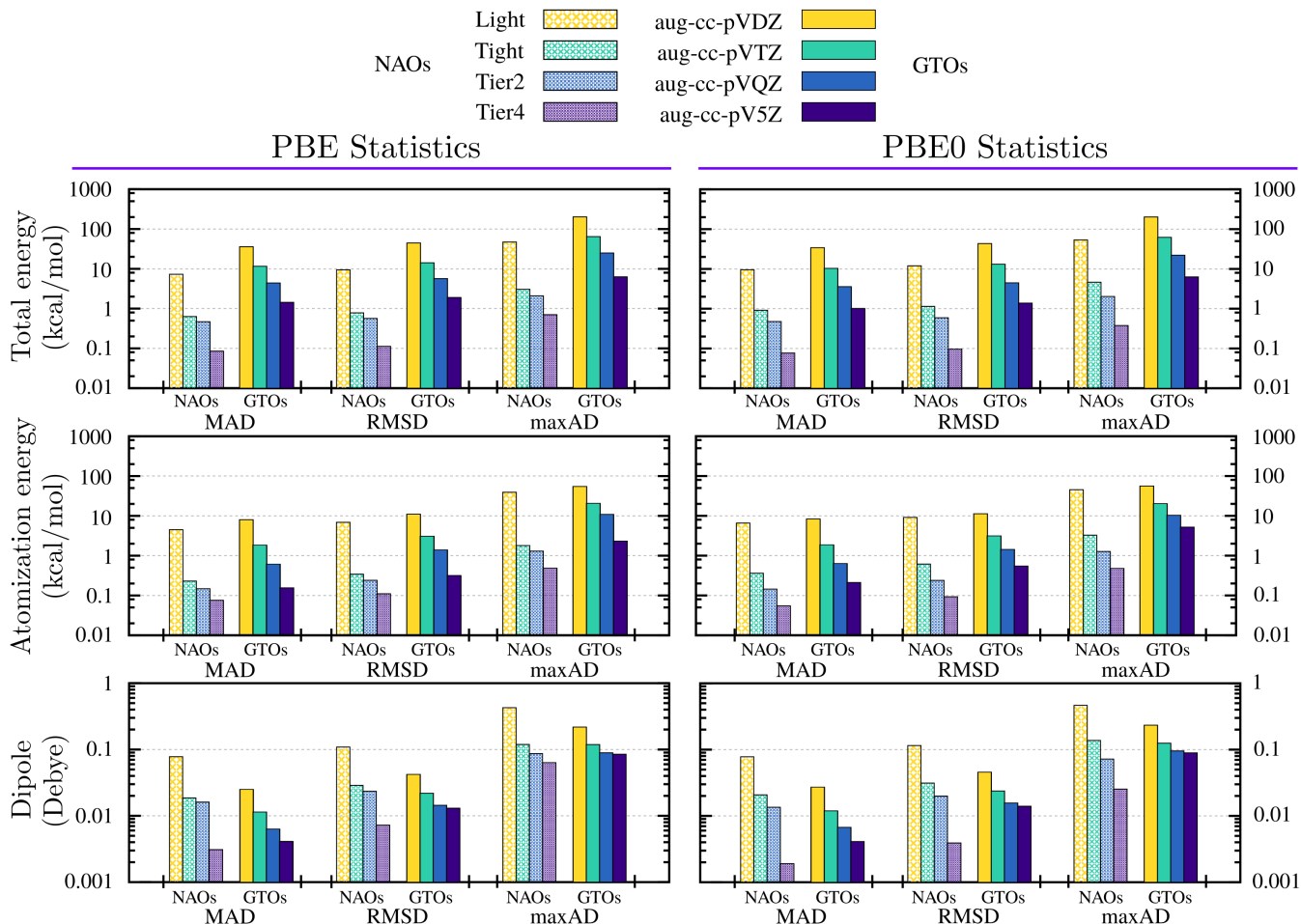


Figure 2: GGA-PBE (left) and hybrid-PBE0 (right) deviations in total energy, atomization energy, and electrostatic dipole moment for the set of 211 molecules with respect to highly accurate values obtained using MRCHEM. MAD, RMSD and maxAD stand for mean absolute deviation, root mean square deviation and maximum absolute deviation, respectively. Results are included for two different DFT codes (NWCHEM and FHI-AIMS) and four bases each: “light”, “tight”, “tier2” and “tier4” for FHI-AIMS and aug-cc-pVXZ (X=D, T, Q, 5) for NWCHEM.

