Optimally tuned range-separated hybrid van der Waals density functional for molecular binding and quasiparticle characterizations

PERSPECTIVE • **OPEN ACCESS**

Optimally tuned range-separated hybrid van der Waals density functional for molecular binding and quasiparticle characterizations

OPEN ACCESS
IOP Publishing
Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **37** (2025) 211501 (35pp)
https://doi.org/10.1088/1361-648X/add0fd

Perspective

Optimally tuned range-separated hybrid van der Waals density functional for molecular binding and quasiparticle characterizations

Elsebeth Schröder*, **Raul Quintero-Monsebaiz**, **Yang Jiao** and **Per Hyldgaard***
Department of Microtechnology and Nanoscience—MC2, Chalmers University of Technology, SE-41296 Gothenburg, Sweden
E-mail: schroder@chalmers.se and *hyldgaar@chalmers.se*

Abstract
We introduce and illustrate use of two closely related range-separated hybrid (RSH) van der Waals density functionals (vdW-DFs), denoted AHBR-mRSH and AHBR-mRSH*, for total-energy and quasiparticle characterizations of molecules using generalized Kohn–Sham (gKS) density functional theory (DFT). For comparison, we also introduce and document a traditional design for a long-range corrected (LRC) vdW-DF, denoted 'B86R-LRC'. All three new vdW-DFs set the exchange potential as free of asymptotic screening in the coupling between the electron and its associated exchange hole. Our two AHBR-mRSHs are key members of a broader class 'AHBR-mRSH(γ)' defined by an inverse length scale γ for a crossover in weighting short- and long-ranged exchange contributions; we obtain a highly accurate predictor of general molecular-energy differences by keeping $\gamma = 0.106$ (inverse Bohr) deliberately fixed in AHBR-mRSH. We obtain an optimally tuned (OT) form AHBR-mRSH* = AHBR-mRSH($\gamma*$) by computing a plausible value of $\gamma*$ for specific types of systems. This AHBR-mRSH* permits characterizations of molecular quasiparticles and generalizes (1) the existing 'OT-RSH' (Stein *et al* 2010 *Phys. Rev. Lett.* **105** 266802; Rafaely–Abramson *et al* 2012 *Phys. Rev. Lett.* **109** 226405) approach by a systematic inclusion of truly nonlocal correlations, and (2) more traditional LRC forms (e.g. B86R-LRC) by setting the short-range exchange description as in vdW-DF2-ahbr (Shukla *et al* 2022 *Phys. Rev. X* **12** 041003). *Importantly,* we may view AHBR-mRSH and AHBR-mRSH* as internally consistent functionals for gKS-DFT, being simultaneously accurate on molecular energies and quasiparticles. This is possible because the 'AHBR-mRSH(γ)' class has enough transferability to almost always limit adverse impacts of tuning γ, as tested here on the GMTKN55 benchmark suite (Goerigk *et al* 2017 *Phys. Chem. Chem. Phys.* **19** 32184). We find that AHBR-mRSH generally outperforms B86R-LRC on molecular problems. To illustrate usage, we complete the OT design of an AHBR-mRSH* for nucleobases and show that it provides quasiparticle predictions that are in good agreement with both literature theory and experimental values for adenine, thymine, cytosine, and guanine.

Supplementary material for this article is available online

Keywords: range-separated hybrid, van der Waals, density functional, molecular quasiparticles, generalized Kohn Sham, density functional theory

* Authors to whom any correspondence should be addressed.

1. Introduction

Hybrid forms of the van der Waals (vdW) density functional (vdW-DF) method [1–5] are delivering a robust performance across many types of problems [6–16]. One motivation for using hybrid vdW-DFs is that density functional theory (DFT) calculations in the Kohn–Sham (KS) framework, with density-explicit functional approximations for the exchange-correlation (XC) energy, can lead to self-interaction and delocalization errors in the total-energy predictions [17–24]. Use of density-explicit vdW-DFs, like vdW-DF-cx [25] (here abbreviated CX) and rev-vdW-DF2 [26] (also called vdW-DF2-b86r and here abbreviated B86R) often produce accurate DFT results [5, 10, 27–29] but can fail on predicting spin configurations and small-molecule adsorption with frustrated electronic configurations [12, 13, 16, 20, 30]. Two new range-separated hybrid (RSH) vdW-DFs, denoted AHCX and AHBR, are helping to improve DFT on some of those problems, e.g. getting both substrate structure and site-preference right for CO adsorption on noble metals and on Pt, correctly reflecting the anti-ferromagnetic ordering of (reactive) copper-peroxide complexes in porous materials, and minimizing errors in some problems of biochemistry relevance [10, 11, 13, 16].

Motivation for expanding use of RSH vdW-DFs exists because they may serve as DFT-based predictors of some quasiparticles [31–36]. To this challenge they bring a consistent inclusion of the truly-nonlocal correlation and weak-interaction effects [3, 5, 13] that often define the nature of molecular systems.

Quasiparticle orbitals, denoted $\Psi_\xi(\mathbf{r})$ or $\Psi_\eta(\mathbf{r})$, characterize the effects of suddenly adding or removing an electron to or from the N-electron ground state $|\Psi_0(N)\rangle$ [37, 38]. They are wavefunctions given by the state overlap of such sudden excitations with a specific $N+1$ state $|\Psi_\xi\rangle$ or $N-1$ state $|\Psi_\eta\rangle$ and we can ascribe orbital energy levels, $\epsilon_\xi^{\rm QP} = -\Delta E_\xi$ or $\epsilon_\eta^{\rm QP} = -\Delta E_\eta$ given by the energy differences, ΔE_ξ or ΔE_η, between these specific excitations and the initial state. This holds at least when ΔE_ξ or ΔE_η remain small, corresponding to so-called frontier orbitals. We may, for example, seek the highest-occupied molecular orbital (HOMO) $\Psi_H(\mathbf{r}) \equiv \langle\Psi_{\eta'}|\hat{\psi}(\mathbf{r})|\Psi_0(N)\rangle$ where $\hat{\psi}(\mathbf{r})$ denotes the operator for electron removal at $\mathbf{r}$ and where η' identifies the $N-1$ state with the minimal value of ΔE_η, thus setting the HOMO level $\epsilon_H^{\rm QP}$. When instead the electron-creation operator $\hat{\psi}^+(\mathbf{r})$ defines the ground-state excitation we find the lowest unoccupied molecular orbital (LUMO) $\Psi_L(\mathbf{r})$ at orbital level $\epsilon_L^{\rm QP}$ (as well as higher-energy LUMO+ quasiparticles). A search for such quasiparticles traditionally requires systematic, costly many-body perturbation theory (MBPT) calculations [39–44]. Meanwhile, in the traditional KS scheme [45, 46] for DFT we work with independent particles moving in a local effective potential. This potential is set by the so-called XC energy [45, 47] that we, in practice, describe in density functional approximations (DFAs) of increasing complexity and capabilities, for example, [1, 5, 23, 24, 48–59]. From the resulting KS-DFA solutions, we can extract correspondingly defined independent-particle orbitals, e.g. $\phi_H(\mathbf{r})$ and $\phi_L(\mathbf{r})$ at level energies ϵ_H and ϵ_L. However, even if we had access to the exact XC functional and could thus complete formally exact KS-DFT studies, we would not have alignment between these independent-particle orbital levels and the quasiparticle levels [21, 31, 46, 60–65]. As discussed, for example, in [21], the limitation of KS-DFT arises because the local potential must contain a derivative discontinuity [31, 32, 64, 66, 67] to accurately reflect, for example, the HOMO-LUMO gap. Nevertheless, it is possible to seek systematic generalizations of [32, 36, 68–70] to navigate this problem: An optimally tuned

(OT) RSH vdW-DF can describe frontier quasiparticles also in weakly interacting systems.

