JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

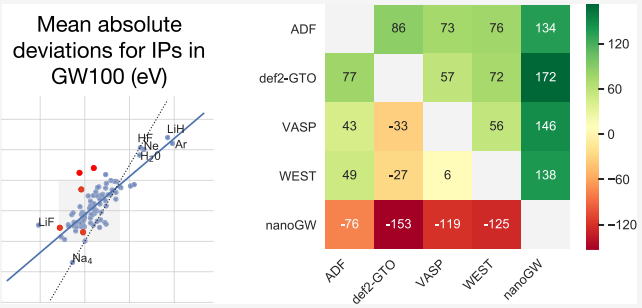
GW100: A Slater-Type Orbital Perspective

Arno Förster* and Lucas Visscher

Cite This: J. Chem. Theory Comput. 2021, 17, 5080−5097 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We calculate complete basis set (CBS) limit-extrapolated ionization potentials (IPs) and electron affinities (EA) with Slater-type basis sets for the molecules in the GW100 database. To this end, we present two new Slater-type orbital (STO) basis sets of triple-(TZ) and quadruple-ζ (QZ) quality, whose polarization is adequate for correlated-electron methods and which contain extra diffuse functions to be able to correctly calculate EAs of molecules with a positive lowest unoccupied molecular orbital (LUMO). We demonstrate that going from TZ to QZ quality consistently reduces the basis set error of our computed IPs and EAs, and we conclude that a good estimate of these quantities at the CBS limit can be obtained by extrapolation. With mean absolute deviations (MAD) from 70 to 85 meV, our CBS limit-extrapolated IP are in good agreement with results from FHI-AIMS, TURBOMOLE, VASP, and WEST, while they differ by more than 130 meV on average from nanoGW. With a MAD of 160 meV, our EA are also in good agreement with the WEST code. Especially for systems with positive LUMOs, the agreement is excellent. With respect to other codes, the STO-type basis sets generally underestimate EAs of small molecules with strongly bound LUMOs. With 62 meV for IPs and 93 meV for EAs, we find much better agreement with CBS limit-extrapolated results from FHI-AIMS for a set of 250 medium to large organic molecules.



1. INTRODUCTION

Over the last years, the *GW* approximation (GWA)[1] has become a popular method to calculate charged excitations in finite systems.[2−9] While formally more rigorous, fully self-consistent *GW* (sc*GW*) calculations are relatively expensive and not necessarily very accurate for the calculation of ionization potentials (IPs) and electron affinities (EAs) for molecular systems.[8,10] Instead, the perturbative G_0W_0 approach is often the preferred alternative since it is computationally less demanding than sc*GW* and can give significantly more accurate quasiparticle (QP) energies, provided that it is based on a suitable reference.[5,8,10]

Ideally, the result of a G_0W_0 calculation should be independent of the particular implementation of the method. In practice, however, choices regarding the numerical representation of the involved quantities must be made, including the choice of a single-particle basis as well as a discretization of frequency and/or time variables. The choice of the single-particle basis also entails a choice regarding the representation of the core electrons and a treatment of virtual states. Both factors are decisive since it is known that core correlation plays a major role in G_0W_0 calculations[7] but also since *GW* QP energies converge very slowly to the complete basis set (CBS) limit.[7,9,11,12] For these reasons, achieving consensus between different G_0W_0 codes is challenging and requires careful convergence of a calculation with respect to all technical parameters. Due to limited resources and/or time constraints, it might not always be possible in applications to only work with converged parameters. In that case, one would like to know how a certain technical parameter affects the final result.

For these reasons, comparison between different codes through systematic benchmarks is highly desirable. First, it allows to verify that the results from these codes agree within a reasonable margin of error. For G_0W_0 calculations, one usually aims for an accuracy of 100 meV. Second, such benchmarks are crucial in order to quantify the influence of the various technical parameters on the QP energies. Significant efforts in this direction have been initiated by van Setten et al.[7] in 2015 with the publication of the GW100 database for finite systems. In their work, van Setten et al. compared the IPs and EAs of 100 small- and medium-sized molecules on the G_0W_0@PBE level of theory, calculated with three different codes, the Gaussian-type orbital (GTO)-based all-electron code TURBOMOLE,[6,13] the numerical atomic orbital (NAO)-based all-electron code FHI-AIMS,[4,14,15] and the plane wave (PW) code BerkeleyGW.[16,17] Later, benchmarks for many more codes followed, including the PW implementations in VASP[18−21] in 2017[22] and WEST[23] in 2018[24] and the real-space finite-

Received: March 30, 2021
Published: July 8, 2021

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

element (RSFE) implementation in nanoGW[25] in 2019.[26] Also, the accuracy of many low-order scaling implementations of the G_0W_0 method was benchmarked against the GW100 database.[27−30]

These studies established the choice of single-particle basis as a crucial factor causing major differences between different implementations. For instance, the results from TURBOMOLE, FHI-AIMS, and MOLGW[31] and also from the low-scaling implementations by Wilhelm et al.[28] in CP2K[32] and by Duchemin and Blase,[30] all using the same def2-GTO type basis sets, agree within a few 10 meV on average for GW100, even though these implementations differ in frequency treatment as well as calculation of four-center integrals. The differences between codes using different basis sets are considerably larger. The discrepancy between the TURBOMOLE and BerkeleyGW results of nearly 300 meV on average reported in ref 7 for the highest occupied molecular orbital (HOMO) was not necessarily insightful since the BerkeleyGW results were not CBS limit-extrapolated. With only around 60 meV on average, the agreement between the CBS limit-extrapolated PW results obtained with VASP and TURBOMOLE was found to be significantly better.[22] However, for EAs, the disagreement between different codes is considerably larger and differences for systems with a positive lowest unoccupied molecular orbital (LUMO) can easily exceed several eV. It has also been pointed out in ref 22 that the type of the GTO-type basis set has a major influence on these EAs and that Dunning's correlation consistent basis sets are more suitable than the def2-series, which has been used in ref 7. Besides the choice of the basis set, the treatment of core electrons (pseudopotentials vs all-electron) also plays a decisive role for many systems.[24]

Against this background, it seems that further benchmark results for GW100 using different basis-set types are useful to advance the current understanding of the dependence of *GW* QP energies on the basis set type. Recently, we presented the first production-level implementation of the G_0W_0 method with Slater-type orbitals (STOs) in the Amsterdam density functional (ADF) module of the Amsterdam modeling suite (AMS).[33] The G_0W_0 implementation in ADF is detailed in ref 29 and is based on the space-time formulation of the *GW* method first proposed by Godby and coworkers.[34,35] It uses nonuniform imaginary time and imaginary frequency grids tailored to the system under investigation[21,36,37] and treats four-point correlation functions with the pair-atomic density fitting (PADF) approximation, resulting in a low-order implementation with a very low prefactor.[28] To demonstrate the correctness of our implementation, we already presented IPs and EAs for the GW100 database in ref 29. However, our results did not allow a meaningful comparison to implementations with other basis set types. Since it is not always straightforward to obtain converged minimax grids,[27,29,36] we used rather small imaginary frequency and time grids which were often not converged. Most importantly, in ref 29, we used standard Slater-type basis sets optimized for independent-electron methods,[38] which do not allow a systematic extrapolation to the CBS limit for correlated methods.[39,40]

To be able to obtain accurate, CBS limit-extrapolated IPs and EAs using STOs, we report here improvements over our original implementation regarding both parameters. For once, we implemented improved imaginary time and frequency grids, which allow a systematic convergence to the limit of an infinite number of grid points. Second, and most importantly, we designed two new Slater-type basis sets for all elements of the periodic table, which we call (aug)-TZ3P and (aug)-QZ6P. They contain extra-diffuse functions to be able to obtain accurate EAs for systems with a LUMO above the vacuum level, and in the choice of polarization functions, we follow the requirements of correlated-electron methods[40,41] as closely as possible: ADF only supports basis functions with angular momenta up to $l = 3$ which is a clear restriction since already for second row atoms, a consistent polarization on the quadruple-ζ (QZ) level requires angular momenta up to $l = 4$ and higher angular momentum functions are necessary for heavier elements.[40] For simplicity, we will refer to these basis sets as correlation-consistent. However, we emphasize that they are not correlation-consistent in a strict sense but rather as correlation-consistent as possible in our current implementation.

Despite these restrictions, as we will demonstrate in the following by comparison to results from other codes, the new basis sets allow for a reliable extrapolation to the CBS limit. Consequently, we present here CBS limit-extrapolated IPs and EAs for the GW100 database with STOs. These results are the focus of the current work and are meant to complement the previous studies on GW100 using GTOs, PWs, and RSFEs. Since G_0W_0@PBE generally do not give accurate QP energies,[8,10,42] we will only focus on numerical aspects and abstain from comparison to experimental or high-level quantum chemistry reference values.[43,44] To complement our results, we also calculate IPs and EAs of 250 molecules from the GW5000 database.[11] This work is organized as follows: in Section 2, we shortly outline the G_0W_0 implementation in ADF and describe our new basis sets. We then present our IPs and EAs for GW100 and point out similarities and differences to other codes in Section 3. Finally, Section 4 summarizes and concludes this work.

2. THEORY

We start this section by shortly reviewing the G_0W_0 approximation (in the following, we refer to this approximation as GWA for simplicity) and describe the main features of the implementation in the ADF. For more details, we refer to ref 29. Here, we mostly focus on the two factors which are the most decisive for this work: the newly designed basis sets and the treatment of imaginary frequency and imaginary time.

2.1. *GW* Approximation. The *GW* approximation is an approximation to the self-energy Σ, which appears in Dyson's equation.[45] It is defined as the difference between a dynamical one-body Hamiltonian H and a static reference Hamiltonian h[46] corresponding to a noninteracting Fermi system (in practice, h will be the Hamiltonian of a generalized[47] KS[48−50] problem)

$$\sum(1, 2) = H(1, 2) - h(1, 2) \tag{1}$$

H is defined by $H(\omega) = \omega - G(\omega)^{-1}$, where G is Green's function of the interacting Fermi system. $1 = (\mathbf{r}_1,\sigma_1,\omega_1)$ collects space, spin, and frequency coordinates of a particle, but, as we do not consider spin−orbit coupling in the current work, we will omit spin for simplicity in the following. If one approximates Σ as diagonal and real and transforms to a molecular orbital (MO) representation, Dyson's equation becomes

$$\omega_n = \langle n|h|n\rangle + \langle n|Re(\sum(\omega_n)|n\rangle \tag{2}$$

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

where n labels a single-particle state. In case n refers to the HOMO (LUMO) level, $-\omega_n$ is equal to the IP (EA) of the system. In the GWA

$$\sum(1, 2) = iG_0(1, 2)*W_0(1, 2) - v_{xc}(1, 2) \tag{3}$$

where * denotes convolution, v_{xc} is the exchange−correlation potential of generalized KS-DFT, W_0 is the dynamically screened Coulomb interaction, which is related to the bare Coulomb interaction V via a Dyson equation

$$W_0(1, 2) = V(1, 2) + V(1, 3)P_0(3, 4)W_0(4, 2) \tag{4}$$

and

$$G_0(\omega) = (\omega - h)^{-1} \tag{5}$$

is the one-body Green's function corresponding to the noninteracting reference system. Integration over repeated indices is implied. The kernel of this Dyson equation, the polarizability P_0, is in turn calculated from the noninteracting Green's function in the random phase approximation (RPA)[46]

$$P_0(1, 2) = -iG_0(1, 2)*G_0(2, 1) \tag{6}$$

In principle, the RPA is distinct from the GWA, but when referring to the *GW* method in a quantum chemistry context, it is usually implied that the RPA is made for P_0.

The G_0W_0 method in ADF is implemented in the space-time formalism introduced by Godby and co-workers.[34,35] This means that eq 4 is solved in imaginary frequency, while P and Σ are calculated in imaginary time. From the imaginary time domain, Σ is brought to the imaginary frequency axis. Since Σ is analytic in the upper-half plane, it can be analytically continued to the real axis. In practice, this requires discrete grids to sample the imaginary time and imaginary frequency axes. In the absence of numerical noise, we can use our knowledge of the self-energy on N_ω points on the imaginary frequency axis to interpolate the input data with a rational function using the Padé technique, as outlined by Vidberg et al.[51] The crucial assumption in this approach is that knowledge of Σ at a small number of points on the imaginary frequency axis allows for an accurate interpolation. In case the single QP picture is valid, this assumption is reasonable for states in the vicinity to HOMO and LUMO energy since the interacting Green's function qualitatively has the same structure as the noninteracting one (see e.g., refs 52 and 53), and thus, the self-energy (by definition) will be smooth.[54] The G_0W_0 implementation in ADF follows the work flow shown in Figure 1.

2.2. Imaginary Time and Frequency Grids. To discretize imaginary frequency and imaginary time in a suitable way, we follow Kresse and co-workers (building on earlier work by Almlöf[55] and Hackbusch and co-workers[56]) and use nonuniform grids for both domains and switch between them

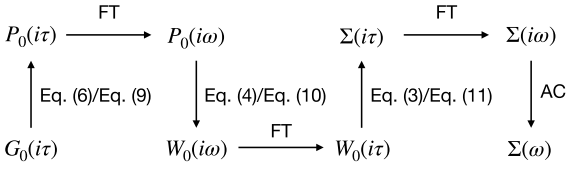


Figure 1. Schematic work flow for the calculation of the self-energy in the whole complex plane in ADF. FT denotes Fourier transform and AC denotes analytical continuation.

using nonuniform Fourier transforms (FTs).[21,36,37] The grid points are selected at run time and are tailored to the system under investigation. Our implementation of these grids closely follows Kresse and co-workers[21,36,37] and is outlined in Appendix A.