Goedecker group for several years.¹⁴

The ground state energy of atoms from Hydrogen ($Z=1$) to Argon ($Z=18$) has been computed with the three chosen functionals. Our results are summarized in Figure 1. For all computational methods employed, the results of this section refer to the largest bases and tighter parameters: μHa for MRCHEM, “tier4” for FHI-AIMS and aug-cc-pV5Z for NWCHEM (see previous section and SI⁷⁹ for details). The top panel reports the LDA-SVWN5 values as absolute errors with respect to the reference values of the NIST²⁰ database for non-relativistic, spin polarized, spherically symmetric atoms. As expected, MWs yield differences which are consistently below the requested accuracy of 1 μHa .

The NAO and APW+lo approaches achieve average errors of $\sim 0.01\text{-}0.1$ mHa and $\sim 0.1\text{-}1$ mHa, respectively. GTOs are limited to around mHa accuracy. The GTO outliers (Li, Be, Na and Mg) have been computed with the aug-cc-pVQZ basis, because the aug-cc-pV5Z basis set is not available for these elements. Had 5Z-quality functions been available for all elements, a more uniform error for GTO would have resulted along the series, but the overall picture would only improve slightly.

In the GGA-PBE (middle) and hybrid-PBE0 (bottom) panels, all 18 atoms (both spherical and not), are included. The non-relativistic, spin-polarized electronic density and the total energy of the ground state, computed using

MRCHEM (converged within μHa) serves as the reference to which the other approaches are compared. For both functionals, NWCHEM performs at the limit of chemical accuracy ($\sim 1\text{ mHa}$). The NAOs in FHI-AIMS achieve 0.1 mHa or better, except for fluorine (0.3 mHa). For closed-shell atoms, FHI-AIMS is essentially exact because the exact radial functions of spherically symmetric, spin-unpolarized atoms are included in the basis sets. For GTOs, we observe that the total energy error grows with the atomic number, Z . In contrast, the accuracy of NAOs is less affected by the nuclear charge, with errors generally below 0.1 mHa for the Z range examined here, irrespective of the choice of functional. For APW+lo, only the LDA-SVWN5 values are included in Figure 1: the corresponding GGA-PBE and hybrid-PBE0 errors achieved in this work are above the threshold of $1\text{e-}03\text{ Ha}$ (dashed line) and were not considered further because it is unclear how much they might be affected by implementation-specific aspects other than the basis set.

The total energies, atomization energies and dipole moments of the 211 molecules considered have been computed within the GGA-PBE and hybrid-PBE0 functionals using MRCHEM with the highest affordable precision (below $1\ \mu\text{Ha}$ throughout). Figure 2 reports the Mean Absolute Deviation (MAD), Root Mean Square Deviation (RMSD) and Maximum Absolute Deviation (maxAD) obtained for total energy (top panel), atomization energy (medium panel) and dipole moment (bottom panel) with NAO and Dunning GTO basis sets w.r.t. MRCHEM for the GGA-PBE and hybrid-PBE0 functionals, respectively.^{||} For all the molecules, the correct ground-state spin multiplicity was specified.

Total energies are a measure of the accuracy achieved by each method/basis pair, whereas the atomization energies deserve special attention for their role in the development of density functionals, generally benchmarked against such thermodynamic values. However, as recently pointed out by Medvedev *et al.*³³ the variational energy is not the optimal measure

^{||}Due to technical reasons in convergence, $\text{CH}_3\text{CH}_2\text{O}$ was excluded from the PBE0 results, while CCH was excluded in both PBE and PBE0.

for the quality of the calculated electronic density, which influences numerous other observables. For this reason, we have included the dipole moment as a non-variational quantity in our benchmarks (dipole errors are linear in the density error, whereas energies are quadratic). Dipole moments also serve as a verification that the different methods converge to the same electronic state and not to a nearby metastable configuration. Although the existence of multiple metastable SCF solutions in Kohn-Sham DFT is well known, it is often not detected by users of electronic structure codes. The solution strategy, also employed in the present paper, is to probe different spin initializations of each molecule to identify the global minimum. In the present work, the correct identification has been validated by ensuring consistency of the dipole moment as well as the KS eigenvalue spectra produced by the three distinct electronic structure methods.

Several important conclusions can be drawn from the results obtained:

1. For total energies, FHI-AIMS with NAOs is able to reach more accurate results than NWCHEM with GTOs, for basis sets of comparable size (e.g. “tier4” vs. aug-cc-pV5Z).
2. For atomization energies, both NAOs and GTOs benefit from error cancellation to some extent. Such a cancellation is however much stronger for GTOs where the RMSD is lowered by a factor 4-8 in most cases, whereas for NAOs only by a factor of 1,5-2. In both cases the cancellation is more marked for the smallest bases. Despite the smaller cancellations, NAOs are still closer to the converged limit than GTOs, for comparable basis sets.
3. The two functionals considered (GGA-PBE and hybrid-PBE0) yield very similar results, and we therefore assume that our conclusions concerning the accuracy of the different approaches (NAOs and GTOs) will hold also for other functionals of the same type.