RSH vdW-DFs open for predictions of quasiparticles, and hence of optical-response properties, because we pursue such calculations within the generalized KS (denoted 'gKS') DFT-solution scheme [21, 46]. That is, we have AHCX and AHBR implemented in QUANTUMESPRESSO [71–74] (QE) which uses orbital-based optimization for the Fock-exchange component; the same holds for the semilocal-correlation HSE [32, 57, 75–77]. Compared with a would-be local-potential implementation [21] the RSHs are thus moved closer in nature to density-matrix formulations of the electronic-structure problem [35, 78–87]. The RSHs of QE are gKS-DFAs and can help us to bypass an explicit compensation for the derivative discontinuity [21, 31, 32, 66, 67].

We note that screening implicitly enters the exchange description [21, 68] in most DFAs as they are at least partly based on the electron-gas tradition [1, 4, 5, 41, 48–52, 88, 89]. The resulting local DFA potentials do not reflect the correct, asymptotically unscreened, nature of electrostatics interactions in molecular system [21, 61, 90, 91] and this issue also affects the RSHs [10, 13, 57, 77] in general. For a compact presentation, we describe exchange in the RSHs in analogy to the formally exact formulation of the XC energy [49, 50],

$$E_{xc} = \frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} n(\mathbf{r})\, n_{xc}\left(\mathbf{r};|\mathbf{r}-\mathbf{r}'|\right) V_{\mathbf{r},\mathbf{r}'}\,. \tag{1}$$

Here $V_{\mathbf{r},\mathbf{r}'} = |\mathbf{r}-\mathbf{r}'|^{-1}$ denotes the Coulomb matrix element and n_{xc} denotes the XC hole [49, 51]. The exchange component E_x of E_{xc} can be represented similarly in a generalized gradient approximation (GGA) (like PBE [89]) and in regular vdW-DFs (like CX, and B86R) by inserting an exchange-hole model [10, 13, 57, 76, 77, 92, 93], denoted $n_x(\mathbf{r};|\mathbf{r}-\mathbf{r}'|)$, into equation (1). To define exchange in corresponding RSHs, we insert instead a form $\tilde{n}_x^{\gamma}$ that is a linear combination of the Fock-exchange expression n_x^{Fo} [23] and n_x [7, 10, 57, 77]. As implied by the superscript, the RSH design permits a cross-over between a short- and long-ranges behavior, as controlled by a choice of an inverse length scale γ.

The extent of screening in the exchange description of DFAs and RSHs can now be made explicit as follows. The Fock-exchange hole $n_x^{\mathrm{Fo}}(\mathbf{r};|\mathbf{r}'-\mathbf{r}|)$ describes a system without screening [7, 94, 95]. In contrast, for constraint-based GGAs [89, 93], HSE [57, 76, 77], and AHBR [10, 13], one keeps XC contributions defined by long-wavelength plasmons together [50] and suppress large-$|\mathbf{r}-\mathbf{r}'|$ contributions to the model exchange hole [76, 93]. In effect, these hole models have full asymptotic screening and we can introduce an effective dielectric constant $\varepsilon(\mathbf{r};|\mathbf{r}-\mathbf{r}'|)$ for counting exchange contributions in GGAs and in the vdW-DFs by setting $\varepsilon^{-1}(\mathbf{r};|\mathbf{r}-\mathbf{r}'|) = \tilde{n}_x^{\gamma}(\mathbf{r};|\mathbf{r}-\mathbf{r}'|)/n_x^{\mathrm{Fo}}(\mathbf{r};|\mathbf{r}-\mathbf{r}'|)$. This leads to an electron-exchange-hole coupling [49, 96],

$$E_x = \frac{1}{2}\int_{\mathbf{r}}\int_{\mathbf{r}'} n(\mathbf{r})\, V_{x,\mathrm{scr}}^{\gamma}\left(\mathbf{r},\mathbf{r}'\right) n_x^{\mathrm{Fo}}\left(\mathbf{r};|\mathbf{r}-\mathbf{r}'|\right)\,, \tag{2}$$

that is given in terms of a screened interaction $V_{x,\mathrm{scr}}^{\gamma}(\mathbf{r},\mathbf{r}') \equiv V_{\mathbf{r},\mathbf{r}'}\varepsilon^{-1}(\mathbf{r};|\mathbf{r}-\mathbf{r}'|)$. Again, γ controls a roll-over to an asymptotic behavior. The nature of PBE, HSE, CX, B86R, and AHBR is reflected by using $\varepsilon^{-1}\to 0$, i.e. using $\varepsilon^{-1}(\mathbf{r};|\mathbf{r}-\mathbf{r}'|\to\infty) = 0$ in $V_{x,\mathrm{scr}}^{\gamma}$ and in equation (2). A long-range-corrected (LRC) hybrid like VV10-LRC [97] uses $\varepsilon^{-1}\to 1$ while global hybrids [6, 94, 95, 98–101] use $\varepsilon^{-1} = \alpha < 1$, that is, set by a Fock-exchange mixing α that is independent of $|\mathbf{r}-\mathbf{r}'|$.

In search for highly accurate DFAs that exclusively address molecular systems, it is appealing to enforce the $\varepsilon^{-1} \to 1$ asymptotic limit on $V^{\gamma}_{x,\mathrm{scr}}$ when describing exchange contributions [22, 61, 68, 90, 91, 97] for molecules. This means moving DFT to the equally valid Hartree–Fock (HF)-KS scheme [46], where the relevant XC functional contains nothing but correlation. With that move the fundamental and HF-KS-DFT gaps are expected to align [21]. We note that exact (near-exact) DFT descriptions of the HOMO (LUMO) states, for a finite system like a molecule, are characterized by an asymptotic $|\mathbf{r}-\mathbf{r'}|^{-1}$ interaction form in equation (2) [61, 90, 91]. This form is also consistent with the observation that exchange must correct for errors that the Hartree term causes when one electron is located far from the set of nuclei [21].

Still, a direct implementation of this HF-KS scheme within the RSH framework of the AHBR approximation [5, 13] is a challenge; such a direct path implies trusting the existing vdW-DF2 correlation [1, 2] to reliably reflect all screening across all length scales and including that which we presently handle in $V^{\gamma}_{x,\mathrm{scr}}(\mathbf{r},\mathbf{r'}) \equiv V_{\mathbf{r},\mathbf{r}'}\varepsilon^{-1}(\mathbf{r};|\mathbf{r}-\mathbf{r'}|)$ and equation (2) [49, 50]. On the one hand, staying with $\varepsilon^{-1} \to 0$ asymptotic screening behavior (as in AHBR) is sometimes considered important [16, 18, 102, 103] to ameliorate impact of static correlation [19, 20, 104]. On the other hand, the nonlocal-correlation part of the vdW-DF method is robust, respects all plasmon constraints [1], and has hole conservation to all orders in expansions of its implicit screening logic [5, 105, 106]. It simultaneously avoids an excessive inclusion of correlation from low-density regions [3, 5, 25, 107–109] and therefore gives a plausible description of correlation contributions arising at asymptotic separations between the electron and its associated (XC) hole. It might work, for molecule-type problems, when matched with an asymptotically unscreened exchange.

In this perspective on RSH vdW-DF use for molecules, we therefore ask and answer two questions: (1) Does the strong AHBR performance [13] survive if we work with a new XC-functional class, denoted 'AHBR-mRSH(γ)', that avoids asymptotic screening in the exchange description *and* (2) Can we thus simultaneously obtain accurate molecular-energy and molecular-quasiparticle descriptions with a consistent inclusion of vdW forces?

Keeping the extent of short-range Fock-exchange mixing fixed at $\alpha = 0.25$, we find that AHBR-mRSH(γ) retains accuracy (in almost all cases) across a broad range of γ values, as tested on the full GMTKN55 benchmark suite of molecular properties [110]. We therefore propose to use a generic version (defined by $\gamma = 0.106$), denoted AHBR-mRSH, as a robust and versatile XC functional for DFT molecular-energy characterizations *and* we suggest use of the OT form AHBR-mRSH* as a consistent, vdW-inclusive predictor of molecule quasiparticles. We illustrate and test such use for frontier-orbital quasiparticles of nucleobases and Watson-Crick (WC) pairs.