Similar grids have also been used by other researchers[27,28,30] to implement the G_0W_0 method for finite systems in the space-time formalism. Recently, Wilhelm et al. benchmarked the convergence of QP energies in the GW100 database with respect to the grid sizes. In our older work,[29] our frequency grids were restricted to not more than 19 grid points, and ref 28 clearly shows such grids to be insufficient to converge all IPs and EAs in GW100. Consequently, we calculated new frequency grids which allow us to use up to 40 imaginary frequency points, which ensures that the results are converged with respect to this parameter.[28] In ref 28, Wilhelm et al. could also reproduce the results obtained by van Setten et al.[7] with the TURBOMOLE code with an accuracy of a few millielectron volts (meV). Since TURBOMOLE performs the frequency integration fully analytically, we conclude that the frequency integration is a numerical parameter which is well under control in our updated implementation.

2.3. Discretization of Real Space. To discretize real space, we use a basis of N_{AO} STOs, $\mathcal{S} = \{\chi_\mu\}_{\mu=1,\ldots,N_{AO}}$, which are related to the MO ϕ by

$$\phi_n = \sum_\mu b_{n\mu}\chi_\mu(\boldsymbol{r}) \tag{7}$$

All quantities which occur in eq 4 are four-point correlation functions, and to this end, solution of eq 4 would scale as N^6 when expanding these quantities in terms of $\mathcal{S}$. In the density-fitting (DF) approximation,[57−63] one approximates the product basis of atomic orbitals $\mathcal{P} = \mathcal{S} \otimes \mathcal{S}$ with an auxiliary basis $\mathcal{A} = \{f_\alpha\}_{\alpha=1,\ldots,N_{aux}}$, where N_{aux} grows linearly with system size and introduces a basis transformation

$$C: \mathcal{P} \to \mathcal{A}, \qquad \chi_\mu(\boldsymbol{r})\chi_\nu(\boldsymbol{r}) = \sum_p c_{\mu\nu p}f_p(\boldsymbol{r}) \tag{8}$$

The equations which need to be solved now are

$$P_{pq,\tau} = c_{\mu\nu p}P_{\mu\nu\kappa\lambda,\tau}c_{\kappa\lambda q} = -ic_{\mu\nu p}[G_0]_{\mu\kappa,\tau}[G_0]_{\nu\lambda,\tau}c_{\kappa\lambda q} \tag{9}$$

$$[W_0]_{pq,\omega} = V_{pq} + V_{pr}[P_0]_{rs,\omega}[W_0]_{sq,\omega} \tag{10}$$

$$[\Sigma_{xc}]_{\mu\nu,\tau} = i\sum_{\kappa\lambda}\sum_{pq}[G_0]_{\kappa\lambda,\tau}c_{\mu\kappa q}[W_0]_{pq,\tau}c_{\nu\lambda q} \tag{11}$$

The computational bottlenecks in these equations are now the basis transformations in eqs 9 and 11, which both scale as N^4 with system size. However, when the map C from eq 8 is constrained such that the number of nonzero-elements in C only grows linearly with system size, both equations only require N^2 operations.[27,29] PADF is a way to achieve this goal by the way of restricting the density fit to pairwise sums only and by subsequently introducing distance cutoffs. Also, other schemes to introduce sparsity into C have been applied successfully to achieve low-order scaling G_0W_0 implementations[27,28] (we refer to ref 28 for a detailed discussion of the pros and cons of different techniques), but PADF is arguably the best practical method to make map C as sparse as possible. Constructing C in that way only requires evaluation of a small number of integrals compared to other approaches,[58] which makes this approach especially suitable for localized basis

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

functions, for which Coulomb integrals cannot be calculated analytically, such as STOs or NAOs. In fact, the computed time spent on evaluating the remaining integrals in our implementation is negligible compared to the time needed to complete a G_0W_0 calculation.

In PADF, only the fit functions centered on atoms A and B are used to expand a product of AOs, which are centered on atoms A and B. This translates into a matrix $\mathbf{c}$, for which $c_{\mu\nu p} \neq 0$ only if $\mu, \nu, p \in (A, B)$. One further introduces a cutoff so that for 2 AOs μ, ν centered on atoms far away from each other, all corresponding matrix elements $c_{\mu\nu p}$ are set to zero. Employing PADF, standard auxiliary basis sets designed for global density fitting are not suitable (they are simply too small), and dedicated fit sets are needed. For an in-depth discussion of challenges in the design of these fit sets for correlated-electron methods, we refer the reader to ref 29 and for a more explicit description of the fit sets, a list of exponents of the fit functions and more technical details to ref 64.

2.4. STO Basis Sets. In eq 7, μ is a composite index, collecting the five defining parameters of a Slater-type function

$$\chi_\mu(r, \theta, \phi) = \chi_{A,\alpha,n,l,m}(r, \theta, \phi) = C(\alpha, n) r^{n-1} e^{-\alpha r} Z_{lm}(\theta, \phi), \qquad r = |\mathbf{r} - \mathbf{R}_A| \tag{12}$$

the exponent α and the quantum numbers (n, l, m), as well as the nucleus A on which the function is centered. In eq 12, $C(\alpha, n)$ is a normalization constant and Z_{lm} denotes a real spherical harmonic. The main difference to Gaussian-type functions is in the dependence on r in the exponential instead of r^2, which mainly results in a different behavior close to the atomic nuclei and a slower decay for large r. While STO-type basis sets are well developed for independent-electron methods,[38,65−68] we are not aware of any attempt to construct general Slater-type basis sets which are consistent with the requirements of correlated-electron methods. Here, we make the first attempt to present such basis sets for the whole periodic table. The concept of correlation consistency has also been applied to NAO basis sets for RPA total-energy calculations.[69]

2.4.1. Correlation-Consistent Atomic Basis Sets. In the construction of our basis sets, we make two assumptions: first, we assume that the principles guiding the construction of GTO-type basis sets for correlated-electron methods which have been developed over the last decades should be applicable to STO-type basis sets as well. This can be justified as follows: based on numerical experiments on correlation-consistent GTO-type basis sets, Helgaker et al.[70] established the relation

$$E_\infty^{\mathrm{corr}} - E_X^{\mathrm{corr}} = aX^{-3} \tag{13}$$

between the correlation energy E_∞^{corr} at the CBS limit and the correlation energy E_X^{corr} calculated with a given correlation consistent basis set with cardinal number $X = l_{\max} + 1$. Conceptually, their formula is based on earlier work by Schwartz[71] and the mathematical, more rigorous work by Hill[72] on the convergence of the ground state of the helium atom in a full configuration interaction calculation with respect to the single-particle basis. Later on, Kutzelnigg and Morgan[73] generalized this result to arbitrary n-electron systems for MP2 calculations.

In principle, eq 13 is only valid in the limit of large X; however, there is numerical evidence that it is already a good approximation for $X = 3$ and $X = 4$.[39,70] Since eq 13 is independent of the type of localized basis functions,[71−73] this should also be the case for STO-type basis sets, provided that they are also constructed in a correlation-consistent fashion as first defined by Dunning[41] such that the total basis set incompleteness error is distributed equally between the different angular momentum functions. This requires that polarization functions are added in well-defined sequences.[40] There is numerical evidence[40] that the consistent polarization for correlated methods is $1(l_{\mathrm{occ}} + 1)$ on the double-ζ (DZ), $2(l_{\mathrm{occ}} + 1)1(l_{\mathrm{occ}} + 2)$ on the TZ, and $3(l_{\mathrm{occ}} + 1)2(l_{\mathrm{occ}} + 2)$ $(l_{\mathrm{occ}} + 3)$ on the QZ level, where l_{occ} denotes the angular momentum of the highest occupied orbital in an atom. This requirement is usually only followed strictly for the first three rows of the periodic table,[40] and the design of correlation-consistent basis sets for heavier elements might follow different principles.[74−78]

2.4.2. Construction of Correlation-Consistent Slater-Type Basis Sets. Based on the considerations above, we construct correlation-consistent basis sets of TZ and QZ quality. We name these basis sets TZ3P and QZ6P, respectively. We also augment these basis sets with additional diffuse functions, and we name these basis sets aug-TZ3P and aug-QZ6P, respectively. The acronym xP refers to the number of polarization functions we use for the elements of the first three rows of the periodic table. This choice is consistent with the requirements for correlation-consistent basis sets stated above.

At this point, we introduce our second assumption: Since ADF only supports basis functions up to $l = 3$, for second- and third-row elements, we chose the polarization 2d1f and 3d3f for TZ and QZ, respectively, and for consistency, also 2p1d and p3d3 for hydrogen and helium. A good justification for the validity of this approximation cannot be given, and as we will see later, our results indeed suggest that the replacement of g with another f function negatively affects the QP energies for small molecules. Generalization of ADF to accommodate use of higher angular momenta functions is therefore desirable.

The TZ2P and QZ4P basis sets which have been described in detail in ref 38 serve as starting points for our new basis sets. TZ2P (QZ4P) is of DZ (TZ) quality in the core region and of TZ (QZ) quality in the valence region. For the TZ3P and QZ6P basis sets, we chose not to optimize exponents but rather to add additional polarization functions in an ad hoc manner. This approach can be justified by the observation that precise values of exponents become less important as a basis set approaches completeness.[40,65] This is especially true for molecules, as opposed to isolated atoms. The latter case generally requires larger and more optimized basis sets than the molecular case. Instead, we rather focus on choosing the exponents in a way that their overlap is small and linear dependency problems are more likely to be avoided.

TZ3P is simply obtained by augmenting the TZ2P basis set by another $l_{\mathrm{occ}} + 1$-function for all elements. The exponent is chosen so that it is twice as large as the exponent of the $l_{\mathrm{occ}} + 1$-function in the TZ2P basis set. This is due to the fact that the exponents of the polarization functions in the TZ2P and QZ4P basis sets are chosen in a way that the basis sets become more accurate in the valence region, which is favorable for the calculation of bonding energies. The calculation of IPs also requires the accurate representation of the electron density closer to the core, and TZ3P should yield a major improvement over TZ2P in this respect. The same reasoning has also been followed in ref 79. In complete analogy, the

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

QZ6P basis set is obtained by adding an additional tight l_{occ} + 1-function and an l_{occ} + 2-function for all elements. For fluorine and chlorine, we added an additional shell of d-function since we found exceptionally slow convergence of QP energies to the CBS limit. This will be discussed in more detail in Section 3.

The exponents α_1, α_2 of the polarization functions in the QZ4P basis set fulfill $\alpha_1/\alpha_2 \approx 2$ for each l. Thus, the polarization functions in the TZ3P and QZ6P basis sets loosely follow an even-tempering scheme[80]

$$\alpha_i = \alpha_1 \beta^{i-1}, \qquad i = 2, \ ..., \ M, \qquad \beta = 2 \tag{14}$$

with M being equal to two (three) for l_{occ} + 1 in TZ3P (QZ6P) and one (three) for l_{occ} + 2. The value of β is rather large to avoid linear dependency problems, and in conjunction with a rather small α_1, it ensures that the exponents span a rather wide range of values in the QZ6P case.[81] Ren et al. recently used a similar reasoning to choose the exponents of Slater-type functions for G_0W_0 calculations for periodic systems.[82]

2.4.3. Adding Diffuse Functions. It is known from electron scattering experiments that molecular EAs are sometimes negative, that is, their anionic state at the geometry of the neutral molecule is unstable.[83−85] This corresponds to a positive LUMO QP energy, which formally corresponds to a non-normalizable continuum orbital. However, as an artifact of working with a finite basis, the orbital will always be constrained to be normalizable.[86] Very diffuse functions (such as plane waves) are then needed to mimic the continuum state, and for this reason, we augment our basis sets with additional diffuse functions (see also refs 87 and 88). The resulting basis sets are denoted as aug-TZ3P and aug-QZ6P and are obtained from TZ3P and QZ6P, respectively, by adding diffuse s, p, and d functions for all element types, with the important exceptions of first-row atoms, where we only add diffuse s and p functions.[89] We decided to choose the exponents of the diffuse functions in line with eq 14, except for a small shift. More precisely, with the exponent of the most diffuse function for a given angular momentum being $\alpha_{l,min}$, the exponents of the diffuse function is chosen to be

$$\alpha_{l,diffuse} = \alpha_{l,min}/2 + 0.05\alpha_{l,min} \tag{15}$$

The exceptions are the elements hydrogen to beryllium, for which we chose the already optimized exponents by Chong[90] These exponents are very close to fulfilling eq 15. Our choice of exponents is a compromise between two requirements: we would like the additional functions to be as diffuse as possible but we still want to be able to fit them accurately with a large auxiliary basis set. As discussed for instance in ref 29, this is already challenging. In practice, this means that $2\alpha_{l,diffuse}$ should be at least equal or preferably slightly larger (for this reason, we added the shift in eq 15) than the exponent of the most diffuse function in the auxiliary basis set.

3. RESULTS

3.1. Computational Details. All calculations have been performed with a locally modified development version of ADF2020[91] using the implementation as described in ref 29 and using the updated imaginary frequency grids as described in this work.[92]

3.1.1. GW100. We follow the protocol for GW100 as described by van Setten et al.[7] and perform nonrelativistic G_0W_0@PBE[93,94] calculations. We use the structures as available on the web page for the GW100 database and also use the updated structures for phenol and vinyl bromide.[95] For consistency with other benchmarks for GW100, we always use the QP energy obtained from the KS LUMO energy, which is usually, but not always, the energetically lowest virtual QP energy. For a detailed discussion of the effect of orbital reordering, we refer the reader to ref 26.