- Dipole moments can be considered accurate if deviations are below 0.01 Debye.^{24,86} Only the largest basis sets in NAO and GTO used in our calculations achieve this target on average, but even such basis sets have outliers with errors close to 0.1 Debye.
- Due to the cumbersome convergence of periodic DFT codes with respect to the box size, we did not include APW+lo results for the entire test set of molecules. Nevertheless, for a small subset of molecules for which the limit of the box size was reached, we found atomization energies with errors of about 1 kcal/mol (see SI⁷⁹). Our experience suggests that it is technically challenging for APW-based codes to reach accuracies below 1 kcal/mol on atomization energies.

As a final remark, we stress that for a few atoms (Li, Be, Na, Mg), the aug-cc-pV5Z basis is not available, as previously mentioned in the atomic calculations part. Had it been available, GTOs might have yielded somewhat higher precision for the affected systems than in our benchmarks. However, considering the large size of our sample, the fact that only a few atoms in a molecule are affected, and the small improvement that can be inferred from the atomic calculations, our main conclusions still hold. On the other hand, such a *de facto* limitation of the availability of GTO basis sets illustrates how demanding it is to generate such basis sets. In contrast, MWs and NAOs are much less affected by such a limitation.

Considering that several families of GTO basis sets are available we have also performed an additional set of calculations for the GGA-PBE functional, in order to compare the performance across such sets. The results are displayed in Figure 3. In particular, we have considered the Pople basis sets 6-31+G** and 6-311++G(3df,3pd),⁸⁷⁻⁹⁰ the pc-2 and pc-3 basis sets⁹¹⁻⁹⁴ (the augmented analogs were also considered, but here we encountered ill-conditioning problems), and finally def2-TZVPD and def2-QZVPD.⁹⁵ This is a selection of the basis sets considered by Mardirossian and Head-Gordon

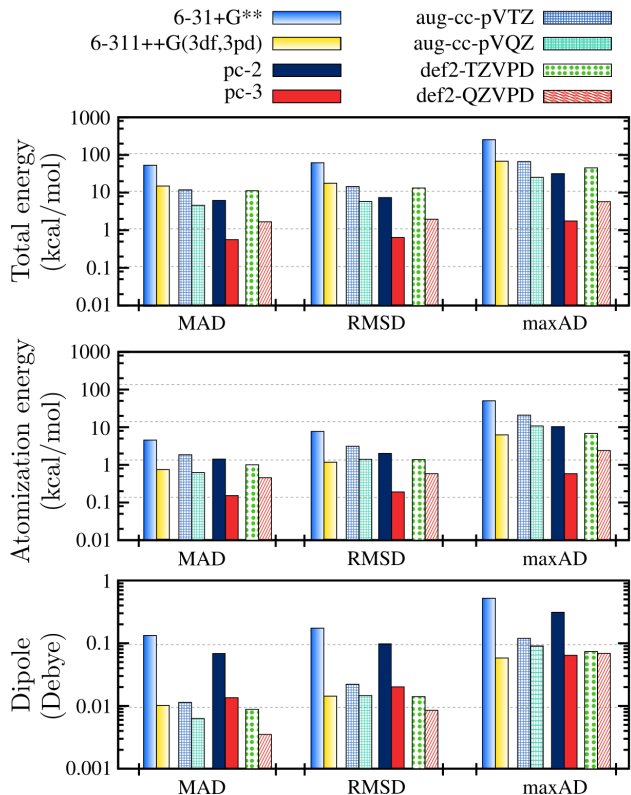


Figure 3: Statistics for total energy, atomization energy, and electrostatic dipole moment for the set of 211 molecules with respect to highly accurate values obtained using MRCHEM. Displayed values are mean absolute deviation (MAD), root mean square deviation (RMSD) and maximum absolute deviation (maxAD). The following basis sets have been considered: 6-31+G**, 6-311++G(3df,3pd), pc-2, pc-3, def2-TZVPD, def2-QZVPD, aug-cc-pVTZ and aug-cc-pVQZ.

for the development of the B97M-V and ω B97M-V functionals.^{17,18} For the comparison one should keep in mind that pc-2 and def2-TZVPD, as well as the large Pople set are comparable in size to aug-cc-pVTZ, while pc-3 and def2-QZVPD are comparable to aug-cc-pVQZ.