The rest of the paper is organized as follows. Sections 2 and 3 contain theory and computational details. Section 4 contains our results and discussion while we present a summary and outlook in section 5. The paper also has a supplementary-information (SI) document [111] that helps detail that the system-specific form AHBR-mRSH* can be expected generally remain accurate also on molecular energy differences.

2. Theory

Using the gKS formulation of QE's hybrid implementation [72–74], we can circumvent the impact of the derivative-discontinuity that exists in the traditional KS solution scheme to DFT [21, 31, 60]. With the exact KS scheme [45, 90] (for a system with N electrons), we can solve for the orbital structure and energies $\epsilon_H(N)$ and $\epsilon_L(N)$ of the HOMO [61] and LUMO [84, 91] levels. We can also

obtain the HOMO energy $\epsilon_H(N+1)$ of the '$N+1$' system. In exact KS-DFT, minus $\epsilon_H(N+1)$ aligns with the exact-DFT description of the electron affinity (EA) $A_{\rm DFT} = E_{\rm DFT}(N) - E_{\rm DFT}(N+1)$ while minus $\epsilon_H(N)$ aligns with the ionization potential (IP) $I_{\rm DFT} = E_{\rm DFT}(N-1) - E_{\rm DFT}(N)$ [31, 112]. However, the fundamental gap $G_{\rm DFT} \equiv I_{\rm DFT} - A_{\rm DFT} = \epsilon_L^{\rm QP} - \epsilon_H^{\rm QP}$ exceeds the exact KS-DFT gap $g_{\rm KS} = \epsilon_L(N) - \epsilon_H(N)$ by the derivative-discontinuity $\Delta_{xc}^{\rm KS} = G_{\rm DFT} - g_{\rm KS} = \epsilon_H(N+1) - \epsilon_L(N)$. The difference arises because there must an abrupt change in the local KS-potential description of exchange exactly as the electron number $N+\delta N$ passes through $\delta N = 0$ [21, 60, 113–118]. This derivative discontinuity complicates direct use of KS-DFT for discussions of the quasiparticle nature. *Fortunately,* the gKS implementations of RSHs have instead an orbital focus and a simpler connection to quasiparticle levels: unlike for a local potential, we expect no singular orbital changes by infinitesimal δN changes [21, 35, 67].

There are many papers on understanding and on navigating the derivative discontinuity problems [60] that exist both when using DFAs and for exact KS-DFT [17, 21, 31–36, 39–41, 48, 62, 63, 66, 67, 69, 70, 82, 84, 113–139]. Our strategy for seeking a new vdW-inclusive predictor of molecular quasiparticles via a hybrid-type gKS vdW-DF mirrors that used in the PBE/HSE-based OT-RSH [32, 70].

We differentiate between three main gKS schemes [46]. They are equivalent ways to get exact-DFT total energies, $E_{\rm DFT}$, but their details impact the extent that we can use DFAs to simultaneously describe total energies $E_{\rm DFA}$, orbitals, and orbital levels $\epsilon_i^{\rm DFA}$. The gKS framework includes the original KS and the HF-KS [46, 82] schemes; the first (second) of these pursue orbital optimization as given by a single-particle Schrödinger equation (as given in the HF scheme while including a local potential that reflects all electron-electron interactions) [21, 46, 114, 135, 136]. The gKS framework for DFT/DFAs furthermore includes hybrid-type 'gKS' approaches, mixing the KS and the HF-KS gKS schemes [21, 32, 35, 46]. For DFA studies, we use that hybrid-type approach for launching a new type of OT-gKS studies with consistent inclusion of nonlocal-correlation effects.

We motivate focus on the AHBR-mRSH(γ) XC functionals as a path to molecular quasiparticles by the following five observations. First, when using the exact functional for the HF-KS solution scheme, the derivative discontinuity is expected to either vanish or be small (i.e. $\Delta_{xc}^{\text{HF-KS}} \approx 0$) and there is no discrepancy between $I_{\rm DFT}$ and $-\epsilon_H^{\text{HF-KS}}$ [36, 135]; unfortunately, we do not have the exact correlational functional for the HF-KS solution scheme [21]. Second, in exact KS-DFT studies we retain the $I_{\rm DFT} = -\epsilon_H^{\rm KS}$ alignment [61] but we loose a direct connection to other quasiparticle energies because $\Delta_{xc}^{\rm KS} > 0$ [21, 60, 135]; also, we do not have the KS potential corresponding to the exact XC functional (except when reconstructed for special atomic cases [61, 66, 91]). Third, DFA studies using GGAs or regular vdW-DFs (like CX and B86R that also have a GGA-type DFA for exchange) give good accounts of

$$I_{\rm DFA} \equiv E_{\rm DFA}\,(N-1) - E_{\rm DFA}\,(N)\,; \tag{3}$$

Such DFA studies are characterized by $\Delta_{xc}^{\rm KS} = 0$ values [35]. However, that does not mean that we can use the resulting KS levels as quasiparticle energies. This follows because the failure to reproduce the correct derivative discontinuity of the exact KS scheme causes mis-alignments of $-\epsilon_H^{\rm KS}$ and $I_{\rm DFA}$ values [140, 141]. Fourth, use of a general RSH (like HSE or AHBR) or global hybrid (like DF2-BR0) in gKS-DFA studies will ameliorate some of the above-summarized frustration in linking gKS levels and quasi particle energies; however, we must enforce $\varepsilon^{-1} \to 1$ for asymptotic exchange contributions in molecules [22] (implying a $|\mathbf{r}-\mathbf{r}'|^{-1}$ coupling of the electron and associated hole) to recoup the $-\epsilon_H^{\rm KS} = I_{\rm DFA}$ success demonstrated in [61]. The $|\mathbf{r}-\mathbf{r}'|^{-1}$ coupling and hence

asymptotic effective-potential form is also required for accuracy in describing unoccupied molecular orbitals [91].

Nevertheless, fifth, we can succeed at crafting a vdW-inclusive quasiparticle predictor via gKS-DFA calculations. This holds if we tune it to approach the exact HF-KS gKS-DFT, as far as possible. Taking inspiration from the OT-RSH, we can summarize the tuning in terms of the γ value that controls the roll-over of Fock-exchange mixing in the effective exchange modeling n^{γ}. A key necessary condition for a AHBR-mRSH(γ) gKS-DFA to mimic exact HF-KS DFT is compliance with the criterion on the HOMO energy position [32, 140]:

$$-\epsilon_{H}^{\mathrm{DFA},\gamma} = I_{\mathrm{DFA}}^{\gamma} \,. \tag{4}$$

To begin an explicit construction for a new broad class, denoted 'AHBR-gRSH', of RSH-vdW-DFs for gKS-DFA studies, we use access to an analytical exchange-hole modeling for B86R and hence for AHBR [10, 13]. As in [32, 68–70], we split the Coulomb matrix element

$$V_{\mathbf{r},\mathbf{r}'} = V\left(y = |\mathbf{r}-\mathbf{r}'|\right) \,, \tag{5}$$

$$V(y) = \frac{W_{\mathrm{FX},\alpha,\beta}\left(\gamma y\right)}{y} + \frac{W_{\mathrm{DF},\alpha,\beta}\left(\gamma y\right)}{y} \,, \tag{6}$$

using separation-dependent weightings,

$$W_{\mathrm{FX},\alpha,\beta}\left(\gamma y\right) = \alpha + \beta\,\mathrm{erf}\left(\gamma y\right) \,, \tag{7}$$

$$W_{\mathrm{DF},\alpha,\beta}\left(\gamma y\right) = 1 - \left[\alpha + \beta\,\mathrm{erf}\left(\gamma y\right)\right] \,. \tag{8}$$