We performed calculations with the augmented versions of the basis sets described in this work and extrapolated them to the CBS limit as described for instance in ref 6: we calculate the QP energies ϵ_n with the aug-TZ3P and aug-QZ6P basis sets and estimate the CBS limit as

$$\epsilon_n^{CBS} = \epsilon_n^{QZ} - \frac{1}{N_{bas}^{QZ}} \frac{\epsilon_n^{QZ} - \epsilon_n^{TZ}}{\frac{1}{N_{bas}^{QZ}} - \frac{1}{N_{bas}^{TZ}}} \tag{16}$$

where ϵ_n^{QZ} (ϵ_n^{TZ}) denotes the value of the QP energy using aug-QZ6P (aug-TZ3P) and N_{bas}^{QZ} and N_{bas}^{TZ} denote the respective numbers of basis functions. Since we work with spherical harmonics, there are 5d and 7f functions. In all calculations, we set the numericalQuality key to good and used 32 imaginary time and 32 imaginary frequency points each. As explained in the appendix, this should be understood as the upper bound for the number of grid points. We report the number of grid points actually used in each calculation in the Supporting Information. Per default, we also used the good auxiliary basis set. However, for all systems in GW100 containing fourth- or fifth-row elements and for all systems smaller than four atoms, we used the normal auxiliary basis set to prevent issues with overfitting (see ref 64). For systems with a positive LUMO, we used the veryGood fit set. This fit set contains additional diffuse fit functions which are necessary to accurately represent the diffuse electron densities of these systems. These fit sets are discussed in ref 64. Also, for guanine, uracil, pentasilane, and ethoxyethane, we used the veryGood fit set since we observed inconsistencies with the good fit set. In, all calculations, we set dependency bas = 5×10^{-4} in the AMS input, as described in ref 29.

3.1.2. GW5000. We also performed G_0W_0@PBE0[96,97] calculations for a subset of 250 molecules from the GW5000 database[11] using the zeroth-order regular approximation (ZORA).[98−101] Calculations were performed with the non-augmented TZ3P and QZ6P basis sets. Equation 16 is used for CBS limit extrapolation. We used 24 points in imaginary frequency and imaginary time each and numericalQuality good. We used the normal auxiliary basis set which is sufficient for nonaugmented basis sets.[29] We set dependency bas = 5×10^{-4}.

3.2. Basis Set Errors. Before we compare the STO results to the ones from other codes in detail, we discuss the basis set errors (BSE) and basis set convergence properties of aug-TZ3P and aug-QZ6P. Using eq 16, one implicitly assumes that when going to a larger basis set, each additional basis function reduces the BSE on average by the same amount. In other words, one assumes uniform convergence of the basis set expansion, a natural property of finite elements in real or reciprocal space. It is clear that such an assumption will only be justified for rather large basis sets, and for the same reason, extrapolation is generally more reliable for larger systems. Usually, one would also like to use three or even more data points instead of using eq 16. It has, however, been pointed out[39] that including a calculation with a basis set of quality

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

lower than TZ will deteriorate the quality of the fit. Therefore, we calculate QP energies at the CBS limit from TZ and QZ results only.

To demonstrate the accuracy of an extrapolation scheme, one would ideally compare the extrapolated result to the one obtained with a very large basis, which already gives a result very close to the CBS limit. Due to the limitations of our basis sets to angular momenta ≤3, this is not possible for us. As a rule of the thumb, extrapolation with basis sets of cardinality X and $X - 1$ can provide the accuracy of a calculation using a basis set of accuracy $X + 1$.[40] Recently, Bruneval et al.[12] found BSEs of around 60 meV for the IPs of a large set of small to medium organic molecules with the cc-pV5Z basis set. This is of the same order as the typical accuracy in a photoionization experiment[12] and considerably lower than the 150 meV for IPs, which are usually found using the cc-pVQZ basis set.[11,12] For EAs, one can usually expect errors of the same order of magnitude than IPs when augmented basis sets are used.[102] It is thus reasonable to use this number as an estimate of the average error in our extrapolation.

The distributions of BSEs with respect to the CBS limit-extrapolated IPs and EAs (excluding noble gases and the hydrogen molecule in the latter) for the GW100 database are shown in Figures 2 and 3, respectively. The average BSE

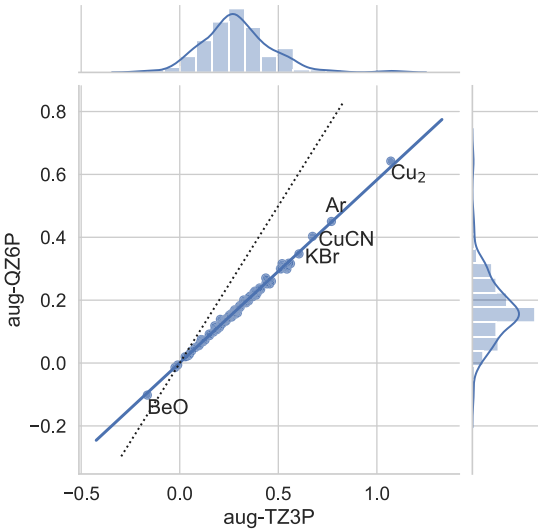


Figure 2. Basis set errors with respect to the CBS limit-extrapolated IPs in the GW100 database using the aug-TZ3P and aug-QZ6P basis sets. The univariate plots show the distributions of errors with respect to the CBS limit-extrapolated values. All values are in electron volts.

reduces from 300 (290) to 170 (170) meV for IPs (EAs), that is, there seems to be no qualitative difference in the convergence to the CBS limit for the IP and EA as one would expect for augmented basis sets. We also notice that with the only exception of the IP of BeO, the BSE at the TZ level is always lower than the one at the QZ level. In both plots, we also highlight some systems for which the convergence to the CBS limit seems to be rather slow, that is, the difference between QP energies at the TZ and QZ level is very large. For these systems, CBS limit extrapolation will be less accurate. Without exception, the problematic systems are composed of only a few atoms.

Also, a good correlation between the basis set errors for IPs and EAs is desirable since it implies a fast convergence of the

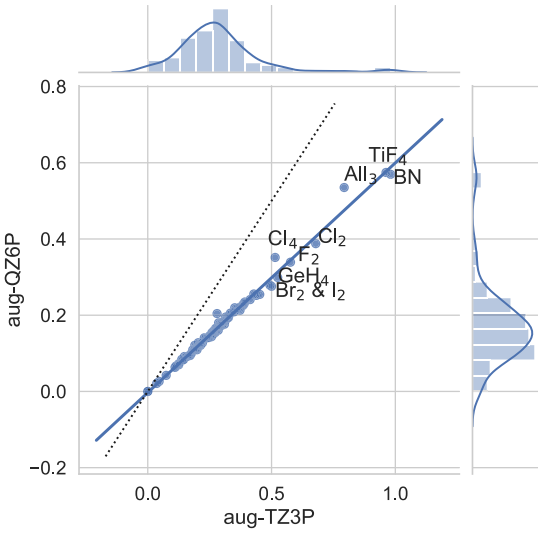


Figure 3. Basis set errors with respect to the CBS limit-extrapolated EAs in the GW100 database using the aug-TZ3P and aug-QZ6P basis sets. The univariate plots show the distributions of errors with respect to the CBS limit-extrapolated values. All values are in electron volts.

HOMO−LUMO QP gap to the CBS limit. Fast convergence of this quantity with basis sets augmented with diffuse functions has been demonstrated before.[29,102] At the first glance, the distributions of IP and EA BSEs appear rather similar, suggesting that such an error cancelation might also be found for our basis sets. To investigate this further, we plot all pairs $(\Delta(\mathrm{QZ} - \mathrm{TZ})_{\mathrm{IP}i}, \Delta(\mathrm{QZ} - \mathrm{TZ})_{\mathrm{EA}i})$ (bivariate plot) together with the corresponding error distributions (univariate plots) in Figure 4 (omitting again all noble gases and H_2). The blue solid line is a linear fit, the dotted black line is defined by the equation $\Delta(\mathrm{QZ} - \mathrm{TZ})_{\mathrm{IP}i} = \Delta(\mathrm{QZ} - \mathrm{TZ})_{\mathrm{EA}i}$, and Gaussian kernels are fitted to the univariate distributions. A few molecules with a large BSE for the IP but small BSE for the EA (Cu_2) or vice versa (BN, TiF_4, F_2, and AlI_3) aside, most

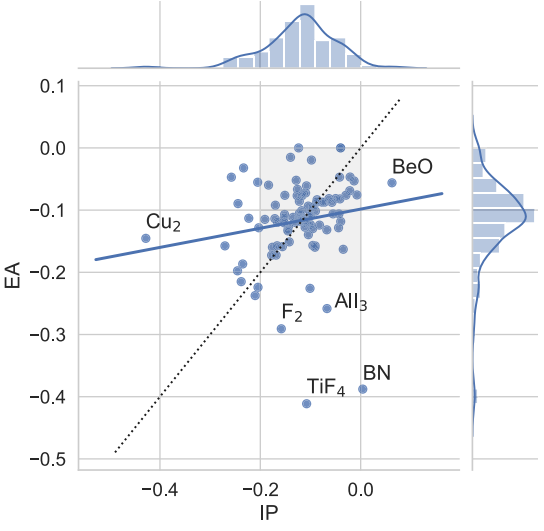


Figure 4. Correlation between differences between the aug-TZ3P and aug-QZ6P results Δ_{T-Q} for EAs (y-axis) and IPs (x-axis). The univariate plots show the corresponding distributions. The blue straight line is a linear fit, and the dotted line is defined by $\Delta_{\mathrm{IP}} = \Delta_{\mathrm{EA}}$. All values are in electron volts.

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

systems cluster around the dotted line in the gray shaded area in which IP and EA BSEs cancel each other to a large extent.

3.3. Comparison to GTO-Type Basis Sets. Additional insights into the convergence properties of the STO-type basis sets is provided when comparing them to GTO-type ones of the same cardinality. Such a comparison is made in Figure 5 for

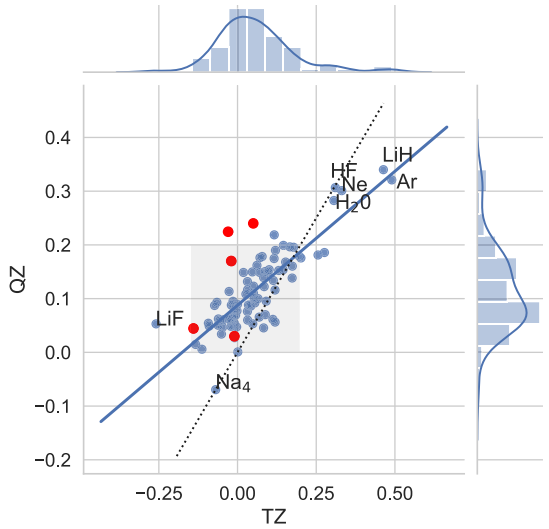


Figure 5. Distribution of deviations of the ADF IPs to the def2-GTO IPs for the GW100 database for TZ and QZ basis sets ($\Delta_{X,i}$, X = TZ,QZ). The univariate plots show the distribution of the $\Delta_{TZ,i}$ (upper histogram) and the $\Delta_{QZ,i}$ (left histogram). The bivariate plots show the pairs ($\Delta_{TZ,i}$, $\Delta_{QZ,i}$), and the blue line is a linear fit. The dotted line is defined by $\Delta_T = \Delta_Q$. Systems containing fifth-row elements are highlighted in red. All values are in electron volts.

results on the TZ and for the QZ level. The univariate plot on top of Figure 5 shows the deviation of the aug-TZ3P IPs to the ones obtained with def2-TZVP, and the univariate plot on the right side of Figure 5 shows the same for aug-QZ6P and def2-QZVP (here and in the remainder of this paper, we use the results calculated with the TURBOMOLE code whenever we refer to def2-GTO basis sets. We could have equally well used results obtained with other codes such as FHI-AIMS or MOLGW). Again, Gaussian kernels are fitted to the univariate distributions. The bivariate plot shows the individual pairs ($\Delta_{T,i}$, $\Delta_{Q,i}$), with $\Delta_{T,i}$ ($\Delta_{Q,i}$) being the differences between the IPs calculated with aug-TZ3P and def2-TZVP (aug-QZ6P and def2-QZVP) for the ith data point in the GW100 database. The dotted line is defined by $\Delta_{T,i} = \Delta_{Q,i}$. The systems for which the QP eq 2 can have multiple solutions (CI_4, KBr, NaCl, BN, O_3, BeO, MgO, Cu_2, and CuCN) are excluded from this comparison.[24] Also, Ag_2 is not shown since the deviations on the TZ and QZ level are exceptionally large.

We observe that the maximum of the Gaussian kernel function is close to 0 eV for the TZ and closer to 0.1 eV for the QZ basis sets. In other words, the aug-QZ6P IPs are consistently smaller than the def2-QZVP IPs [with a mean deviation (MD) of 120 eV], while the aug-TZ3P ones are with a MD of 60 meV, only slightly smaller than the def2-TZVP IPs on average. The missing l_{occ} + 3 function in the aug-QZ6P basis set might be a reason for this discrepancy. We will see in the next subsection that the CBS limit-extrapolated IPs calculated with our basis sets are on average lower than the def2-GTO results. This section clearly shows that this discrepancy is mostly caused by the differences at the QZ level.

The deviations of the STO-type to the respective def2-basis sets are strongly correlated. In cases in which the aug-TZ3P IPs are considerably smaller than their def2-TZVP counterparts, also the aug-QZ6P IPs will be much smaller than the def2-QZVP ones. Good examples are the five molecules represented by the points in the upper right corner of the bivariate plot. Since it is known that the GTO-type basis sets allow for a reliable CBS limit extrapolation, this fact is highly important since it guarantees that the same CBS limit extrapolation can be performed using our STO-type basis sets.

We also shortly comment on the systems containing the fifth-row elements, highlighted in red. With the exception of one of them (Rb_2), the agreement on the TZ level is rather good, while the agreement on the QZ level is significantly worse. This is again due to the inconsistent polarization for the heavier elements mentioned above: Reaching the CBS limit for heavier elements becomes difficult without Slater functions with angular momentum larger than l = 3.