Although the detailed analysis of the performance of each basis set family is a relevant issue, it is outside the scope of this letter and will be considered in another study. Suffice it here to say that, among the considered basis sets, only pc-3 stands out when it comes to energy calculations, actually outperforming the much larger aug-cc-pV5Z basis, and competitive with NAOs. However, due to the lack of diffuse functions the

pc-n series suffers when dealing with dipole moments, where pc-3 performs only on par with the smaller aug-cc-pVTZ basis. The Pople sets are those which benefit the most from error cancellation when atomization energies are considered, while the def2-QZVPD basis yields better dipole moments, which is an indication of higher versatality in the basis. In practical calculations, the relatively weak performance of this broad range of production quality GTO basis sets for DFT-based **total and atomization energies (except pc-3)** can translate into real problems for subtle energy differences, e.g. conformational energy differences of hydrogen bonded systems.⁹⁶

To the best of our knowledge, this work presents the most accurate atomization energies calculated to date, for a large benchmark set of molecules. We conclude that moderately sized GTO basis sets, frequently used in quantum chemistry applications, suffer from average total energy errors much larger than 10 kcal/mol, and while very large GTO basis sets yield the desired accuracy on average, there are still significant outliers. Moreover, it may not always be feasible to employ such basis sets for systems much larger or chemically more diverse than those included in this study. In addition to cost, ill-conditioning can be a practical limitation for large GTO basis sets. Applying large, high-accuracy GTO basis sets to all-electron calculations for elements beyond the lightest few (e.g., $Z=1-18$ as covered here) is also not straightforward. In contrast, there is no practical limitation to extend the NAO, APW, or MW paradigm to all-electron DFT calculations across the full periodic table, for which NAOs and APWs are routinely used.

NAOs give much better accuracy even for moderately large bases (“tight” and beyond) since they can be constructed to possess the numerically correct behavior for a given XC functional, both in the nuclear as well as in the tail region. When feasible, APW+lo-based calculations achieve errors around 1 mHa for total energies, and 1 kcal/mol for atomization energies. However, this level of convergence is difficult to reach for general molecular systems.

With our MW results as reference,⁷⁹ it will be possible to unambiguously assess the accuracy

of any given basis for the computation of total energies and atomization energies. This will help to shed light on the quality of the currently available basis sets, and the underlying reasons for their shortcomings. It will also guide towards the development of more accurate basis sets.

Another central conclusion of our work is that the basis set error can dominate over errors arising from the choice of XC functional under many circumstances, in particular if some of the most advanced and accurate functionals are used. Our results set therefore new standards in the verification and validation of electronic structure methods. We expect that results of this work and the method described will be used to assess the accuracy of all future developments in Density Functional Theory methods.

Acknowledgement This research was partly supported by the Research Council of Norway through a Centre of Excellence Grant (Grant No. 179568/V30) with computing time provided by NOTUR (Grant No. NN4654K), and partly by the NCCR MARVEL, funded by the Swiss National Science Foundation (SNSF) with computing time provided by CSCS under project s707. S. Saha acknowledges support from the SNSF. J.A.F.-L. acknowledges fruitful discussions with John Kay Dewhurst and Andris Gulans on the APW method.

References

- (1) Hohenberg, H.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864.
- (2) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- (3) Goerigk, L.; Grimme, S. A General Database for Main Group Thermochemistry, Kinetics, and Noncovalent Interactions - Assessment of Common and Reparameterized (meta-)GGA Density Functionals. *J. Chem. Theory Comput.* **2010**, *6*, 107–126.