The first (second) of these weightings is then used when counting contributions arising from the Fock-exchange hole n_x^{Fo} (from the B86R exchange-hole model [13] n_x). The resulting RSH approach implies using an effectively screened electron–hole interaction with a hybrid exchange-hole modeling $V_{x,\mathrm{scr}}^{\gamma}(\mathbf{r},\mathbf{r}') \equiv V_{\mathbf{r},\mathbf{r}'}\varepsilon^{-1}(\mathbf{r};y=|\mathbf{r}-\mathbf{r}'|)$ given as a range-dependent dielectric constant

$$\epsilon^{-1}\left(\mathbf{r};y\right) = W_{\mathrm{FX},\alpha,\beta}\left(\gamma y\right) + W_{\mathrm{DF},\alpha,\beta}\left(\gamma y\right)\frac{n_x\left(\mathbf{r};y\right)}{n_x^{\mathrm{Fo}}\left(\mathbf{r};y\right)} \,, \tag{9}$$

in equation (2). The crossover in the screening nature, at electron–hole separation $y \sim \gamma^{-1}$, reflects a RSH exchange-hole modeling

$$\begin{aligned} n_x^{\gamma}\left(\mathbf{r};y\right) = {} & W_{\mathrm{FX},\alpha,\beta}\left(\gamma y\right) n_x^{\mathrm{Fo}}\left(\mathbf{r};y\right) \\ & + W_{\mathrm{DF},\alpha,\beta}\left(\gamma y\right) n_x\left(\mathbf{r};y\right) \,, \end{aligned} \tag{10}$$

that is more general than the model used in setting up the AHBR [10, 13, 57, 76, 77]. Specifically, to land at a parameter-free form of the AHBR-gRSH class, for a given type of system, we must now set a total of three parameters (α, β, and γ). This can be done, for example, by looking at predictions for physical properties [36, 68–70, 99, 100, 142, 143] and demanding internal consistency or by using a coupling-scaling analysis to assert plausible values for the expected hybrid mixing [7, 94, 95, 144, 145]. We note that the error-function erf(y) switches between zero (at $y=0$) and unity (at $y\to\infty$) so that α determines the extent that we include short-range Fock-exchange. Also, for a finite γ, it follows that $\varepsilon^{-1} \equiv \alpha+\beta$ sets the extent of Fock-exchange at large separations and therefore the asymptotic screening.

We define our molecule-relevant class of gKS DFA 'AHBR-mRSH(γ)' subclass simply as 'AHBR-gRSH($\alpha=0.25,\beta=0.75$)', i.e. when enforcing an absences of asymptotic screening, $\varepsilon^{-1}=\alpha+\beta=1$. For a simplified discussion, we suppress expressing the γ variation when discussing the variation arising in the broader functional class 'AHBR-gRSH(α,β)'. We clarify that the

'AHBR-gRSH(α,β)' class contains both B86R and a traditional definition of a LRC form, denoted B86R-LRC. This B86R-LRC has a cross over (defined by using the fixed, LRC-standard value $\gamma = 0.4$) between a total absence of Fock exchange and (as in the AHBR-mRSHs) an assumed absence of asymptotic screening of exchange.

The top panel of figure 1 provides further details on the AHBR-gRSH(α,β) class of gKS DFAs. It highlights that the existing AHBR (black dot) and the new AHBR-mRSH (gold dot) emerge as the left and right ends of $\alpha = 0.25$ (dashed line). We also note that this line furthermore runs through DF2-BR0 (olive dot at $\beta = 0$) while the B86R (green triangle) and B86R-LRC (light-green square) instead exist on the $\alpha = 0$ line. The figure makes clear that AHBR-mRSH and B86R-LRC (B86R and AHBR) are the two options that comply with a molecule-consistent (metal-consistent) description, being situated on the gold line $\alpha + \beta = 1$ (black line $\alpha + \beta = 0$). However, it does not highlight the observation that DF2-BR0 also emerges as the AHBR-mRSH($\gamma = 0$) value (because the choice of β looses significance at $\gamma \to 0$).

As for the practical identification of the OT AHBR-mRSH* version, we seek the γ that means that AHBR-mRSH(γ) has compliance with equation (4), as far as possible also subject to additional conditions, below. We note that because the process seeks alignment with the exact HF-KS DFT, we are set up for quasiparticle predictions. We shall illustrate completing an OT of γ to define a system-specific version 'AHBR-mRSH*' and we have implemented AHBR-mRSH(γ) in an in-house QE version. We can therefore extract the quasiparticle predictions directly as in any QE DFT-RSH vdW-DF calculation, see [10, 11, 13, 16]. Practical use will always be limited to frontier levels: deeper quasiparticles exhibit a more complex structure [66, 146].

We note that we could perform the OT process for AHBR-mRSH* with attention to γ-dependent gKS-DFA results for the pair of IP values that correspond to the neutral and charged systems [32, 68]. However, by limiting the OT-process to just the neutral IP value of one member of a given system type, (e.g. nucleobases) we are in a position to validate the AHBR-mRSH* directly by characterizing the rest of the same-system frontier orbitals (before proceeding to predict those of the rest of related systems, e.g. other nucleobases and WC pairs).

We also note that we can see the resulting AHBR-mRSH functional as an effectively parameter-free description. The molecular nature sets the value of $\alpha + \beta$, while γ is tuned by a self-consistency criterion on the HOMO-level description. Presently we keep $\alpha = 0.25$ fixed, but we also observe that a coupling-constant analysis of the vdW-DFs can be used to establish a plausible choice also for the short-range Fock-exchange weight α, following the ideas of [7, 95, 98, 144].

Furthermore, we observe that even though the AHBR-mRSH(γ) is generally found to exhibit limited sensitivity towards tuning γ to values $\gamma \leqslant 0.2$, we are weary of permitting large-γ values for use in an OT AHBR-mRSH*, even if suggested by full compliance with equation (4). This reservation applies when considering vdW-dominated problems. We have defined a general-purpose CX version of vdW-DF by stressing balance of exchange and correlation, and a similar argument can be made also for the B86R (and hence for AHBR). When we use a $\gamma = 0.166$ (inverse Bohr [147]) value in an AHBR-mRSH* we are placing the cross over to the unscreened asymptotic Fock-exchange behavior at 6 Bohr, corresponding to about 3.1 Å. This is close to layer binding distances for graphite and vdW heterostructures [148]. The 3.1 Å length scale is still beyond the extent of the regions that dominate in setting the major nonlocal-correlation binding contributions [5]. If we further increase γ (for example, by using the $\gamma = 0.4$ value of B86R-LRC), however, we could be rubbing the XC balance for the vdW inclusion. More analysis is necessary to detail implications and, in this