To summarize the key points of this section, our STO-type basis sets seem to behave qualitatively similar to the GTO-type basis sets, although the improvement when going from TZ to QZ is smaller for the STO- than for the GTO-type basis sets. Together with the good correlation of deviations on the TZ and QZ level, this indicates that our basis sets allow indeed for a meaningful CBS limit extrapolation. However, the CBS limit-extrapolated IPs from the STO-type basis sets will on average be lower than their counterparts calculated using GTOs.

3.4. Comparison to Other Codes. *3.4.1. KS Eigenvalues.* Before we dive into our comparison of the *GW* QP energies, we shortly compare our KS eigenvalues for the systems in GW100 to the ones from other codes. Here and in the following, we do not include phenol and vinyl bromide in the statistical analysis since different structures have been used in the past for both systems.[22] Our results (see Supporting Information for the raw data for all 100 systems) only confirm what is already well known; for KS eigenvalues, the agreement between different codes is generally excellent. For example, the CBS limit-extrapolated KS HOMO energies from the WEST code agree with the ones obtained from def-GTO calculations within 30 meV on average, with a maximum deviation of 176 meV.[24] These figures reduce to 24 and 92 meV when the plane-wave results are compared to def2-QZVP.[24] With only 19 meV, the agreement from VASP to def2-QZVP is even better.[22] One should keep in mind that the extrapolation schemes for correlated-electron methods and localized basis functions are not necessarily useful to extrapolate KS eigenvalues, as has already been pointed out in refs 22 and 24. Such a comparison should rather be based on nonextrapolated results.[103,104]

Our KS HOMO eigenvalues calculated on the aug-QZ6P level of theory show a mean absolute deviation (MAD) of 26 meV to the ones on the def2-QZVP level and of 22 meV to the CBS limit extrapolated values calculated with the WEST code. Our LUMO eigenvalues only differ to the ones from WEST by 35 meV on average. Major deviations of more than 150 meV are only found for helium (340 meV), H_2 (280 meV), and Ag_2 (340 meV). With deviations of more than 420 meV to WEST and 490 meV to def-QZVP, the latter system is also the only outlier for IPs. However, when the ZORA is made, the deviations to WEST reduce to 15 meV for the HOMO and 17 meV for the LUMO, respectively. Also, the deviation to the def2-QZVP IP reduces to an acceptable value of 50 meV,

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

Table 1. G_0W_0@PBE IP for the GW100 Database (Third Column). Columns Four to Seven Denote Deviations of the ADF IPs to the Ones from Reference X, $\Delta_X = \mathrm{IPX} - \mathrm{IP}_{\mathrm{ADF}}$[a]

	name	ADF	$\Delta_{\mathrm{def2\text{-}GTO}}$	Δ_{VASP}	Δ_{WEST}	Δ_{nanoGW}
1	helium	23.31	0.18	0.07	0.11	−0.11
2	neon	20.06	0.27	0.11	0.27	0.39
3	argon	15.26	0.02	0.06	0.11	0.19
4	krypton	13.71	0.18	0.22	0.05	−0.06
5	xenon	11.88	0.45	0.26	1.24	0.23
6	hydrogen	15.88	−0.03	−0.03	−0.04	−0.13
7	lithium dimer	5.08	−0.02	0.01	−0.04	−0.04
8	sodium dimer	4.89	0.02	0.04	0.09	0.05
9	sodium tetramer	4.25	−0.04	−0.08	−0.01	−0.03
10	sodium hexamer	4.30	0.04	0.04	0.07	0.05
11	potassium dimer	3.99	0.09	0.13	0.15	0.07
12	rubidium dimer	3.78	0.10	0.24	0.23	0.08
13	nitrogen	14.79	0.26	0.14	0.15	0.04
14	phosphorus dimer	10.26	0.12	0.09	0.17	0.06
15	arsenic dimer	9.58	0.08	0.01	−0.03	−0.09
16	fluorine	14.99	0.12	−0.06	0.01	0.03
17	chlorine	11.18	0.13	0.14	0.23	0.13
18	bromine	10.46	0.10	0.11	−0.02	−0.17
19	iodine	9.04	0.54	0.48	1.37	0.23
20	methane	13.89	0.11	0.13	0.10	0.26
21	ethane	12.35	0.11	0.15	0.09	−0.19
22	propane	11.85	0.04	0.05	−0.01	−0.32
23	butane	11.52	0.07	0.09	−0.11	−0.27
24	ethylene	10.28	0.12	0.14	0.11	−0.01
25	acetylene	11.14	−0.05	−0.07	−0.05	−0.21
26	tetracarbon	10.74	0.17	0.15	0.16	0.06
27	cyclopropane	10.63	0.02	0.09	0.04	−0.15
28	benzene	9.07	0.03	0.04	0.01	−0.11
29	cyclooctatetraene	8.16	0.02	0.03	0.00	−0.14
30	cyclopentadiene	8.43	0.02	0.04	0.01	−0.13
31	vinyl fluoride	10.24	0.08	0.04	0.05	−0.06
32	vinyl chloride	9.82	0.08	0.10	0.12	−0.01
33	vinyl bromide	9.03	0.11	0.72	0.61	0.49
34	vinyl iodide	8.93	0.26	0.34	0.88	0.09
35	tetrafluoromethane	15.43	0.17	−0.02	0.08	−0.04
36	tetrachloromethane	11.24	−0.03	−0.04	0.05	−0.00
37	tetrabromomethane	10.05	0.17	0.20	0.06	−0.01
38	tetraiodomethane	8.74	0.31	0.37		0.07
39	silane	12.37	0.02	0.03	0.05	−0.12
40	germane	12.09	0.03	0.04	0.23	−0.18
41	disilane	10.46	−0.05	−0.02	0.06	−0.14
42	pentasilane	9.07	−0.02	0.06	0.12	−0.08
43	lithium hydride	6.52	0.07	−0.06	0.11	0.15
44	potassium hydride	4.88	0.09	0.09	0.09	−0.20
45	borane	12.93	0.03	0.02	0.02	−0.15
46	diborane (6)	11.92	0.01	0.02	−0.00	−0.25
47	ammonia	10.27	0.12	0.05	−0.09	−0.20
48	hydrazoic acid	10.44	0.11	0.06	0.04	−0.26
49	phosphine	10.27	0.08	0.08	0.16	0.04
50	arsine	10.31	−0.11	−0.05	0.02	−0.11
51	hydrogen sulfide	10.12	0.01	−0.01	0.11	0.07
52	hydrogen fluoride	15.14	0.23	0.23	0.09	0.12
53	hydrogen chloride	12.36	0.00	0.09	0.12	0.11
54	lithium fluoride	10.03	0.24	0.04	0.08	0.03
55	magnesium fluoride	12.44	0.07	−0.03	0.02	−0.14
56	titanium tetrafluoride	13.93	0.15	0.08		0.00
57	aluminum fluoride	14.33	0.15	0.00	0.07	−0.13
58	boron monofluoride	10.59	0.14	−0.13	−0.03	−0.17
59	sulfur tetrafluoride	12.34	0.05	−0.14	−0.02	−0.18
60	potassium bromide	7.85	−0.24	−0.05		−0.74

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

Table 1. continued

	name	ADF	$\Delta_{\text{def2-GTO}}$	Δ_{VASP}	Δ_{WEST}	Δ_{nanoGW}
61	gallium monochloride	9.81	−0.07	0.08	0.36	−0.06
62	sodium chloride	8.32	0.03	0.15		−0.14
63	magnesium chloride	11.07	0.12	0.12	0.18	0.04
64	aluminum iodide	9.38	0.20	0.20	0.93	−0.07
65	boron nitride	10.94	0.15			0.25
66	hydrogen cyanide	13.10	0.22	0.19	0.12	0.12
67	phosphorus mononitride	11.10	0.20	0.14	0.16	0.11
68	hydrazine	9.40	−0.03	−0.07	−0.13	−0.30
69	formaldehyde	10.46	0.00	−0.04	−0.05	−0.25
70	methanol	10.54	0.13	0.07	0.07	−0.23
71	ethanol	10.17	0.10	0.04	0.04	−0.20
72	acetaldehyde	9.61	0.05	0.02	0.00	−0.22
73	ethoxy ethane	9.41	0.02	0.02	−0.02	−0.30
74	formic acid	10.84	0.03	−0.03	−0.03	−0.13
75	hydrogen peroxide	10.96	0.14	−0.00	0.04	0.02
76	water	11.94	0.11	−0.10	−0.07	−0.12
77	carbon dioxide	13.37	0.09	−0.01	0.00	−0.25
78	carbon disulfide	9.80	0.15	0.16	0.25	0.09
79	carbon oxide sulfide	10.90	0.21	0.16	0.26	0.08
80	carbon oxide selenide	10.40	0.03	0.02	−0.03	−0.22
81	carbon monoxide	13.66	0.05	−0.04	−0.00	−0.18
82	ozone	11.74	−0.23			0.33
83	sulfur dioxide	11.86	0.20	0.05	0.10	0.05
84	beryllium monoxide	8.99	−0.48			0.76
85	magnesium monoxide	6.82	−0.07			0.24
86	toluene	8.72	0.00	0.03	−0.01	−0.14
87	ethylbenzene	8.65	0.01	0.04	0.01	−0.16
88	hexafluorobenzene	9.70	0.04	−0.07	−0.05	−0.13
89	phenol	8.47	0.04	−0.09		−0.24
90	aniline	7.72	0.07	0.06	0.01	−0.14
91	pyridine	9.18	0.00	−0.02	−0.05	−0.19
92	guanine	7.77	0.10	0.08	0.05	−0.12
93	adenine	8.14	0.01	−0.02	−0.05	−0.20
94	cytosine	8.44	0.01	−0.04	−0.04	−0.19
95	thymine	8.86	0.01	−0.03	−0.04	−0.22
96	uracil	9.26	0.12	0.10	−0.07	−0.02
97	urea	9.25	0.21	0.10	0.14	−0.03
98	silver dimer	7.06	0.91	0.77	0.98	0.81
99	copper dimer	8.20	−1.57	−1.01		−0.33
100	copper cyanide	10.25	−0.73			0.40

[a]All values are in electron volts.

which is rather strange, given that the latter has been calculated without relativistic corrections.

3.4.2. Ionization Potentials. We now turn our attention to the IPs calculated within the GWA. In addition to phenol and vinyl bromide, we again exclude the systems for which the QP eq 2 can have multiple solutions (CI_4, KBr, NaCl, BN, O_3, BeO, MgO, Cu_2, and CuCN)[24] from the following statistical comparison, but also TiF_4 and OCS for which no IPs from the WEST code are available. Due to large discrepancies between WEST and the VASP, we also exclude the systems containing iodine, gallium, and xenon. Finally, we also exclude all remaining systems containing the fifth-row elements from our analysis since for these systems (especially the ones containing iodine and Ag_2), the different treatments of relativistic effects have been shown to significantly affect QP energies.[22] This leaves us with a set of 81 molecules, whose IPs we include in the statistical analysis in this section.

Table 1 shows the results for all 100 IPs in the GW100 database obtained with our code next to the deviations to def2-GTO, VASP, WEST, and nanoGW, if available. To facilitate a discussion of the data, Figure 6 shows median absolute deviations (MADs) and maximum absolute errors between all codes, while Figure 7 visualizes the distribution of the deviations of the ADF IPs to the ones from other codes. The IPs from ADF, def2-GTO, VASP, and WEST are all in good agreement with each other, with MADs between 56 and 86 meV, while the deviations to nanoGW are around twice as large. Figure 7 also shows that the deviations of the ADF IPs to the ones from other codes (again, with the exception of nanoGW) show a small spread and no outliers can be found. We note again that we assume the CBS limit extrapolation error to be of the order of at least 60 meV on average and all the values reported and compared here should only be interpreted with these error bars. For the plane-wave codes, the CBS limit extrapolation error is likely to be smaller than for the

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

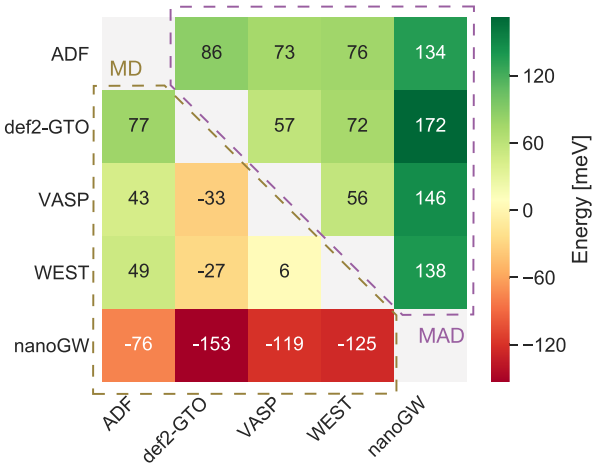


Figure 6. MADs (upper triangle) and MDs (lower triangle) of the CBS limit-extrapolated IPs in the GW100 database (18 molecules have been excluded from the comparison; see explanations above) computed with different codes. All values are in millielectron volts.