- (4) Karton, A.; Tarnopolsky, A.; Lamere, J.-F.; Schatz, G. C.; Martin, J. M. L. Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.
- (5) Zhao, Y.; Truhlar, D. G. Design of Density Functionals that are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. *J. Phys. Chem. A* **2005**, *109*, 5656–5667.
- (6) Perdew, J.; Burke, K.; Ernzerhof, M. GGA Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (7) Perdew, J. P.; Schmidt, K. Jacob’s Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conf. Proc.* **2001**, *577*, 1–20.
- (8) Li, C.; Zheng, X.; Cohen, A. J.; Morisánchez, P.; Yang, W. Local Scaling Correction for Reducing Delocalization Error in Density Functional Approximations. *Phys. Rev. Lett.* **2015**, *114*, 053001.
- (9) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.
- (10) Su, N. Q.; Xu, X. Toward the Construction of Parameter-free Doubly Hybrid Density Functionals. *Int. J. Quantum Chem.* **2015**, *115*, 589–595.
- (11) Peverati, R.; Truhlar, D. G. Quest for a Universal Density Functional: The Accuracy of Density Functionals Across a Broad Spectrum of Databases in Chemistry and Physics. *Philos. Trans. R. Soc. London A: Math., Phys. and Eng. Sci.* **2014**, *372*, 20120476.
- (12) Mardirossian, N.; Head-Gordon, M. How Accurate are the Minnesota Density Functionals for Noncovalent Interactions, Isomerization Energies, Thermochemistry, and Barrier Heights Involving Molecules Composed of Main-Group Elements? *J. Chem. Theory Comput.* **2016**, *12*, 4303–4325, PMID: 27537680.
- (13) Singh, D. J.; Nordstrom, L. *Planewaves, Pseudopotentials, and the LAPW method*; Springer Science & Business Media, 2006.
- (14) Goedecker, S.; Maschke, K. Transferability of Pseudopotentials. *Phys. Rev. A* **1992**, *45*, 88.
- (15) Willand, A.; Kvashnin, Y. O.; Genovese, L.; Mayagoitia, Á. V.; Deb, A. K.; Sadeghi, A.; Deutsch, T.; Goedecker, S. Norm-conserving Pseudopotentials with Chemical Accuracy Compared to All-electron Calculations. *J. Chem. Phys.* **2013**, *138*, 104109.
- (16) Lejaeghere, K.; Bihlmayer, G.; Björkman, T.; Blaha, P.; Blügel, S.; Blum, V.; Caliste, D.; Castelli, I. E.; Clark, S. J.; Dal Corso, A. et al. Reproducibility in Density Functional Theory Calculations of Solids. *Science* **2016**, *351*, aad3000.
- (17) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A Combinatorially Optimized, Range-separated Hybrid, Meta-GGA Density Functional with VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144*, 214110.
- (18) Mardirossian, N.; Head-Gordon, M. Mapping the Genome of Meta-Generalized Gradient Approximation Density Functionals: the Search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.
- (19) Moncrieff, D.; Wilson, S. Computational Linear Dependence in Molecular Electronic Structure Calculations Using Universal Basis Sets. *Int. J. Quantum Chem.* **2005**, *101*, 363–371.
- (20) Kotochigova, S.; Levine, Z. H.; Shirley, E. L.; Stiles, M. D.; Clark, C. W. *Atomic reference data for electronic*

- structure calculations*; National Institute of Standards and Technology, Physics Laboratory, 1997.
- (21) Ballone, P.; Galli, G. Accurate Pseudopotential Local-Density-Approximation Computations for Neutral and Ionized Dimers of the IB and IIB Groups. *Phys. Rev. B* **1990**, *42*, 1112–1123.
 - (22) Becke, A. D. Density-Functional Thermochemistry. I. The Effect of the Exchange-only Gradient Correction. *J. Chem. Phys.* **1992**, *96*, 2155–2160.
 - (23) Daykov, I. P.; Arias, T. A.; Engeness, T. D. Robust ab initio Calculation of Condensed Matter: Transparent Convergence through Semicardinal Multiresolution Analysis. *Phys. Rev. Lett.* **2003**, *90*.
 - (24) NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, Editor: Russell D. Johnson III. See <http://cccbdb.nist.gov/>, 2015.
 - (25) Papas, B. N.; Schaefer, H. F. Concerning the Precision of Standard Density Functional Programs: Gaussian, Molpro, NWChem, Q-Chem and Gamess. *J. Mol. Struct.: THEOCHEM* **2006**, *768*, 175–181.
 - (26) Harrison, R. J.; Fann, G. I.; Yanai, T.; Beylkin, G. Multiresolution Quantum Chemistry in Multiwavelet Bases. *Comput. Sci. - ICCS 2003: Int. Conf. Proc.* **2003**, 103–110.
 - (27) Harrison, R. J.; Fann, G. I.; Yanai, T.; Gan, Z.; Beylkin, G. Multiresolution Quantum Chemistry: Basic Theory and Initial Applications. *J. Chem. Phys.* **2004**, *121*, 11587.
 - (28) Yanai, T.; Fann, G. I.; Gan, Z.; Harrison, R. J.; Beylkin, G. Multiresolution Quantum Chemistry in Multiwavelet Bases: Hartree–Fock Exchange. *J. Chem. Phys.* **2004**, *121*, 6680.
 - (29) Yanai, T.; Fann, G. I.; Gan, Z.; Harrison, R. J.; Beylkin, G. Multiresolution Quantum Chemistry in Multiwavelet Bases: Analytic Derivatives for Hartree–Fock and Density Functional Theory. *J. Chem. Phys.* **2004**, *121*, 2866.
 - (30) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200.
 - (31) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods Without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
 - (32) Paier, J.; Hirschl, R.; Marsman, M.; Kresse, G. The Perdew–Burke–Ernzerhof Exchange–Correlation Functional Applied to the G2-1 Test Set Using a Plane-Wave Basis Set. *J. Chem. Phys.* **2005**, *122*, 234102.
 - (33) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Density Functional Theory is Straying From the Path Toward the Exact Functional. *Science* **2017**, *355*, 49–52.
 - (34) Pask, J. E.; Sterne, P. A. Finite Element Methods in ab initio Electronic Structure Calculations. *Modelling and Simulation in Materials Science and Engineering* **2005**, *13*, R71.
 - (35) Losilla, S. A.; Sundholm, D. A Divide and Conquer Real-space Approach for All-electron Molecular Electrostatic Potentials and Interaction Energies. *J. Chem. Phys.* **2012**, *136*, 214104.
 - (36) Toivanen, E. A.; Losilla, S. A.; Sundholm, D. The Grid-based Fast Multipole Method - a Massively Parallel Numerical Scheme for Calculating Two-electron Interaction Energies. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31480–31490.
 - (37) Parkkinen, P.; Losilla, S. A.; Solala, E.; Toivanen, E. A.; Xu, W.-H.; Sundholm, D.