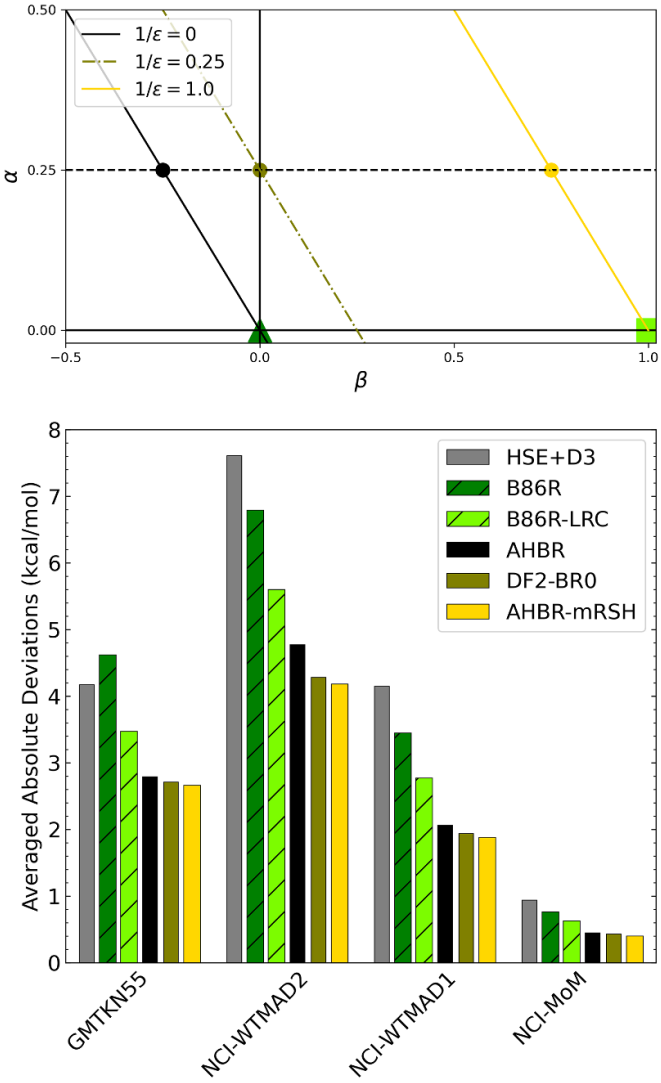

Figure 1. Nature, mutual relation, and Fock-exchange mixing details for molecular-relevant RSH vdW-DFs, top panel *and* corresponding functional performance on broad molecular properties [110], bottom panel. The RSH vdW-DF nature is summarized by the α-β position, where the former (latter) controls the extent of short-range exchange inclusion (enters in an effective dielectric constant, $\varepsilon^{-1} = \alpha + \beta$ that in turn controls the extent of long-range exchange inclusion, see text). Using the exchange-hole model [13] for the regular vdW-DF 'B86R' (nature and performance identified by green triangle and bar) we can define the fully-asymptotically screened AHBR [13] (identified by a black dot in the top panel and black performance bars in the bottom panel), the new molecule-relevant AHBR-mRSH (gold dot and performance bars), and a more traditional long-ranged corrected B86R-LRC (light green square and performance bars). The RSH-vdW-DF framework furthermore encompasses the global hybrid DF2-BR0 (olive dot and bar) and, for comparison, we repeat a performance assessment for HSE+D3 [13] (gray bars in bottom panel). The performance comparisons reflect mean absolute deviations from reference data using the WTMAD1 averaging measure for the entire GMTKN55 suite (as in [13]). For the subclass of all GMTKN55 benchmarks that reflect noncovalent interactions (NCIs), we report the assessment in the set of often-used standard measures, see text.

perspective on potential for wider RSH-vdW-DF use, we limit the range of possible quasiparticle predictors to tempered AHBR-mRSH($\gamma \leqslant 0.166$) choices.

Finally, we turn to discuss (1) the nature of the quasiparticle approximations that we can hope to capture by using our OT gKS-DFA method AHBR-mRSH*, and (2) transition dipole moment descriptors that reveal differences between, for example, the set of DNA nucleobases and that can track changes that emerge with nucleobase assembly into, for example, WC pairs. We shall compute

AHBR-mRSH* results for such descriptors, i.e. wavefunction overlap matrix elements, to permit a simple qualitative discussion of trends in exciton formation and in radiative exciton decay.

The description of LUMO and LUMO+ quasiparticles in molecules can formally present size-consistency challenges for characterizations in AHBR-mRSH* as well as in other DFT-based approximations and in GW. The issue arises in the many molecular cases where some or perhaps all of these electron-addition quasiparticles are unbound. For example, in some of the four DNA nucleobases, even the LUMO level exists merely as resonances observed in modified photoemission experiments [149–155]; we refer to [34, 44] for an overview of experimental results on nucleobases and for G_0W_0(PBE) and Koopmans-DFT studies that provide theory for the set of frontier orbitals. When we characterize such molecule resonances in a model framework using periodic boundary conditions (in GW and in our gKS-DFA) there will be a concentration of the LUMO or LUMO+ at positions with minimal values in the potential well defined by the molecular structure. Periodic-cell calculations are not then providing a quantitative description of such quasiparticle resonance since the relative magnitude of the concentration depends on the choice of unit-cell size. However, periodic-cell calculations (for example, AHBR-mRSH*) still permit a qualitative discussion of the nature of the resonance trapping. There is inherent value in making theoretical characterization of, e.g. nucleobase resonances [34, 44, 156], and we launch AHBR-mRSH* to get a simple (gKS-DFA-based) path to approximate the description of such orbitals.

Optical probing of biomolecular systems (like DNA) often focus on photon absorption in exciton creation, a neutral excitation, and luminescence by ensuing exciton decay. Quasiparticle formation (for example, by photoemission) are instead charged excitations that cause more damage. The rate of exciton formation (and decay) is proportional to the absolute square of the matrix element $M^j_{i\to f}$ for dipole optical transitions between the initial ground state and final state j with the same number of electrons, at minimal photon energy; as identified by the superscript, we consider here transitions driven by or producing electrical radiation fields directed in a given cartesian direction x_j. The rate of exciton creation (and radiative decay) can be approximated by computing the dipole transition moment [157, 158]

$$M^j_{i\to f} \propto \langle \Psi_L | \hat{x}_j | \Psi_H \rangle = \int_{\mathbf{r}} x_j \Psi^*_L(\mathbf{r})\, \Psi_H(\mathbf{r})\,, \tag{11}$$

that is defined by the electron-type LUMO and hole-type HOMO quasiparticles. The approximation equation (11) holds to lowest order in the potential V_{e-h} that describes the electrostatic coupling between the excited electron and the exciton hole; higher-order corrections arise when this potential also changes (primarily, we expect) the electron component, denoted Ψ_e, of the exciton beyond the Ψ_L approximation. Meanwhile, our new AHBR-mRSH* aspires to align its gKS-DFA frontier-orbital description with corresponding quasiparticles. We therefore characterize the gKS-DFA HOMO and LUMO through the corresponding wavefunction-overlap measure

$$M^j_{\mathrm{L,H}} \equiv \langle \phi_L | \hat{x}_j | \phi_H \rangle = \int_{\mathbf{r}} x_i \phi^*_L(\mathbf{r})\, \phi_H(\mathbf{r})\,, \tag{12}$$

$$\bar{M}^2_{\mathrm{L,H}} \equiv \frac{1}{3} \sum_j |M^j_{\mathrm{L,H}}|^2\,. \tag{13}$$

The dipole-transition moment predictor, equations (12) and (13), works, for example, for descriptions of exciton creation rates when taking the orbitals from a

non-hybrid KS-DFA study of water molecules in various interaction environments [158]. For a quantitative discussion one must use a prefactor $S < 1$ to account for the fact that the underlying MBPT description [157] cannot be fully captured with integer-weight wavefunctions extracted from a Slater-determinant KS-DFA solution (as used in equation (12)) or even from the exact KS-DFT or exact gKS-DFT descriptions. It is not clear how much S would change between different types of biomolecules.

For our intended demonstrator focus on nucleobase and WC pair characterizations, we see reasons to also seek more refined discussions when characterizing the rate of exciton creation and radiative decay. For example, we expect the unbound or weakly trapped LUMOs to be susceptible to the electron–hole coupling V_{e-h} and we should, ideally, go to higher orders in a perturbative treatment. We delay a discussion of what is the relevant HOMO and LUMO states when seeking to model exciton transitions in WC pairs.

We can sometimes provide qualitative measures of the scope of V_{e-h} impact on Ψ_e and hence on the exciton transition rates as described in AHBR-mRSH*. We note that the LUMO and some LUMO+ quasiparticles of the nucleobases have a dipole-trapped nature [34, 44], i.e. they are sitting beyond the spatial region containing the nuclei. The dipole-trapped nature is in contrast to the molecule-localized nature of so-called valence-bound orbitals [34, 156] that are found in non-hybrid DFA characterizations and in equation-of-motion coupled-cluster (EoMCC) studies [124, 128, 139] that also implement atomic relaxations. In general, we characterize the latter form as 'valence trapped' as the more localized states may still be unbound, for example, as found for LUMO+1 states in Koopmans-DFT [34]. We shall show that the AHBR-mRSH* correctly reproduces the Koopmans-DFT result and experimental finding of having a dipole-trapped character in the LUMO states of the nucleobases [34]. However, such states will be pulled in by the electron–hole coupling V_{e-h} in the exciton. Accordingly, we find that the LUMO-based predictor, equations (12) and (13), should be supplemented for exciton descriptions.