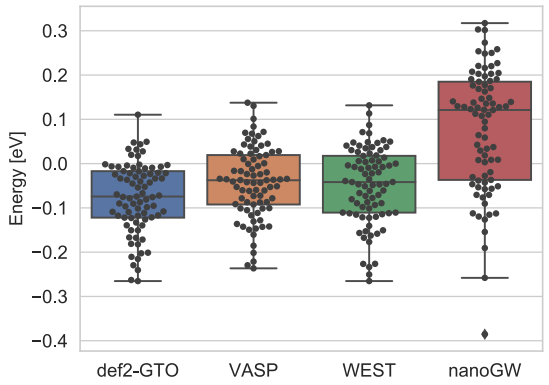


Figure 7. Deviation of IPs calculated with ADF to different codes. Black dots denote the individual data points. The horizonal line in each box denotes the median deviation, the box contains all data points between the first quartile ($Q1$) and third quartile ($Q2$) and the whiskers are at $Q1 \pm \frac{3}{2}|Q1 - Q3|$ (in case of a normal distribution, the whiskers include 99.3% of all data points). All values are in electron volts.

localized basis sets, but there are additional sources of error, most notably pseudopotentials and box-size effects. In light of these uncertainties, the agreement between all four codes can be considered as excellent.

Looking at the MDs to the other codes in the lower triangle of Figure 6 as well as at the box plots in Figure 7, we see that ADF IPs are generally smaller than the ones from def2-GTO, VASP, and WEST. The reasons for the discrepancy between ADF and def2-GTO have already been discussed above. We also see that the nanoGW IPs are on average much smaller than the ones from all others codes. This is in line with the fact that the nanoGW results were apparently obtained without basis set extrapolation,[26] although the numerical parameters determining the convergence to the CBS limit (grid spacing, chosen cutoff for virtual states and radius of the sphere around a given finite systems) were tested for convergence separately. Still, our data analysis suggests that the nanoGW IPs are not as well converged as the ones from the other codes.

Systems containing fluorine and nitrogen generally show rather pronounced disagreements between different codes (e.g., consider the following deviations from ADF to def2-GTO: N_2: 260 meV, HCN: 220 meV, TiF_4: 150 meV, AlF_3: 150 meV, LiF: 240 meV, HF: 230 meV). For fluorine, this has already been observed by Maggio and Kresse in ref 22, who pointed at the default pseudopotentials in VASP as a potential source for these discrepancies. Another possible explanation might be found in a recent study by Bruneval et al.,[12] which suggests that molecules predominantly composed of carbon and hydrogen converge to the CBS limit rather quickly, whereas the convergence is considerably slower for systems to a large part composed of fluorine and nitrogen.

Finally, we shortly comment on some systems which we have been excluded from the statistical comparisons in Figures 6 and 7. We find large differences for system containing fifth-row elements, for example, Xe, Ag_2, I_2, or CI_4. Here, the ADF IPs are considerably lower than the ones from other codes, indicating that the ADF results are not properly converged to the CBS limit, which is due to the missing basis functions with angular momentum higher than $l = 3$.

With 1.57 eV, the largest deviation reported in Table 1 by far can be found between ADF (8.20 eV) and TURBOMOLE (6.63 eV) for Cu_2. FHI-AIMS gives an IP of 7.78 eV for this system,[7] which is in considerably better agreement with ADF. FHI-AIMS relies on an analytical continuation from the imaginary to the real frequency axis with 16 sampling points (AIMS-P16), not much different from the procedure in ADF. We can conclude that for this particular system, the large deviations of ADF and FHI-AIMS to TURBOMOLE are caused by non-converged frequency grids. The ADF IP reported in Table 1 has been obtained from aug-TZ3P and aug-QZ6P calculation with 24 and 27 imaginary frequency points, respectively. When 24 imaginary frequency points are used for the aug-QZ6P calculation as well, the IP of Cu_2 reduces to 8.05 eV, which is already in reasonable agreement with FHI-AIMS. Similar conclusions can be drawn for CuCN. Another interesting case is BeO, with a deviation of 0.48 eV from def-GTO. Again, the large deviation is due to non-converged frequency grids. The ADF IP is with 8.99 eV, very close to the 9.07 eV obtained by AIMS-P16. With 128 imaginary frequency points (AIMS-P128), FHI-AIMS gives an IP of 9.63 eV which is then in perfect agreement with TURBOMOLE. Furthermore, in ref 7, three solutions are reported for BeO, while ADF only recovers one of them. These three examples show that the frequency treatment currently implemented in ADF cannot properly describe the IPs of systems, for which the single QP picture breaks down in the valence region.

3.4.3. Electron Affinities. We now turn our attention to the EAs. As for the IPs, Table 2 shows the EAs calculated with ADF and the differences from the other four codes excluding all noble gases and H_2. However, it is known, that the def2-GTO basis sets sometimes severely overestimate positive LUMO energies, which then deviate from results from plane-wave codes by more than 1 eV. Furthermore, since EAs converge slower to the CBS limit than IPs when non-augmented basis sets are used,[12,29] also basis set extrapolation errors are larger for the remaining systems. On the other hand, plane-wave calculations require very large box sizes for these systems, which makes it harder to converge the EAs with respect to this parameter, and for this reason, results from VASP are often not available.[22]

Thus, for the full database, only comparison to WEST and nanoGW is possible. Excluding again all compounds

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

Table 2. G_0W_0@PBE EA for the GW100 Database (Third Column). Columns Four to Seven Denote Deviations of the ADF EAs to the Ones from Reference *X*, $\Delta_X = EAX - EA_{ADF}$[a]

	name	ADF	$\Delta_{def2\text{-}GTO}$	Δ_{VASP}	Δ_{WEST}	Δ_{nanoGW}
7	lithium dimer	0.52	0.23	0.09	0.12	0.14
8	sodium dimer	0.64	0.02	−0.04	−0.03	−0.00
9	sodium tetramer	0.92	0.23	0.15	0.16	0.20
10	sodium hexamer	0.95	0.18	0.12	0.09	0.11
11	potassium dimer	0.59	0.16	0.15	0.16	0.16
12	rubidium dimer	0.67		0.07	0.06	0.07
13	nitrogen	−2.40	0.28		0.25	0.17
14	phosphorus dimer	0.64	0.44	0.35	0.45	0.38
15	arsenic dimer	1.08	0.44	−0.01	0.01	−0.04
16	fluorine	0.54	0.69		0.52	0.58
17	chlorine	0.83	0.57	0.42	0.55	0.52
18	bromine	1.40	0.56	0.59	0.48	0.29
19	iodine	1.56		0.65	1.65	
20	methane	−0.78	−1.25	0.15	0.02	−0.10
21	ethane	−0.77	−1.16		−0.01	−0.11
22	propane	−0.72	−1.15		−0.03	−0.11
23	butane	−0.70	−1.13		−0.04	−0.13
24	ethylene	−1.91	0.09		0.11	−0.04
25	acetylene	−2.48	−0.08		−0.02	−0.20
26	tetracarbon	2.62	0.53	0.47	0.48	0.52
27	cyclopropane	−0.73	−1.23		−0.02	−0.14
28	benzene	−0.96	0.07		0.03	−0.10
29	cyclooctatetraene	0.03	0.09	0.02	0.04	−0.10
30	cyclopentadiene	−0.91	0.06		0.01	−0.15
31	vinyl fluoride	−1.92	0.04		0.03	−0.09
32	vinyl chloride	−1.31	0.14	0.12	0.10	0.00
33	vinyl bromide	−1.23	0.12		0.17	0.09
34	vinyl iodide	−0.77		0.40	0.55	0.44
35	tetrafluoromethane	−0.88	−3.00		0.05	−0.10
36	tetrachloromethane	0.04	0.50	0.28	0.37	0.41
37	tetrabromomethane	0.99	0.57	0.48	0.46	0.41
38	tetraiodomethane	2.16		0.28	0.88	0.32
39	silane	−0.72	−1.54		−0.04	−0.11
40	germane	−0.47	−1.38		−0.14	−0.36
41	disilane	−0.75	−0.76		−0.02	−0.83
42	pentasilane	−0.08	0.08	0.05	0.15	−0.02
43	lithium hydride	0.05	0.11	0.02	0.02	0.13
44	potassium hydride	0.17	0.15	0.08	0.08	0.35
45	borane	−0.26	0.23	0.23	0.25	0.21
46	diborane (6)	−0.87	0.13		0.15	−0.00
47	ammonia	−0.76	−1.24		−0.05	−0.09
48	hydrazoic acid	−1.40	0.30		0.25	0.23
49	phosphine	−0.67	−1.59		−0.03	−0.06
50	arsine	−0.58	−1.36		−0.08	−0.16
51	hydrogen sulfide	−0.73	−1.52		−0.05	−0.11
52	hydrogen fluoride	−1.06	−0.98		−0.05	−0.12
53	hydrogen chloride	−1.19	−0.34		0.10	−0.03
54	lithium fluoride	−0.04	0.05	−0.13	−0.03	0.11
55	magnesium fluoride	0.26	0.05	0.03	0.07	0.18
56	titanium tetrafluoride	0.09	0.97		0.83	1.14
57	aluminum fluoride	0.06	0.17	−0.14	0.10	0.01
58	boron monofluoride	−1.21	0.16		0.28	0.24
59	sulfur tetrafluoride	−0.29	0.39	0.22	0.37	0.35
60	potassium bromide	0.34	0.08	−0.02	0.06	0.59
61	gallium monochloride	0.02	0.37	0.17	0.42	0.20
62	sodium chloride	0.42	−0.00	0.04	0.05	0.61
63	magnesium chloride	0.68	−0.00	−0.07	0.02	0.05
64	aluminum iodide	1.18		−0.16	0.48	0.02
65	boron nitride	4.05	−0.10		0.03	0.05
66	hydrogen cyanide	−2.31	0.09		0.06	−0.06

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

Table 2. continued

	name	ADF	$\Delta_{\text{def2-GTO}}$	Δ_{VASP}	Δ_{WEST}	Δ_{nanoGW}
67	phosphorus mononitride	0.12	0.47		0.40	0.35
68	hydrazine	−0.70	−0.98		−0.02	−0.08
69	formaldehyde	−1.06	0.35		0.30	0.15
70	methanol	−0.81	−1.00		−0.10	−0.19
71	ethanol	−0.73	−0.94		−0.11	−0.19
72	acetaldehyde	−1.16	0.33	0.29	0.29	0.25
73	ethoxy ethane	−0.62	−1.08		−0.09	−0.18
74	formic acid	−1.82	0.23	0.18	0.18	0.04
75	hydrogen peroxide	−2.06	0.11		0.26	0.28
76	water	−0.88	−1.13	−0.16	−0.03	−0.08
77	carbon dioxide	−1.03	0.10		0.06	−0.03
78	carbon disulfide	0.10	0.45	0.32	0.40	0.32
79	carbon oxide sulfide	−1.22	0.39		0.28	0.20
80	carbon oxide selenide	−0.93	0.41		0.29	0.22
81	carbon monoxide	−0.84	0.47		0.40	0.38
82	ozone	2.03	0.66	0.47	0.53	0.52
83	sulfur dioxide	0.86	0.63	0.39	0.51	0.36
84	beryllium monoxide	1.99	0.73	0.74	0.52	0.54
85	magnesium monoxide	1.74	0.39	0.31	0.21	0.52
86	toluene	−0.91	0.08		0.04	−0.09
87	ethylbenzene	−0.87	0.00		−0.03	−0.18
88	hexafluorobenzene	−0.03	−0.33		0.00	−0.08
89	phenol	−0.78	0.04		−0.07	−0.18
90	aniline	−0.91	−0.03		−0.07	−0.19
91	pyridine	−0.44	0.14		0.08	−0.05
92	guanine	−0.48	0.02		−0.03	−0.15
93	adenine	−0.28	0.07		−0.01	−0.13
94	cytosine	−0.18	0.17	0.06	0.09	−0.03
95	thymine	0.02	0.16	0.04	0.08	−0.06
96	uracil	0.05	0.20	0.06	0.10	−0.02
97	urea	−0.49	−0.68		−0.04	−0.12
98	silver dimer	0.91		0.44	0.58	0.41
99	copper dimer	1.00	0.23	0.24	0.41	0.25
100	copper cyanide	1.47	0.38	0.44	0.51	0.34
MAD			0.48	0.21	0.16	0.21

[a]All values are in electron volts.

containing iodine, copper, gallium, and xenon as well as the remaining systems containing fifth-row elements, we find a MAD of 160 meV to the former and of 210 meV to the latter code. These MADs are around twice as large as for IPs but in light of the difficulties mentioned above, certainly not surprising and in line with the previous benchmark studies on GW100.[7,22,24,26] Figure 8 shows that the ADF EAs, as for the IPs, are on average smaller than the ones from WEST, while no trend in that direction can be observed when compared to nanoGW.

One clearly identifies a group of molecules with significant deviations from ADF to def2-GTO of up to several electron volts. These are the already mentioned systems with positive LUMOs. Contrariwise, the raw data in Table 2 show that the agreement with WEST is especially good for systems with positive LUMOs. As an example, consider the series of linear alkane chains, C_nH_{2n+1} for $n = 1, ..., 4$. With deviations from 10 to 40 meV, the agreement with WEST is excellent, while def2-GTO overestimates the EAs of these systems by more than 1 eV. In this context, it is interesting to investigate the effect of the diffuse functions. This is shown in Table 3 for some systems with LUMOs well above the vacuum level. Only comparing the basis set extrapolated values, the effect of the

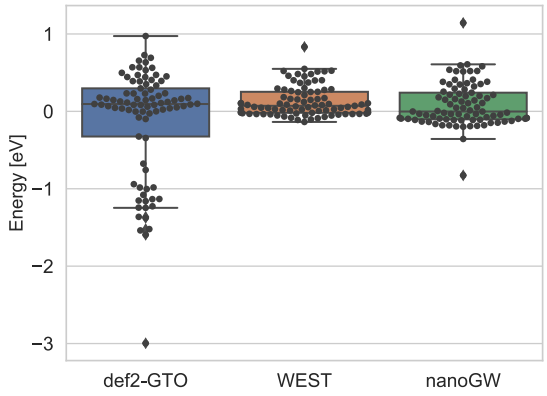


Figure 8. Deviation of EAs calculated with ADF from different codes. Black dots denote the individual data points. The horizonal line in each box denotes the median deviation; the box contains all data points between the first quartile ($Q1$) and third quartile ($Q2$), and the whiskers are at $Q1 \pm \frac{3}{2}|Q1 - Q3|$ (in case of a normal distribution, the whiskers include 99.3% of all data points). All values are in electron volts. The deviations to VASP have been excluded due to the lack of reference values for too many systems in GW100.