- A Generalized Grid-Based Fast Multipole Method for Integrating Helmholtz Kernels. *J. Chem. Theory Comput.* **2017**, DOI:10.1021/acs.jctc.6b01207.
- (38) Lin, L.; Lu, J.; Ying, L.; E, W. Adaptive Local Basis Set for Kohn–Sham Density Functional Theory in a Discontinuous Galerkin Framework I: Total Energy Calculation. *J. Comput. Phys.* **2012**, *231*, 2140–2154.
- (39) Frediani, L.; Sundholm, D. Real-space Numerical Grid Methods in Quantum Chemistry. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31357–31359.
- (40) Sundholm, D.; Olsen, J. Finite-Element Multiconfiguration Hartree-Fock Calculations of the Atomic Quadrupole Moments of $C^+(^2P)$ and $Ne^+(^2P)$. *Phys. Rev. A* **1994**, *49*, 3453–3456.
- (41) Alpert, B. K.; Beylkin, G.; Coifman, R. Wavelets for the Fast Solution of Second-kind Integral Equations. *SIAM J. Sci. Stat. Comput.* **1993**, *14*, 159–174.
- (42) Alpert, B. K. A Class of Bases in L^2 for the Sparse Representation of Integral Operators. *SIAM J. Math. Analysis* **1999**, *24*, 246–262.
- (43) Beylkin, G.; Mohlenkamp, M. Numerical Operator Calculus in Higher Dimensions. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10246.
- (44) Beylkin, G.; Cramer, R.; Fann, G.; Harrison, R. J. Multiresolution Separated Representations of Singular and Weakly Singular Operators. *Appl. Comput. Harmon. A.* **2007**, *23*, 235–253.
- (45) Fann, G.; Beylkin, G.; Harrison, R. J.; Jordan, K. E. Singular Operators in Multiwavelet Bases. *IBM J. Research and Development* **2004**, *48*, 161–171.
- (46) Beylkin, G. On Approximation of Functions by Exponential Sums. *Appl. Comput. Harmon. A.* **2005**, *19*, 17–48.
- (47) Gines, D.; Beylkin, G.; Dunn, J. LU Factorization of Non-standard Forms and Direct Multiresolution Solvers 1. *Appl. Comput. Harmon. A.* **1998**, *5*, 156–201.
- (48) Beylkin, G.; Cheruvu, V.; Perez, F. Fast Adaptive Algorithms in the Non-standard Form for Multidimensional Problems. *Appl. Comput. Harmon. A.* **2008**, *24*, 354–377.
- (49) Kalos, M. H. Monte Carlo Calculations of the Ground State of Three- and Four-body Nuclei. *Phys. Rev.* **1962**, *128*, 1791.
- (50) Harrison, R. J. Krylov Subspace Accelerated Inexact Newton Method for Linear and Nonlinear Equations. *J. Comput. Chem.* **2004**, *25*, 328–334.
- (51) Yanai, T.; Harrison, R. J.; Handy, N. C. Multiresolution Quantum Chemistry in Multiwavelet Bases: Time-dependent Density Functional Theory with Asymptotically Corrected Potentials in Local Density and Generalized Gradient Approximations. *Mol. Phys.* **2005**, *103*, 413–424.
- (52) Yanai, T.; Fann, G. I.; Beylkin, G.; Harrison, R. J. Multiresolution Quantum Chemistry in Multiwavelet Bases: Excited States from Time-dependent Hartree-Fock and Density Functional Theory via Linear Response. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31405–31416.
- (53) Kottmann, J. S.; Höfener, S.; Bischoff, F. A. Numerically Accurate Linear Response-Properties in the Configuration-Interaction Singles (CIS) Approximation. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31453–31462.
- (54) Sekino, H.; Maeda, Y.; Yanai, T.; Harrison, R. J. Basis Set Limit Hartree-Fock and Density Functional Theory Response Property Evaluation by Multiresolution Multiwavelet Basis. *J. Chem. Phys.* **2008**, *129*, 034111.