In practice we proceed at the level of adjusting the nature of the quasiparticle orbital(s) entering the dipole-transition dipole description. A more advanced study of how electron–hole coupling changes the nature of the electron component of the exciton wavefunction is beyond the present scope of presenting possible simple modeling usage of AHBR-mRSH*. We shall instead exploit the fact that AHBR-mRSH* also finds valence-localized (empty) frontier orbitals for thymine and cytocine and therefore also for both of the WC pairs. We denote such AHBR-mRSH* orbitals as ϕ_{L+v} (as they are always found at levels beyond the LUMO) and proceed to compute and use adjusted overlap integrals

$$M^j_{L+v,H} \equiv \int_{\mathbf{r}} x_j \phi^*_{L+v}(\mathbf{r})\, \phi_H(\mathbf{r})\,. \tag{14}$$

$$\bar{M}^2_{L+v,H} \equiv \frac{1}{3}\sum_j |M^j_{L+v,H}|^2\,, \tag{15}$$

The resulting descriptor, equations (14) and (15), remains crude. When available, it does provide an AHBR-mRSH*-based estimate of the rate of exciton formation and radiative decay and can again be used to identify trends, for example, the extent that WC pairing changes the rate of radiative decay of excitons in nucleobases. Also, contrasting the pair of descriptors, $|\bar{M}_{\mathrm{L+v,H}}|^2$ and $|\bar{M}_{\mathrm{L,H}}|^2$, allows us to discuss the scope of impact of wavefunction contractions driven by V_{e-h}.

3. Computational details

The potential for using the AHBR-mRSH* for quasiparticle descriptions hinges on whether we retain the AHBR accuracy when using a molecule-type exchange description and while completing (optimal) tuning. We shall therefore track how the AHBR-mRSH(γ) functional fares as a predictor of molecular total energy differences at increasing γ. In effect, we extend the performance assessment of B86R, HSE+D3, DF2-BR0 and AHBR that we reported in [13]. That is, we rely on the reference data of the GMTKN55 suite of benchmarks on broad molecular properties [110] but we now cover also the γ variation as well as the also-new AHBR-LRC form. We keep consistency (e.g. in terms of pseudopotential choices and grid convergence) with [13] and need only discuss the following four computational details.

First, a small QE code adjustment is needed to enable the exchange description to have a general cross-over between different degrees of Fock-exchange inclusion in a general AHBR-gRSH and for AHBR-mRSH(γ) forms. For the AHBR version, the Fock-exchange mixing is

$$\begin{aligned} W_{\rm FX}^{\rm AHBR} &= W_{{\rm FX},\alpha=0.25,\beta=-\alpha}\left(\gamma y\right) , \\ &= 1-{\rm erf}\left(\gamma y\right) = {\rm erfc}\left(\gamma y\right) . \end{aligned} \tag{16}$$

That form is already coded in the QE subroutine for fast evaluations of the Fock-exchange contribution (used in both HSE and in AHBR). For our present focus, AHBR-mRSH(γ), we need instead

$$\begin{aligned} W_{\rm FX}^{\rm mRSH} &= W_{{\rm FX},\alpha=0.25,\beta=1-\alpha}\left(\gamma y\right) , \\ &= \alpha\,{\rm erfc}\left(\gamma y\right) + {\rm erf}\left(\gamma y\right) , \end{aligned} \tag{17}$$

and so we have made those changes in the `exx` subroutine of QE. We have also made corresponding re-weighting of contributions arising from the B86R exchange-hole model, as evaluated in the `ahsx` subroutine in QE, [10]. The changes mirror those that permit an adapted form of QE to perform OT-RSH studies for PBE/HSE-based quasiparticle explorations [69].

Second, the GMTKN55 benchmark suite concerns broad molecular properties and we need procedures to weight all of 55 different benchmarks sets when seeking a statistical gauge of deviations between computed results and reference data. For an individual benchmark 'b' and a given functional 'f' we compute (and in interesting cases report and discuss) mean absolute deviation (MAD) measures, denoted MAD_b^f. These MAD_b^f results constitute a lowest (and thus most sensitive) rung in our statistical testing of the accuracy performance of the set of AHBR-mRSH(γ) gKS-DFAs.

The GMTKN55 suite organizes 55 benchmarks into groups as follows: small-molecular properties (Group 1, 'Small'), large-molecule properties (Group 2, 'Large'), barrier-height problems (Group 3, 'Barrier'), as well as intermolecular and intramolecular NCI problems (Groups 4 and 5 denoted 'Inter-NCI' and 'Intra-NCI'). This grouping of benchmarks means that we can get access to more statistics on functional performance of related types of problems. However, the average of the characteristic energy difference differs between the individual benchmarks even within any such group. For a meaningful comparison of functional performance we must therefore also discuss options for, and implication of, defining a set of *weighted total* (WT) MAD measures [110] as well as using the simpler Mean of MADs (MoM) measure, sometimes called the Total-MAD value [10, 159]. This is true for these benchmark groups as well as for the combined-inter-/intra-molecular NCI group (here denoted just NCI) and for the full GMTKN55 suite.

The definitions of two main options, WTMAD1 and WTMAD2, are given in [13, 110]. The MoM approach ignores the variation of the average reference energies. The WTMADs instead implement a per-group statistical assessment so that a functional deviation, a given MAD_b^f value, is associated with a larger (smaller) weight if the average of the reference energy data in benchmark *b* is small (large). The logic of using one of the WTMADs is that a 1 kcal mol^{-1} MAD value is expected to be much less acceptable when we address problems in which reactions or properties are set by subtle energy differences [110]. Such 'small-energy problems' are typical of the systems in the group(s) of NCI benchmarks.

The bottom panel of figure 1 reports examples of the statistical performance assessments that we are using. The first set of bars in the panel compares the performance of the fixed-γ AHBR-mRSH version with those of HSE+D3, B86R, B86R-LRC, AHBR, and DF2-BR0 for the entire GMTKN55 suite. This comparison is done using the WTMAD1 measure that is also used in [13].

The remaining set of performance comparisons in the panel makes it clear that the choice of statistical measures is important in functional comparisons and that this dependence itself contains information. The sets of bars here characterize functional accuracy across the NCI benchmark group, using the sequence of WTMAD2, WTMAD1, and the MoM measures. The WTMAD2 measure is the most sensitive measure for the NCI group, i.e. gives the largest nominal values; this generally holds for all groups of problems having small energy differences. The observation that the fixed-γ AHBR-mRSH is a robust, highly accurate functional for molecules is eventually made below because it performs at least at the level of the AHBR and DF2-BR0, for all of the GMTKN55 benchmarks groups, and for all of the statistical measures, i.e. holds for WTMAD2, WTMAD1, and MoM.

Prioritizing the WTMAD2 (over the MoM) measure for our statistical survey means that we can more easily detect whether a particular type of NCI problem (collected in a specific benchmark) develops errors as we increase γ in the AHBR-mRSH(γ) subclass. However, a sensitivity towards γ tuning might also emerge first in one of the large-energy problems. In that case we must instead prioritize the MoM measure, to prevent the impact from being hidden in the WTMAD2 representation of statistical errors. Accordingly, we use both the WTMAD2 and the MoM measures so that we can also detect what type of (large or small) problems that are adversely affected by tuning to a large γ, e.g. in B86R-LRC. We note that we do not consider use of a given AHBR-mRSH* form as internally consistent if the implied γ tuning causes a clear loss of accuracy in total-energy descriptions of the same type of systems.

Third, some of the 55 benchmarks include either individual or an entire class of negative-ion/-radical calculations that challenge a direct DFT-based assessment. The challenges that arise with negatively charged systems go beyond the task of compensating for electrostatic energies [10, 13, 160]; for those negatively charged systems it is important to deal with the adverse effects of self-interaction errors [22] on the electronic-structure in the numerical convergence [161–163]. It is primarily the G21EA benchmark set on electron-affinity energies (in the 'Small' group) and the OH^- radical in the WATER27 benchmarks set (in the inter-NCI group) that challenge a direct DFT approach. Shukla *et al* [13] summarizes the electrostatic-embedding scheme (using the Environ module [161, 162] of QE) that allows a full GMTKN55 performance assessment.