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

Table 3. Comparison of EAs for Selected Molecules from the GW100 Database Calculated with the ADF with and without Diffuse Functions[a]

		aug.				non.aug.						
index		T	Q	Ex.	Δ_{TQ}	T	Q	Ex.	Δ_{TQ}	Δ_{TT}	Δ_{QQ}	Δ_{EE}
13	N_2	−2.65	−2.54	−2.40	−0.11	−3.00	−2.58	−2.15	−0.42	0.36	0.04	−0.25
20	CH_4	−0.97	−0.89	−0.78	−0.08	−2.20	−1.59	−0.95	−0.61	1.23	0.70	0.17
21	C_2H_6	−0.96	−0.88	−0.77	−0.08	−2.13	−1.52	−0.87	−0.62	1.17	0.64	0.10
22	C_3H_8	−0.92	−0.83	−0.72	−0.08	−2.06	−1.44	−0.78	−0.62	1.14	0.61	0.06
23	C_4H_{10}	−0.89	−0.81	−0.70	−0.08	−2.04	−1.41	−0.74	−0.63	1.15	0.60	0.04
24	C_2H_4	−2.12	−2.03	−1.91	−0.09	−2.40	−2.12	−1.81	−0.28	0.28	0.09	−0.10
25	C_2H_2	−2.76	−2.65	−2.48	−0.11	−3.24	−2.87	−2.44	−0.37	0.48	0.22	−0.04
27	C_3H_3	−0.98	−0.88	−0.73	−0.10	−2.29	−1.61	−0.87	−0.68	1.31	0.73	0.14
31	C_2H_3F	−2.21	−2.09	−1.92	−0.12	−2.50	−2.22	−1.91	−0.28	0.29	0.13	−0.01
39	SiH_4	−0.92	−0.83	−0.72	−0.09	−1.75	−1.42	−1.08	−0.33	0.83	0.59	0.36
47	NH_3	−0.93	−0.85	−0.76	−0.08	−1.85	−1.35	−0.85	−0.51	0.93	0.50	0.09
66	HCN	−2.59	−2.48	−2.31	−0.12	−2.96	−2.61	−2.23	−0.35	0.37	0.14	−0.08
70	CH_4O	−1.05	−0.95	−0.81	−0.10	−2.00	−1.49	−0.93	−0.52	0.95	0.54	0.12
71	C_2H_6O	−0.97	−0.87	−0.73	−0.10	−1.90	−1.39	−0.84	−0.51	0.93	0.52	0.11
76	H_2O	−1.02	−0.96	−0.88	−0.06	−1.75	−1.34	−0.89	−0.41	0.73	0.38	0.01
	MD				0.09				0.48	0.81	0.43	0.05

[a]All values are in electron volts.

diffuse functions seems to be rather small; for N_2, the EA calculated from the basis set without the diffuse functions is with 250 meV even higher than the augmented basis sets. Overall, the average difference is only 50 meV, which is well within the expected error range from the CBS limit extrapolation. However, comparing the results from the finite basis sets, the differences are exorbitant. Especially, on the TZ level, the addition of diffuse functions results in a lowering of the EAs by nearly 1 eV on average. For the nonaugmented basis sets, the average difference between TZ and QZ basis set is 480 meV, resulting in differences of sometimes more than 1 eV between the EAs on the TZ and on the extrapolated level. In light of these differences, the good agreement between the CBS limit-extrapolated EAs is remarkable. Despite this good agreement, the augmented basis sets should be the preferred choice to calculate EAs of systems with unbound LUMOs since the extrapolation procedure is generally less reliable when the differences between the results for the finite basis sets are larger.

Finally, we compare our EAs for systems with a bound LUMO to WEST, nanoGW, and def-GTO results. The MADs and MDs in Figure 9 show that def2-GTO, nanoGW, and WEST are generally in good agreement for these systems, while the MADs to ADF are large. As the MDs show, ADF significantly overestimates these EAs compared to the other codes, which indicates that the results are not entirely converged to the CBS limit. This interpretation is also in line with the raw data in Table 2, showing that the deviations are generally the largest for di- and triatomic systems as well as molecules containing fluorine and chlorine, while the agreement for the medium organic molecules such as the nucleobases is satisfactory. Adding additional diffuse or tight functions would possible not result in an improved description of the EAs of the former systems. Instead, reaching the CBS limit is most likely only possible using basis functions with higher angular momenta than $l = 3$, which are not available to us. However, we can try to simulate the effect of these functions by adding additional off-center Slater functions to the basis set, which can be achieved conveniently by adding ghost atoms. This approach is reminiscent of bond-centered

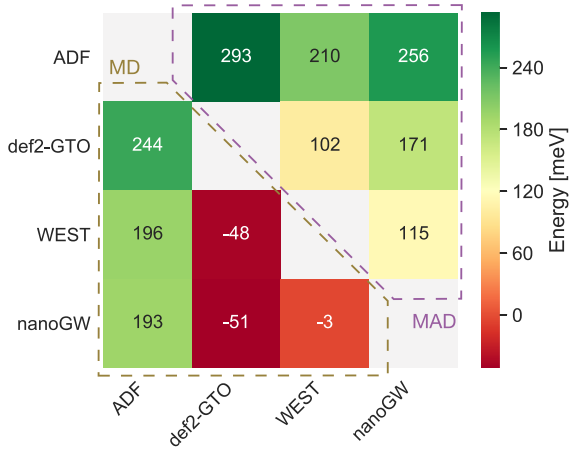


Figure 9. MADs (upper triangle) and MD (lower triangle) of the CBS limit-extrapolated EAs of the subset of systems with bound LUMOs and medium-sized organic molecules (in total 48 data points) in the GW100 data set computed with different codes. All values are in electron volts.

basis functions, which have, for example, been used by Dunlap and co-workers.[105]

As examples, we consider the dihalogens F_2, Cl_2, Br_2, and SO_2, as well as TiF_4, and we augment these structures with Ne ghost atoms, for which we use the same basis set than for the real atoms. For each atom A in a system, we place two ghost atoms G, G' on a straight line defined by the position of A and every neighboring atom B so that the distance between A and G (G') equals one-third (minus one-third) of the distance between A and B. The results of this augmentation are shown in Table 4. We clearly see that it reduces the BSEs considerably, as compared to the other codes. Of course, such an augmentation should not be seen as a practical solution, but it shows that agreement between ADF and the other codes can in principle be also reached for these systems.

3.5. IP and EA for a Subset of GW5000. For the GW100 database, discrepancies to other codes are more pronounced for smaller systems, while the agreement for medium systems is

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

Table 4. Effect of the Addition of Additional Off-Center Slater Functions via Ne Ghost Atoms on the EAs of Selected Systems from GW100[a]

system	T	Q	ex.	ex. (no gh.)	def2-GTO	WEST
F_2	0.20	0.51	0.90	0.54	1.23	1.06
Cl_2	0.74	0.95	1.21	0.83	1.40	1.38
Br_2	1.27	1.45	1.69	1.40	1.96	1.88
TiF_4	−0.33	−0.10	0.19	0.09	1.06	0.92
SO_2	0.84	1.07	1.38	0.86	1.49	1.37

[a]All values are in electron volts.

significantly better. To confirm this observation for a larger number of systems, we also calculated IPs and EAs of a subset of 250 medium to large organic molecules from the GW5000 database,[11] for which CBS limit-extrapolated references values calculated with FHI-AIMS using the def2-GTO basis sets are available. We have already considered this subset in ref 29. We used here the TZ3P and QZ6P basis sets without diffuse functions since the LUMOs of the considered systems are negative. As expected, we find better agreement with FHI-AIMS than for GW100: for IPs, ADF deviated to FHI-AIMS by 62 meV on average as opposed to the 86 meV for GW100. For EAs, the MAD is with 93 meV, slightly worse, but the agreement is much better than for GW100.

4. CONCLUSIONS

The GW100 benchmark by van Setten et al.[7] has become the most important resource for comparing different implementations of the *GW* method for finite systems. IPs and EAs calculated with implementations using def2 GTO-type basis sets,[7] PWs,[22,24] and real-space finite elements[26] have been compared in the recent past, and good agreement between CBS limit extrapolated results has been found for IPs and EAs of molecules with bound LUMOs. In this work, we extended the list of available results and calculated the IPs and EAs in the GW100 data set using STOs. For this purpose, we have developed new Slater-type basis sets which allow us to extrapolate our results to the CBS limit. These basis sets are available online.[106]

Our study confirms once more that it is possible, although difficult, to reach consensus between different implementations. All implementations compared in this work do not only implement the GWA with different basis sets but also differ in other technical parameters such as frequency treatment, description of core electrons, the algorithm used to solve the QP equations, and the numerical treatment of four-point correlation functions. In light of these differences, the observed agreement between 55 and 85 meV on average for IPs between STO, def2-GTO, and plane-wave results is excellent. Reaching the CBS limit is more difficult for EAs than for IPs. Still, EAs calculated with ADF are in excellent agreement with the plane-wave results from the WEST code for systems with positive LUMOs, with an overall MAD between both codes of 160 meV. These deviations mostly stem from large BSEs for the EAs of small molecules with bound LUMOs with the ADF code. For larger, organic molecules, agreement to def2-GTO results is with a MAD of 93 meV, significantly better.

We observe that ADF tends to give lower EAs and IPs than the def2-GTO and plane-wave results. This is an indicator that the results are not entirely converged with respect to the CBS limit. Our STO-type basis sets are restricted to angular momenta smaller or equal to $l = 3$, which sets a limit to the accuracy currently attainable with STOs. It would be desirable to obtain results for the GW100 database using even larger STO-type basis sets containing functions also with higher angular momenta. This would enable a more accurate estimate of the CBS limit and might hopefully lead to an even better agreement with other implementations. On the other hand, the best IPs for GW100 in our earlier work showed a MAD of more than 300 meV to def-GTO and nanoGW.[29] Thus, the improvements which are due to our improved basis sets and imaginary time and frequency grids are already immense.

The close agreement between different codes is highly important in practice since it allows researchers to interpret the results of their *GW* calculations without worrying that they might be skewed by technical aspects. The slow convergence of *GW* calculations to the CBS limit remains an obstacle in practical applications. Especially for *GW* calculations for finite systems with hundreds of atoms, which have become the focus of much research in the last years,[27−30] calculations at (or even close to) the CBS limit for individual QP energies are currently often out of reach and will most likely not become routine anytime soon. To overcome this issue, finite basis-set corrections[12,107−109] hold much promise and it is to hope that these techniques will be further developed and become more widespread available in the near future. For differences of QP energies, the situation is already much better. When augmented basis sets are used, BSEs usually cancel out to a large extent and already at the DZ level, fundamental gaps are often sufficiently converged to the CBS limit.[2,29,102,110−113]

A

Implementation of Imaginary Time and Frequency Grids

In this appendix, we state the formulas we use to calculate imaginary time and imaginary frequency grids. For derivations, we refer to ref 36. Since we want to switch between imaginary time and imaginary frequency, we write the discrete FT $f(i\omega_k)$ of an arbitrary function $f(i\tau_j)$ as

$$f(i\omega_k) = -i\sum_j^{N_\tau} \{\gamma_{kj}^{(c)} \cos(\omega_k\tau_j)(f(i\tau_j) + f(-i\tau_j)) - i\gamma_{kj}^{(s)} \sin(\omega_k\tau_j)(f(i\tau_j) - f(-i\tau_j))\} \quad (17)$$

FT from imaginary frequency to imaginary time can be done by inverting the matrices with the elements $\gamma_{kj}^{(c)} \cos(\omega_k\tau_j)$ and $\gamma_{kj}^{(s)} \sin(\omega_k\tau_j)$ (in case $N_\tau \neq N_\omega$, a pseudoinverse can be calculated). At the beginning of a GW calculation, we find optimal sets of imaginary frequency $\{\omega_k\}_{k=1,\ldots,N_\omega}$ and imaginary time points $\{\tau_j\}_{j=1,\ldots,N_\tau}$. For imaginary time, we use the algorithm described in ref 114 to minimize

$$\left\| \eta^{(\tau)}(x;\alpha,\tau) \right\|_\infty, \qquad \eta^{(\tau)}(x;\alpha,\tau) = \frac{1}{2x} - \sum_j^{N_\tau} \alpha_j e^{-2x\tau_j} \quad (18)$$

and for imaginary frequency, we first minimize

$$\left\| \eta^{(\omega)}(x;\sigma,\omega) \right\|_2,$$

$$\eta^{(\omega)}(x;\sigma,\omega) = \frac{1}{x} - \frac{1}{\pi}\sum_k^{N_\omega} \sigma_k \left(\frac{2x}{x^2 + {\omega_k}^2}\right)^2 \quad (19)$$

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

using a Levenberg−Marquardt algorithm (LMA). In both minimizations, x denotes the electron−hole pair transition energy; both, $\eta^{(\tau)}$ and $\eta^{(\omega)}$ are minimized in the interval $[1, x_{min}/x_{max}]$, with x_{min} (x_{max}) being the smallest (largest) considered electron−hole transition energy. In both optimizations, the interval $[x_{min}, x_{max}]$ is represented on a logarithmic grid. After minimizing the L_2 norm on the logarithmic grid, we minimize the L_2 norm for the positions of the minima and maxima and repeat this procedure until self-consistency is reached. The coefficients $\{\sigma_k\}\}_{k=1,...,N\omega}$ and $\{\{\alpha_j\}\}_{j=1,...,N\tau}$ are byproducts of this optimization, but they can be used to evaluate MP2 or RPA correlation energies.