- (55) Sekino, H.; Yokoi, Y.; Harrison, R. J. A New Implementation of Dynamic Polarizability Evaluation Using a Multiresolution Multiwavelet Basis Set. *J. Phys.: Conf. Series* **2012**, *352*, 012014.
- (56) Jensen, S. R.; Flå, T.; Jonsson, D.; Monstad, R. S.; Ruud, K.; Frediani, L. Magnetic Properties with Multiwavelets and DFT: The Complete Basis Set Limit Achieved. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21145–21161.
- (57) Marques, M. A.; Oliveira, M. J.; Burnus, T. Libxc: A Library of Exchange and Correlation Functionals for Density Functional Theory. *Comput. Phys. Comm.* **2012**, *183*, 2272–2281.
- (58) Ekström, U.; Visscher, L.; Bast, R.; Thorvaldsen, A. J.; Ruud, K. Arbitrary-Order Density Functional Response Theory from Automatic Differentiation. *J. Chem. Theory Comput.* **2010**, *6*, 1971–1980.
- (59) Alpert, B. K.; Beylkin, G.; Gines, D.; Vozovoi, L. Adaptive Solution of Partial Differential Equations in Multiwavelet Bases. *J. Comput. Phys.* **2002**, *182*, 149–190.
- (60) MultiResolution Chemistry (MRChem) program package. See <http://mrchemdoc.readthedocs.org/en/latest/>, 2016.
- (61) Feller, D. The Role of Databases in Support of Computational Chemistry Calculations. *J. Comput. Chem.* **1996**, *17*, 1571–1586.
- (62) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. Basis Set Exchange: A Community Database for Computational Sciences. *J. Chem. Information and Modeling* **2007**, *47*, 1045–1052.
- (63) Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; van Dam, H.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. et al. NWChem: A Comprehensive and Scalable Open-source Solution for Large Scale Molecular Simulations. *Comput. Phys. Comm.* **2010**, *181*, 1477.
- (64) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio Molecular Simulations with Numeric Atom-centered Orbitals. *Comput. Phys. Comm.* **2009**, *180*, 2175.
- (65) Ren, X.; Rinke, P.; Blum, V.; Wieferink, J.; Tkatchenko, A.; Sanfilippo, A.; Reuter, K.; Scheffler, M. Resolution-of-identity Approach to Hartree-Fock, Hybrid Density Functionals, RPA, MP2 and GW with Numeric Atom-centered Orbital Basis Functions. *New Journal of Physics* **2012**, *14*, 053020.
- (66) An All-electron Full-Potential Linearised Augmented-Plane Wave (FP-LAPW) Code. See <http://elk.sourceforge.net/>, 2015.
- (67) Blaha, P.; Schwarz, K.; Madsen, G.; Kvasnicka, D.; Luitz, J. *WIEN2K: An Augmented Plane Wave Plus Local Orbitals Program for Calculating Crystal Properties*, Rev. Ed. *WIEN2K 13.1*; Vienna University of Technology, Austria, 2001.
- (68) Gulans, A.; Kontur, S.; Meisenbichler, C.; Nabok, D.; Pavone, P.; Rigamonti, S.; Sagmeister, S.; Werner, U.; Draxl, C. Exciting: a Full-potential All-electron Package Implementing Density-Functional Theory and Many-body Perturbation Theory. *J. Phys.: Condensed Matter* **2014**, *26*, 363202.
- (69) Delley, B. An All-electron Numerical Method for Solving the Local Density Functional for Polyatomic Molecules. *J. Chem. Phys.* **1990**, *92*, 508–517.
- (70) Delley, B. From Molecules to Solids With the DMol3 Approach. *J. Chem. Phys.* **2000**, *113*, 7756–7764.
- (71) Koepernik, K.; Eschrig, H. Full-potential Nonorthogonal Local-orbital Minimum-basis Band-structure Scheme. *Phys. Rev. B* **1999**, *59*, 1743–1757.