Finally, to use the OT AHBR-mRSH(γ) to predict nucleobase optical transitions in practice, we need to access the computed quasiparticles. Given the formal construction, making the AHBR-mRSH* appear as a HF-KS DFT (as far as it goes), our use of QE makes this straightforward: the orbital energies and info on their spatial structure are in part standard QE output and in part available

through the set of QE post-processing tools [71, 74]. However, we rely on (further) approximations when predicting relevant nucleobase optical transitions, as discussed in section II, above.

4. Results and discussion

We include broad documentation of AHBR-mRSH(γ) accuracy and robustness on molecular total-energy differences as well as demonstrations of both its internal consistency and its accuracy on quasiparticle descriptions. Our demonstrator example primarily concerns predictions of nucleobase frontier orbitals and estimates of matrix elements for their optical transitions. However, in DNA, the nucleobases are first combined into WC pairs and then into so-called WC steps, i.e. a dimer formed by stacking two WC pairs in geometries resembling their organization in DNA. A systematic inclusion of vdW forces is essential to understand any of such nucleobase assembly. Since, furthermore we illustrate an impact of WC pair assembly on the optical transitions, we think it important to maintain mutual AHBR-mRSH and AHBR-mRSH* consistency when seeking a better understanding of the electronic structure in the nucleobase systems.

The arguments for robustness and consistency is primarily given in a statistical sense on broad molecular properties [110]. That is, we give clear weight to checking whether (1) the molecule-optimized (but fixed γ) AHBR-mRSH is almost always at least as accurate as the existing AHBR for molecules, and that (2) the γ-tuning in AHBR-mRSH(γ) almost always has limited impact on the AHBR-mRSH accuracy. Validation is a requirement for seeing AHBR-mRSH and the AHBH-mRSH* as variant functionals for DFT descriptions of molecular energies (and by implication of molecular structure). For nucleobases, we also test AHBR-mRSH* consistency on having piecewise linear variation of the energy with the electron count, i.e. on a thermodynamic criterion [60].

4.1. AHBR-mRSH as a robust predictor of molecular-energy differences

Our first objective is to test if the fixed-γ pick AHBR-mRSH = AHBR-mRSH(0.106) is, in fact, a general-purpose DFA for characterizations of molecular total-energy differences. Since there are cases where one must expect that the AHBR-mRSH(γ) class fails [16], the test and validation can only be made in a statistical sense. In effect, we inquire if AHBR-mRSH is almost always performing at or beyond the performance of the (dispersion corrected) B3LYP+D3 [164–166], better than what arises with the more traditional definition of the B86R-LRC, and at least as good as the already accurate AHBR.

The bottom panel of figure 1 shows a comparison of functional performances for the full GMTKN55 suite and for the (combined) 'NCI' benchmark groups. We repeat that in using the AHBR-mRSH, we are keeping the $\gamma = 0.106$ value deliberately fixed and that DF2-BR0 can be seen as an intermediary between AHBR and AHBR-mRSH, section 2. The full-GMTKN55 comparison shows that DF2-BR0 has, on a statistical level, a performance that is between that of AHBR and AHBR-mRSH. We also note that the RSH vdW-DFs perform better than HSE+D3 and B3LYP+D3 (not shown, see data in [13]). Comparing the AHBR and AHBR-mRSH results, we find that it is of some but not dramatic consequence to prioritize the molecular-type screening logic (using $\varepsilon^{-1} \to 1$). That is, for descriptions of total energy differences, the improvements are marginal at the broadest statistical level.

Figure 2 shows that more pronounced performance differences emerge when we make the comparisons instead at the level of the five different types of GMTKN55 benchmark groups. The left panel of figure 2 shows a so-called radar plot comparison, now using the WTMAD2, i.e. the weighted measure that is most

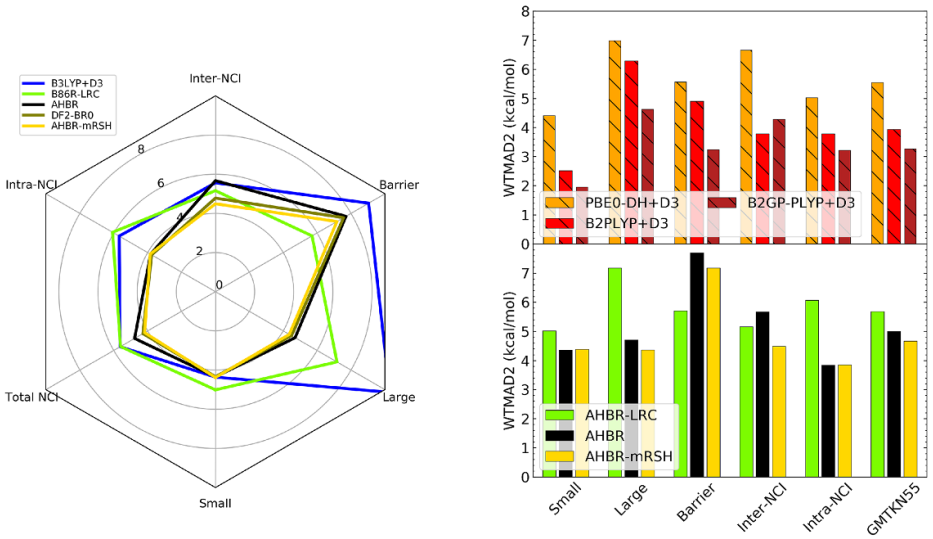

Figure 2. Left: radar plot of molecular-system performance (WTMAD2 [110] in kcal mol^{-1}), contrasting our three RSH vdW-DFs for general molecular descriptions (AHBR, AHBR-mRSH, and B86R-LRC) with dispersion corrected B3LYP+D3 [164–166] and DF2-BR0 [6, 13, 26]. Right: overview of performance, contrasting a non-empirical as well as two semi-empirical dispersion-corrected double hybrids [110, 167, 168] (top half) and the set of best RSH vdW-DF designs (bottom panel).

sensitive to errors made on benchmarks with small average energies in the reference data (i.e. to errors in the NCI class of problems). The left panel confirms that the performance of all the RSH vdW-DFs—and especially AHBR-mRSH and AHBR—are better than B3LYP+D3 in general. This is also true as for the challenging group of molecular barrier benchmarks, when, as here, we rely on the WTMAD2 performance measure; figure S.1 of the SI document shows that the performance of the RSH vdW-DFs matches that of the B3LYP+D3 when the functional performance is instead asserted in the WTMAD1 measure. Similarly, figure 2 shows that the B86R-LRC does slightly outperform the AHBR-mRSH when it comes to the Barrier group of benchmarks as asserted in WTMAD2, while figure S.1 shows that these functionals have the same Barrier-group performance as asserted in the WTMAD1 measure.

The right panel of figure 2 documents that the WTMAD2 performance of the best of the RSH vdW-DFs approaches that of (dispersion-corrected) double hybrids for molecule problems [110, 167, 168], at much less cost.

Figure S.2 of the SI document details differences in the performance of AHBR-mRSH and B86R-LRC for the Barrier benchmark group. It shows, in a radar plot comparing individual-benchmark MAD values, that the overall B86R-LRC accuracy edge (for Barriers), summarized in figure 2, cannot be seen as a general trend for transition-state problems. While the B86R-LRC is the more accurate choice for the BH76 benchmark, and to a lesser extent for the BHDIV10 benchmark, the AHBR-mRSH is the more accurate description for the INV24 and BHPERI benchmarks.