$\eta^{(\omega)}$ has multiple minima, and especially for larger N_ω, the LMA only converges to a useful minimum when it is initialized with σ and ω, which are already sufficiently close to the parameters which minimize $\eta^{(\omega)}$. In practice, we found that at a useful (but not necessarily global) minimum, $\eta^{(\omega)}$ has $2N_\omega$ or $2N_\omega - 1$ extrema. This behavior can be exploited to find good starting points for the LMA for different ratios x_{min}/x_{max} and N_ω. We always start the LMA from pretabulated values.[115] These values have been obtained by a simple metropolis algorithm for several x_{min}/x_{max} and for N_ω between 1 and 40 and are chosen so that $\eta^{(\omega)}$ has $2N_\omega$ or $2N_\omega - 1$ extrema for a given N_ω. In a GWA calculation, the LMA is then initialized with the pretabulated parameters which are closest to the x_{min}/x_{max} of the calculation.

To avoid unnecessarily large grids, the N_ω and N_τ are determined at run time and the user only specifies an upper limit of points for both, imaginary time and frequency grids. More precisely, for small N_ω, grid points and weights are calculated and the L_2 norm of η^{ω} is calculated. Then, we increase the imaginary frequency grid until the least square error is smaller than $\epsilon_\omega = 1e^{-10}$. The number of points which are required to reach this accuracy strongly depends on x_{min}/x_{max}. In our experience, the QP energies converge faster with respect to the imaginary time grid than with respect to the ω-grid; since we find it convenient to work with grids of the same size (although this is not necessary), we set $N_\tau = N_\omega$. For example, for a hydrogen molecule in a triple-ζ (TZ) basis, $N_\omega = 7$ will already be sufficient to reach the desired accuracy, while for the iodine molecule in a QZ basis, $N_\omega = 31$ will be necessary.

To calculate the matrices $\gamma^{(c)}$, we minimize

$$\left\| \eta^{(c)}(x;\gamma^{(c)}) \right\|_2 ,$$

$$\eta^{(c)}(x,\gamma^{(c)}) = \frac{2x}{x^2+\omega_k^2} - \sum_{j=1}^{N_\tau} \gamma_{kj}^{(c)} \cos(\omega_k\tau_j) e^{-x\tau_j} \tag{20}$$

for all ω_k and for $\gamma^{(s)}$, we minimize

$$\left\| \eta^{(s)}(x;\gamma^{(s)}) \right\|_2 ,$$

$$\eta^{(s)}(x,\gamma^{(s)}) = \frac{2\omega_k}{x^2+\omega_k^2} - \sum_{j=1}^{N_\tau} \gamma_{kj}^{(c)} \sin(\omega_k\tau_j) e^{-x\tau_j} \tag{21}$$

with a LMA. The procedure is the same as described in ref 21. In all cases, the algorithm converges smoothly from arbitrary starting values.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00308.

KS HOMO and LUMO eigenvalues, QP energies for finite basis sets, raw data for the subset of 250 molecules from GW5000, and technical parameters of all calculations for GW100 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Arno Förster − *Theoretical Chemistry, Vrije Universiteit, NL-1081 HV Amsterdam, The Netherlands;* orcid.org/0000-0002-0957-4081; Email: a.t.l.foerster@vu.nl

Author

Lucas Visscher − *Theoretical Chemistry, Vrije Universiteit, NL-1081 HV Amsterdam, The Netherlands;* orcid.org/0000-0002-7748-6243

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c00308

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research received funding from the Netherlands Organisation for Scientific Research (NWO) in the framework of the Innovation Fund for Chemistry and from the Ministry of Economic Affairs in the framework of the "TKI/PPS-Toeslagregeling". The authors thank Erik van Lenthe for fruitful discussion.

■ REFERENCES

(1) Hedin, L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **1965**, *139*, A796.

(2) Blase, X.; Attaccalite, C.; Olevano, V. First-principles GW calculations for fullerenes, porphyrins, phtalocyanine, and other molecules of interest for organic photovoltaic applications. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2011**, *83*, 115103.

(3) Faber, C.; Attaccalite, C.; Olevano, V.; Runge, E.; Blase, X. First-principlesGWcalculations for DNA and RNA nucleobases. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2011**, *83*, 115123.

(4) 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 J. Phys.* **2012**, *14*, 053020.

(5) Bruneval, F.; Marques, M. A. L. Benchmarking the starting points of the GW approximation for molecules. *J. Chem. Theory Comput.* **2013**, *9*, 324−329.

(6) Van Setten, M. J.; Weigend, F.; Evers, F. The GW-method for quantum chemistry applications: Theory and implementation. *J. Chem. Theory Comput.* **2013**, *9*, 232−246.

(7) Van Setten, M. J.; Caruso, F.; Sharifzadeh, S.; Ren, X.; Scheffler, M.; Liu, F.; Lischner, J.; Lin, L.; Deslippe, J. R.; Louie, S. G.; Yang, C.; Weigend, F.; Neaton, J. B.; Evers, F.; Rinke, P. GW100: Benchmarking G0W0 for Molecular Systems. *J. Chem. Theory Comput.* **2015**, *11*, 5665−5687.

(8) Knight, J. W.; Wang, X.; Gallandi, L.; Dolgounitcheva, O.; Ren, X.; Ortiz, J. V.; Rinke, P.; Körzdörfer, T.; Marom, N. Accurate Ionization Potentials and Electron Affinities of Acceptor Molecules III: A Benchmark of GW Methods. *J. Chem. Theory Comput.* **2016**, *12*, 615−626.

(9) Golze, D.; Dvorak, M.; Rinke, P. The GW Compendium: A Practical Guide to Theoretical Photoemission Spectroscopy. *Front. Chem.* **2019**, *7*, 377.
(10) Caruso, F.; Dauth, M.; Van Setten, M. J.; Rinke, P. Benchmark of GW Approaches for the GW100 Test Set. *J. Chem. Theory Comput.* **2016**, *12*, 5076−5087.
(11) Stuke, A.; Kunkel, C.; Golze, D.; Todorović, M.; Margraf, J. T.; Reuter, K.; Rinke, P.; Oberhofer, H. Atomic structures and orbital energies of 61,489 crystal-forming organic molecules. *Sci. Data* **2020**, *7*, 58.
(12) Bruneval, F.; Maliyov, I.; Lapointe, C.; Marinica, M.-C. Extrapolating unconverged GW energies up to the complete basis set limit with linear regression. *J. Chem. Theory Comput.* **2020**, *16*, 4399.
(13) Balasubramani, S. G.; Chen, G. P.; Coriani, S.; Diedenhofen, M.; Frank, M. S.; Franzke, Y. J.; Furche, F.; Grotjahn, R.; Harding, M. E.; Hättig, C.; Hellweg, A.; Helmich-Paris, B.; Holzer, C.; Huniar, U.; Kaupp, M.; Marefat Khah, A.; Karbalaei Khani, S.; Müller, T.; Mack, F.; Nguyen, B. D.; Parker, S. M.; Perlt, E.; Rappoport, D.; Reiter, K.; Roy, S.; Rückert, M.; Schmitz, G.; Sierka, M.; Tapavicza, E.; Tew, D. P.; van Wüllen, C.; Voora, V. K.; Weigend, F.; Wodyński, A.; Yu, J. M. TURBOMOLE : Modular program suite for ab initio quantum-chemical and condensed-matter simulations. *J. Chem. Phys.* **2020**, *152*, 184107.
(14) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. The Fritz Haber Institute ab initio molecular simulations package (FHI-aims). 2009, http://www.fhi-berlin.mpg.de/aims (accessed March 24, 2021).
(15) 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. Commun.* **2009**, *180*, 2175−2196.
(16) Hybertsen, M. S.; Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *34*, 5390.
(17) Deslippe, J.; Samsonidze, G.; Strubbe, D. A.; Jain, M.; Cohen, M. L.; Louie, S. G. BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. *Comput. Phys. Commun.* **2012**, *183*, 1269−1289.
(18) Kresse, G.; Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47*, 558−561.
(19) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 11169−11186.
(20) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15−50.
(21) Liu, P.; Kaltak, M.; Klimeš, J.; Kresse, G. Cubic scaling GW: Towards fast quasiparticle calculations. *Phys. Rev. B* **2016**, *94*, 165109.
(22) Maggio, E.; Liu, P.; Van Setten, M. J.; Kresse, G. GW100: A Plane Wave Perspective for Small Molecules. *J. Chem. Theory Comput.* **2017**, *13*, 635−648.
(23) Govoni, M.; Galli, G. Large Scale GW Calculations. *J. Chem. Theory Comput.* **2015**, *11*, 2680−2696.
(24) Govoni, M.; Galli, G. GW100: Comparison of Methods and Accuracy of Results Obtained with the WEST Code. *J. Chem. Theory Comput.* **2018**, *14*, 1895−1909.
(25) Tiago, M. L.; Chelikowsky, J. R. Optical excitations in organic molecules, clusters, and defects studied by first-principles Green's function methods. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2006**, *73*, 205334.
(26) Gao, W.; Chelikowsky, J. R. Real-Space Based Benchmark of G0W0 Calculations on GW100: Effects of Semicore Orbitals and Orbital Reordering. *J. Chem. Theory Comput.* **2019**, *15*, 5299−5307.
(27) Wilhelm, J.; Golze, D.; Talirz, L.; Hutter, J.; Pignedoli, C. A. Toward GW Calculations on Thousands of Atoms. *J. Phys. Chem. Lett.* **2018**, *9*, 306−312.
(28) Wilhelm, J.; Seewald, P.; Golze, D. Low-scaling GW with benchmark accuracy and application to phosphorene nanosheets. *J. Chem. Theory Comput.* **2021**, *17*, 1662−1677.
(29) Förster, A.; Visscher, L. Low-Order Scaling G0W0 by Pair Atomic Density Fitting. *J. Chem. Theory Comput.* **2020**, *16*, 7381−7399.
(30) Duchemin, I.; Blase, X. Cubic-scaling all-electron GW calculations with a separable density-fitting space-time approach. *J. Chem. Theory Comput.* **2021**, *17*, 2383.
(31) Bruneval, F.; Rangel, T.; Hamed, S. M.; Shao, M.; Yang, C.; Neaton, J. B. MOLGW 1: Many-body perturbation theory software for atoms, molecules, and clusters. *Comput. Phys. Commun.* **2016**, *208*, 149−161.
(32) Kühne, T. D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F.; Golze, D.; Wilhelm, J.; Chulkov, S.; Bani-Hashemian, M. H.; Weber, V.; Borštnik, U.; Taillefumier, M.; Jakobovits, A. S.; Lazzaro, A.; Pabst, H.; Müller, T.; Schade, R.; Guidon, M.; Andermatt, S.; Holmberg, N.; Schenter, G. K.; Hehn, A.; Bussy, A.; Belleflamme, F.; Tabacchi, G.; Glöß, A.; Lass, M.; Bethune, I.; Mundy, C. J.; Plessl, C.; Watkins, M.; VandeVondele, J.; Krack, M.; Hutter, J. CP2K: An Electronic Structure and Molecular Dynamics Software Package − Quickstep: Efficient and Accurate Electronic Structure Calculations. *J. Chem. Phys.* **2020**, *152*, 194103.
(33) 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.
(34) Rojas, H. N.; Godby, R. W.; Needs, R. J. Space-Time Method for Ab Initio Calculations ofSelf-Energies and Dielectric Response Functions of Solids. *Phys. Rev. Lett.* **1995**, *74*, 1827−1830.
(35) Rieger, M. M.; Steinbeck, L.; White, I. D.; Rojas, H. N.; Godby, R. W. GW space-time method for the self-energy of large systems. *Comput. Phys. Commun.* **1999**, *117*, 211−228.
(36) Kaltak, M.; Klimeš, J.; Kresse, G. Low scaling algorithms for the random phase approximation: Imaginary time and laplace transformations. *J. Chem. Theory Comput.* **2014**, *10*, 2498−2507.
(37) Kaltak, M.; Klimeš, J.; Kresse, G. Cubic scaling algorithm for the random phase approximation: Self-interstitials and vacancies in Si. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *90*, 054115.
(38) Van Lenthe, E.; Baerends, E. J. Optimized Slater-type basis sets for the elements 1−118. *J. Comput. Chem.* **2003**, *24*, 1142−1156.
(39) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N2 , and H2O. *Chem. Phys. Lett.* **1998**, *286*, 243−252.
(40) Jensen, F. Atomic orbital basis sets. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, 3, 273−295.
(41) 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.
(42) Rezaei, M.; Öğüt, S. Photoelectron spectra of early 3d-transition metal dioxide molecular anions from GW calculations. *J. Chem. Phys.* **2021**, *154*, 094307.
(43) Krause, K.; Harding, M. E.; Klopper, W. Coupled-cluster reference values for the GW27 and GW100 test sets for the assessment of GW methods. *Mol. Phys.* **2015**, *113*, 1952−1960.
(44) Lange, M. F.; Berkelbach, T. C. On the Relation between Equation-of-Motion Coupled-Cluster Theory and the GW Approximation. *J. Chem. Theory Comput.* **2018**, *14*, 4224−4236.
(45) Hirata, S.; Doran, A. E.; Knowles, P. J.; Ortiz, J. V. One-particle many-body Green's function theory: Algebraic recursive definitions, linked-diagram theorem, irreducible-diagram theorem, and general-order algorithms. *J. Chem. Phys.* **2017**, *147*, 044108.
(46) Martin, R. M.; Reining, L.; Ceperley, D. M. *Interacting Electrons*; Cambridge University Press, 2016.
(47) Seidl, A.; Görling, A.; Vogl, P.; Majewski, J. A.; Levy, M. Generalized Kohn-Sham schemes and the band-gap problem. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *53*, 3764−3774.