- (72) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.
- (73) Kenny, S.; Horsfield, A. Plato: A Localised Orbital Based Density Functional Theory Code. *Comput. Phys. Comm.* **2009**, *180*, 2616 – 2621.
- (74) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian 16 Revision A.03. 2016; Gaussian Inc. Wallingford CT.
- (75) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. et al. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* *14*, 1347–1363.
- (76) Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P. et al. The Dalton Quantum Chemistry Program System. *WIREs Comput. Mol. Sci.* **2013**, *4*, 269–284.
- (77) Aquilante, F.; De Vico, L.; Ferre, N.; Ghigo, G.; Malmqvist, P.-A.; Neogrady, P.; Pedersen, T. B.; Pitonak, M.; Reiher, M.; Roos, B. O. et al. Software News and Update MOLCAS 7: The Next Generation. *J. Comput. Chem.* **2010**, *31*, 224–247.
- (78) Harrison, R. J.; Beylkin, G.; Bischoff, F. A.; Calvin, J. A.; Fann, G. I.; Fosso-Tande, J.; Galindo, D.; Hammond, J. R.; Hartman-Baker, R.; Hill, J. C. et al. MADNESS: A Multiresolution, Adaptive Numerical Environment for Scientific Simulation. *SIAM J. Sci. Comput.* **2016**, *38*, S123–S142.
- (79) Jensen, S. R.; Saha, S.; Flores-Livas, J. A.; Huhn, W.; Blum, V.; Goedecker, S.; Frediani, L. GGA-PBE and Hybrid-PBE0 Energies and Dipole Moments with MRChem, FHI-aims, NWChem and ELK. 2017; <http://dx.doi.org/10.18710/OEMOEL>.
- (80) Schmider, H. L.; Becke, A. D. Optimized Density Functionals from the Extended G2 Test Set. *J. Chem. Phys.* **1998**, *108*, 9624–9631.
- (81) Jensen, F. Atomic Orbital Basis Sets. *WIREs Comput. Mol. Sci.* **2013**, *3*, 273–295.
- (82) Woon, D. E.; Dunning, T. H., Jr Gaussian Basis Sets for Use in Correlated Molecular Calculations V. Core-valence Basis Sets for Boron Through Neon. *J. Chem. Phys.* **1995**,
- (83) Woon, D. E.; Dunning, T. H. Gaussian-Basis Sets for Use in Correlated Molecular Calculations III. The Atoms Aluminum Through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- (84) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron-Affinities of the 1st-Row Atoms Revisited - Systematic Basis-Sets and Wave-Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (85) Dunning, T. H. Gaussian-Basis Sets for Use in Correlated Molecular Calculations I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (86) Bak, K. L.; Gauss, J.; Helgaker, T.; Jørgensen, P.; Olsen, J. The Accuracy of Molecular Dipole Moments in Standard Electronic Structure Calculations. *Chem. Phys. Lett.* **2000**, *319*, 563–568.
- (87) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; De-frees, D. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods XXIII. A Polarization-Type Basis Set for 2nd-Row Elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.

- (88) Dill, J. D.; Pople, J. A. Self-Consistent Molecular Orbital Methods XV. Extended Gaussian-type Basis Sets for Lithium, Beryllium and Boron. *J. Chem. Phys.* **1975**,
- (89) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J Chem Phys* **1972**, *56*, 2257–2261.
- (90) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-Consistent Molecular-Orbital Methods XX. Basis Set for Correlated Wave-Functions. *J. Chem. Phys.* **1980**, *72*, 650–654.
- (91) Jensen, F. Polarization Consistent Basis Sets IV: The Elements He, Li, Be, B, Ne, Na, Mg, Al, and Ar. *J. Phys. Chem. A* **2007**, *111*, 11198–11204.
- (92) Jensen, F.; Helgaker, T. Polarization Consistent Basis Sets V. The Elements Si-Cl. *J. Chem. Phys.* **2004**, *121*, 3463–3470.
- (93) Jensen, F. Polarization Consistent Basis Sets II. Estimating the Kohn-Sham Basis Set Limit. *J. Chem. Phys.* **2002**, *116*, 7372–7379.
- (94) Jensen, F. Polarization Consistent Basis Sets: Principles. *J. Chem. Phys.* **2001**, *115*, 9113–9125.
- (95) Rappoport, D.; Furche, F. Property-optimized Gaussian Basis Sets for Molecular Response Calculations. *J. Chem. Phys.* **2010**, *133*, 134105.
- (96) Chutia, S.; Rossi, M.; Blum, V. Water Adsorption at Two Unsolvated Peptides with a Protonated Lysine Residue: From Self-Solvation to Solvation. *J. Phys. Chem. B* **2012**, *116*, 14788–14804, PMID: 23171405.