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

(48) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133.
(49) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, 864−871.
(50) Engel, E.; Dreizler, R. M. *Density Functional Theory an Advanced Course*; Springer, 2013.
(51) Vidberg, H. J.; Serene, J. W. Solving the Eliashberg equations by means of N-point Padé approximants. *J. Low Temp. Phys.* **1977**, *29*, 179−192.
(52) Kopietz, P. *Bosonization of Interacting Fermions in Arbitrary Dimensions*; Springer, 1997; Chapter 1
(53) Mattuck, R. D. *A Guide to Feynman Diagrams in the Many-body Problem*, 2nd ed.; Dover Publications Inc.: New York, 1992.
(54) Cancés, E.; Gontier, D.; Stoltz, G. A mathematical analysis of the GW0 method for computing electronic excited energies of molecules. *Rev. Math. Phys.* **2016**, *28*, 1650008.
(55) Almlöf, J. Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319−320.
(56) Takatsuka, A.; Ten-no, S.; Hackbusch, W. Minimax approximation for the decomposition of energy denominators in Laplace-transformed Møller−Plesset perturbation theories. *J. Chem. Phys.* **2008**, *129*, 044112.
(57) Billingsley, F. P.; Bloor, J. E. Limited expansion of diatomic overlap (LEDO): A Near-accurate approximate Ab Initio LCAO MO Method. I. Theory and preliminary investigations. *J. Chem. Phys.* **1971**, *55*, 5178−5190.
(58) Baerends, E. J.; Ellis, D. E.; Ros, P. Self-consistent molecular Hartree—Fock—Slater calculations I. The computational procedure. *Chem. Phys.* **1973**, *2*, 41−51.
(59) Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496−4501.
(60) Sambe, H.; Felton, R. H. A new computational approach to Slater's SCF-X α equation. *J. Chem. Phys.* **1975**, *62*, 1122−1126.
(61) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On some approximations in applications of Xα theory. *J. Chem. Phys.* **1979**, *71*, 3396−3402.
(62) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On first-row diatomic molecules and local density models. *J. Chem. Phys.* **1979**, *71*, 4993−4999.
(63) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514−518.
(64) Förster, A.; Franchini, M.; van Lenthe, E.; Visscher, L. A Quadratic Pair Atomic Resolution of the Identity Based SOS-AO-MP2 Algorithm Using Slater Type Orbitals. *J. Chem. Theory Comput.* **2020**, *16*, 875−891.
(65) Cohen, A. J.; Handy, N. C. Density functional generalized gradient calculations using Slater basis sets. *J. Chem. Phys.* **2002**, *117*, 1470−1478.
(66) Watson, M. A.; Handy, N. C.; Cohen, A. J. Density functional calculations, using Slater basis sets, with exact exchange. *J. Chem. Phys.* **2003**, *119*, 6475−6481.
(67) Watson, M. A.; Handy, N. C.; Cohen, A. J.; Helgaker, T. Density-functional generalized-gradient and hybrid calculations of electromagnetic properties using slater basis sets. *J. Chem. Phys.* **2004**, *120*, 7252−7261.
(68) Chong, D. P.; Van Lenthe, E.; Van Gisbergen, S.; Baerends, E. J. Even-tempered slater-type orbitals revisited: From hydrogen to krypton. *J. Comput. Chem.* **2004**, *25*, 1030−1036.
(69) Zhang, I. Y.; Ren, X.; Rinke, P.; Blum, V.; Scheffler, M. Numeric atom-centered-orbital basis sets with valence-correlation consistency from H to Ar. *New J. Phys.* **2013**, *15*, 123033.
(70) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639−9646.
(71) Schwartz, C. Importance of Angular Correlations between Atomic Electrons. *Phys. Rev.* **1962**, *126*, 1015−1019.
(72) Hill, R. N. Rates of convergence and error estimation formulas for the Rayleigh-Ritz variational method. *J. Chem. Phys.* **1985**, *83*, 1173−1196.
(73) Kutzelnigg, W.; Morgan, J. D., III Rates of convergence of the partial-wave expansions of atomic correlation energies correlation energies. *J. Chem. Phys.* **1992**, *96*, 4484.
(74) Dyall, K. G. Regular article Relativistic and nonrelativistic finite nucleus optimized double zeta basis sets for the 4 p, 5 p and 6 p elements. *Theor. Chem. Acc.* **1998**, *99*, 366−371.
(75) Dyall, K. G. Regular article Relativistic and nonrelativistic finite nucleus optimized triple-zeta basis sets for the 4 p , 5 p and 6 p elements. *Theor. Chem. Acc.* **2002**, *108*, 335−340.
(76) Dyall, K. G. Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the 5 d elements Hf − Hg. *Theor. Chem. Acc.* **2004**, *112*, 403−409.
(77) Dyall, K. G. Relativistic quadruple-zeta and revised triple-zeta and double-zeta basis sets for the 4p , 5p , and 6p elements. *Theor. Chem. Acc.* **2006**, *115*, 441−447.
(78) Gomes, A. S. P.; Dyall, K. G.; Visscher, L. basis sets for the lanthanides La − Lu. *Theor. Chem. Acc.* **2010**, *127*, 369−381.
(79) Autschbach, J.; Ziegler, T. Nuclear spin-spin coupling constants from regular approximate relativistic density functional calculations. II. Spin-orbit coupling effects and anisotropies. *J. Chem. Phys.* **2000**, *113*, 9410−9418.
(80) Reeves, C. M.; Harrison, M. C. Use of gaussian functions in the calculation of wavefunctions for small molecules. II. The ammonia molecule. *J. Chem. Phys.* **1963**, *39*, 1−10.
(81) Silver, D. M.; Nieuwpoort, W. C. Universal atomic basis sets. *Chem. Phys. Lett.* **1978**, *57*, 421−422.
(82) Ren, X.; Merz, F.; Jiang, H.; Yao, Y.; Rampp, M.; Lederer, H.; Blum, V.; Scheffler, M. All-electron periodic G0W0 implementation with numerical atomic orbital basis functions: Algorithm and benchmarks. *Phys. Rev. Mater.* **2021**, *5*, 013807.
(83) Schulz, G. J. Resonances in electron impact on diatomic molecules. *Rev. Mod. Phys.* **1973**, *45*, 423−486.
(84) Jordan, K. D.; Burrow, P. D. Studies of the Temporary Anion States of Unsaturated Hydrocarbons by Electron Transmission Spectroscopy. *Acc. Chem. Res.* **1978**, *11*, 341−348.
(85) Aflatooni, K.; Gallup, G. A.; Burrow, P. D. Electron attachment energies of the DNA bases. *J. Phys. Chem. A* **1998**, *102*, 6205−6207.
(86) Tozer, D. J.; De Proft, F. Computation of the Hardness and the Problem of Negative Electron Affinities in Density Functional Theory. *J. Phys. Chem. A* **2005**, *109*, 8923−8929.
(87) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(88) Zheng, J.; Xu, X.; Truhlar, D. G. Minimally augmented Karlsruhe basis sets. *Theor. Chem. Acc.* **2011**, *128*, 295−305.
(89) This is slightly different from usual augmentation approaches encountered for Gaussian-type basis sets. Dunning and coworkers defined the prefix "-aug" to mean adding an additional diffuse functions for all angular momenta already present in a basis set for every atom type.[87] On the contrary, in a process which they denoted as minimal augmentation, Truhlar and coworkers only added diffuse s and p functions to all elements heavier than hydrogen.[88] Our approach can be seen as a compromise between both.
(90) Chong, D. P. Augmenting basis set for time-dependent density functional theory calculation of excitation energies: Slater-type orbitals for hydrogen to krypton. *Mol. Phys.* **2005**, *103*, 749−761.
(91) Baerends, E.; Ziegler, T.; Atkins, A.; Autschbach, J.; Baseggio, O.; Bashford, D.; Bérces, A.; Bickelhaupt, F.; Bo, C.; Boerrigter, P.; Cavallo, L.; Daul, C.; Chong, D.; Chulhai, D.; Deng, L.; Dickson, R.; Dieterich, J.; Ellis, D.; van Faassen, M.; Fan, L.; Fischer, T.; Förster, A.; Guerra, C. F.; Franchini, M.; Ghysels, A.; Giammona, A.; van Gisbergen, S.; Goez, A.; Götz, A.; Groeneveld, J.; Gritsenko, O.; Grüning, M.; Gusarov, S.; Harris, F.; van den Hoek, P.; Hu, Z.; Jacob, C.; Jacobsen, H.; Jensen, L.; Joubert, L.; Kaminski, J.; van Kessel, G.; König, C.; Kootstra, F.; Kovalenko, A.; Krykunov, M.; van Lenthe, E.; McCormack, D.; Michalak, A.; Mitoraj, M.; Morton, S.; Neugebauer,

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097

J.; Nicu, V.; Noodleman, L.; Osinga, V.; Patchkovskii, S.; Pavanello, M.; Peeples, C.; Philipsen, P.; Post, D.; Pye, C.; Ramanantoanina, H.; Ramos, P.; Ravenek, W.; Reimann, M.; Rodríguez, J.; Ros, P.; Rüger, R.; Schipper, P.; Schlüns, D.; van Schoot, H.; Schreckenbach, G.; Seldenthuis, J.; Seth, M.; Snijders, J.; Solà, M.; Stener, M.; Swart, M.; Swerhone, D.; Tognetti, V.; te Velde, G.; Vernooijs, P.; Versluis, L.; Visscher, L.; Visser, O.; Wang, F.; Wesolowski, T.; van Wezenbeek, E.; Wiesenekker, G.; Wolff, S.; Woo, T.; Yakovlev, A. *ADF2020, Locally Modified Development Version*.

(92) The code to generate the imaginary time grids can be found on Github. (https://github.com/bhelmichparis/laplace-minimax). Please also see refs 56 and 114.

(93) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(94) Perdew, J. P.; Burke, K.; Ernzerhof, M. [ERRATA] Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(95) Data downloaded from the website of the GW100 project by van Setten et al., https://gw100.wordpress.com.

(96) Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew−Burke−Ernzerhof exchange-correlation functional. *J. Chem. Phys.* **1999**, *110*, 5029.

(97) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.

(98) Lenthe, E. v.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597.

(99) Van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic total energy using regular approximations. *J. Chem. Phys.* **1994**, *101*, 9783−9792.

(100) Van Lenthe, E.; Snijders, J. G.; Baerends, E. J. The zero-order regular approximation for relativistic effects: The effect of spin-orbit coupling in closed shell molecules. *J. Chem. Phys.* **1996**, *105*, 6505−6516.

(101) Van Lenthe, E.; Van Leeuwen, R.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component hamiltonians. *Int. J. Quantum Chem.* **1996**, *57*, 281−293.

(102) Wilhelm, J.; Del Ben, M.; Hutter, J. GW in the Gaussian and Plane Waves Scheme with Application to Linear Acenes. *J. Chem. Theory Comput.* **2016**, *12*, 3623−3635.

(103) Jensen, F. Polarization consistent basis sets: II. Estimating the Kohn-Sham basis set limit. *J. Chem. Phys.* **2002**, *116*, 7372−7379.

(104) Kraus, P. Basis Set Extrapolations for Density Functional Theory. *J. Chem. Theory Comput.* **2020**, *16*, 5712−5722.

(105) Dunlap, B. I.; Mei, W. N. Basis set effects on spectroscopic constants for C2 and Si2 and the symmetry dilemma in the Xα model. *J. Chem. Phys.* **1983**, *78*, 4997.

(106) The basis set files can be downloaded from http://www.scm.com/wp-content/uploads/Correlation-basis-sets.zip.

(107) Bruneval, F. Optimized virtual orbital subspace for faster GW calculations in localized basis. *J. Chem. Phys.* **2016**, *145*, 234110.

(108) Riemelmoser, S.; Kaltak, M.; Kresse, G. Plane wave basis set correction methods for RPA correlation energies. *J. Chem. Phys.* **2020**, *152*, 134103.

(109) Loos, P.-F.; Pradines, B.; Scemama, A.; Giner, E.; Toulouse, J. Density-Based Basis-Set Incompleteness Correction for GW Methods. *J. Chem. Theory Comput.* **2020**, *16*, 1018−1028.

(110) Boulanger, P.; Jacquemin, D.; Duchemin, I.; Blase, X. Fast and accurate electronic excitations in cyanines with the many-body bethe-salpeter approach. *J. Chem. Theory Comput.* **2014**, *10*, 1212−1218.

(111) Faber, C.; Boulanger, P.; Attaccalite, C.; Cannuccia, E.; Duchemin, I.; Deutsch, T.; Blase, X. Exploring approximations to the GW self-energy ionic gradients. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2015**, *91*, 155109.

(112) Jacquemin, D.; Duchemin, I.; Blase, X. Benchmarking the Bethe-Salpeter Formalism on a Standard Organic Molecular Set. *J. Chem. Theory Comput.* **2015**, *11*, 3290−3304.

(113) Bruneval, F.; Hamed, S. M.; Neaton, J. B. A systematic benchmark of the ab initio Bethe-Salpeter equation approach for low-lying optical excitations of small organic molecules. *J. Chem. Phys.* **2015**, *142*, 244101.

(114) Helmich-Paris, B.; Visscher, L. Improvements on the minimax algorithm for the Laplace transformation of orbital energy denominators. *J. Comput. Phys.* **2016**, *321*, 927−931.

(115) The pretabulated values are available on Github (https://github.com/ArnoFoerster/Imaginary-Frequency-Grids-GW-and-RPA).

https://doi.org/10.1021/acs.jctc.1c00308
J. Chem. Theory Comput. 2021, 17, 5080−5097