Table 1 reports individual-benchmark performance results (MAD values in kcal mol^{-1}), allowing comparisons for the set of here-investigated functionals. The benchmarks are selected from the GMTKN55 suite [110] because they are expected to be difficult [13], i.e. may represent generic challenges for AHBR-mRSH(γ) usage. The first three benchmarks reside in (and often limit functional performance on) the Barrier group, the next five plus four benchmarks are from the groups of 'Small' and 'Large'-system problems. The data in the last table section are benchmarks from the NCI groups [110]. Given this individual-benchmark focus, the statistics is poorer but the table does make it

Table 1. Contrast in performance between regular (B86R) and hybrid vdW-DFs (AHBR, DF2-BR0, AHBR-mRSH, B86R-LRC), and dispersion-corrected DFAs, based on a regular functional GGA (revPBE+D3), on the PBE-based RSH (HSE+D3), or on a popular empirical global hybrid (B3LYP+D3). The set of vdW-DFs can be seen as a here-introduced class 'AHBR-gRSH' of environment-sensitive RSH vdW-DFs based on the exchange-hole model for B86R. The benchmarks sets are chosen because they can be seen as sensitive to density-driven DFT errors [13, 22] or otherwise difficult [13, 110]. Boldface (italic) entries identify functionals within a group having leading (runner-up or competitive) performances, as asserted in terms of minimal MAD values (in kcal mol^{-1}) for each of these benchmark challenges. A broader comparison of the dependence of the AHBR-mRSH(γ) functional class is given in table S.I of the SI document.

	revPBE+D3[a]	B3LYP-D3[b]	HSE+D3[a]	B86R[a]	AHBR[a]	DF2-BR0[a]	AHBR-mRSH	B86R-LRC
BH76	7.38	5.70	4.21	9.22	4.14	4.05	*3.77*	**1.94**
PX13	9.29	*4.33*	7.38	11.36	*4.56*	*4.58*	4.66	*3.97*
BHPERI	5.74	**1.18**	2.83	6.08	*1.84*	*1.81*	*1.78*	3.28
W4-11	*5.88*	**3.40**	6.77	6.97	9.45	9.49	9.61	12.45
G21EA	2.83	**1.91**	3.40	4.50	**2.32**	*2.54*	3.34	6.02
SIE4x4	21.67	18.06	*13.58*	23.52	15.09	14.94	*14.49*	**10.12**
G2RC	5.60	**2.73**	6.48	5.58	**3.30**	**3.30**	*3.37*	*3.42*
DC13	9.38	10.14	8.24	7.26	**6.18**	**6.18**	*6.63*	9.75
MB16-43	25.28	24.84	**15.48**	22.37	*18.95*	*18.88*	*18.64*	35.00
RSE43	1.93	1.72	1.25	2.14	*0.74*	*0.72*	**0.66**	**0.42**
ISOL24	4.82	5.80	2.42	3.68	**1.82**	**1.59**	**1.65**	4.88
C60ISO	8.93	**2.22**	*2.51*	11.95	*2.72*	**2.15**	3.15	15.48
WATER27	2.63	4.07	5.73	5.10	*2.52*	*2.50*	*2.39*	**1.99**
PNICO23	0.84	0.48	0.86	0.56	**0.25**	**0.24**	**0.23**	*0.42*
HAL59	0.82	*0.57*	0.64	1.00	*0.58*	*0.57*	**0.53**	**0.50**
IDISP	3.25	3.57	*2.96*	2.69	*1.50*	**1.26**	**0.90**	3.63
Amino20x4	0.35	*0.21*	0.29	0.22	**0.19**	**0.18**	**0.19**	0.31

[a] [13].
[b] [110].

possible to discern if there are patterns in functional capabilities. For example, the first three entries exemplify the observation that B86R-LRC is not systematically better than AHBR-mRSH.

Figure 2 shows that AHBR-mRSH outperforms the B86R-LRC for the non-barrier benchmark groups. The strength of the fixed-γ AHBR-mRSH and the promise of the AHBR-mRSH(γ) class are exemplified by the individual per-benchmark MAD values reported in sections two to four of table 1. In the table rows, boldface and italic entries mark the best performer(s) and the cluster of second-best alternatives in our functional survey, respectively. We find that the B86R-LRC is only a clear performance leader for the BH76, SIE4x4, and WATER27 benchmarks, but here the AHBR-mRSH still represents a viable alternative. Meanwhile, one of the related AHBR, DF2-BR0, and new AHBR-mRSH gKS-DFAs are otherwise among the performance leaders.

At this point we can answer the first of our questions concerning use of AHBR-mRSH as an accurate DFA for predicting molecular energies. We find that moving to an absence of asymptotic screening for exchange (as in AHBR-mRSH) will not remove all residual functional error in the AHBR descriptions for molecules. However the AHBR accuracy [13] does indeed survive to the extend that the performance of the generic AHBR-mRSH (like AHBR) approaches that of the nonempirical (dispersion-corrected) double hybrids, left panel of figure 2. That is, at the statistical level(s), we find that AHBR-mRSH, like AHBR, is more accurate than B86R and B86R-LRC on (generic) molecular problems. Striving for consistency in electronic-structure predictions, we next ask if these AHBR-mRSH qualities survive γ tuning.

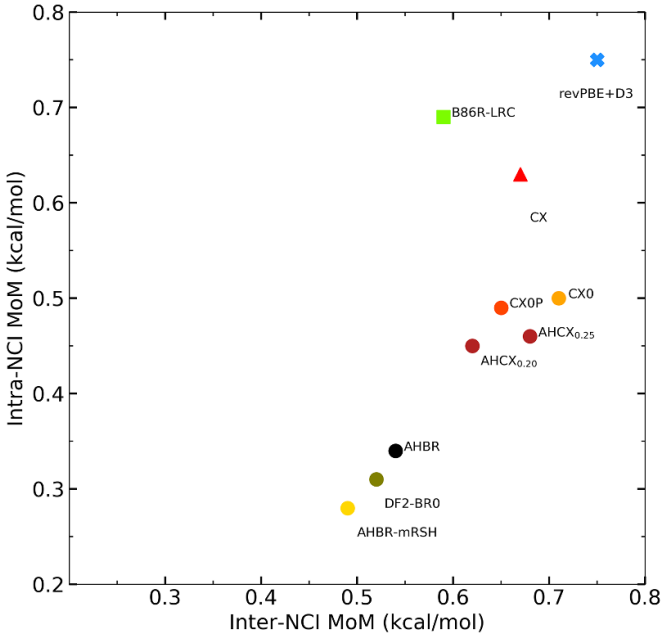

Figure 3. Performance as asserted for the two groups of noncovalent interactions (NCI) of the GMTKN55 suite [110], focusing on accuracy leaders; see the SI for full comparison. The coordinates for each hybrid vdW-DF entry (and those for each of CX, B86R, and revPBE-D3) reflect mean of mean average deviation 'MoM' measures [10, 13, 159].

4.2. Transferability of the molecular-energy description by AHBR-mRSH(γ) DFAs

Figure 3 highlights performance differences that arise for NCI problems. The figure shows that the change in screening (implied in going from AHBR to AHBR-mRSH) does not have a uniform impact on accuracy for molecular problems. It also suggests, to be explained, that there is an opportunity for completing an OT process, crafting a system-specific AHBR-mRSH* gKS DFA for (fast, vdW-inclusive) quasiparticle predictions.

Figure 3 reveals structure in the functional performance by the following analysis. There are two benchmark groups for the NCI class of problems, and for each we can extract the average performance measure MoM (also sometimes reported as T-MAD [10, 159]). The figure correlates performance variations in the set of best-performing vdW-DFs by assigning each functional a map position defined by its MoM performance on the inter- and intra-molecular NCI benchmark groups; a high-quality functional is identified by its proximity to the lower-left corner. Part of the data for the figure is taken from [13] and besides the here-introduced and -discussed functionals, the figure also includes CX0p and $AHCX_{0.20}$, i.e. versions of CX0 and AHCX where we set the (short-range) Fock-exchange inclusion at $\alpha = 0.20$. The figure makes it clear that the B86R-LRC is not ideal for NCI problems, whereas the AHBR/DF2-BR0/AHBR-mRSH set are the leading performers.

More importantly, figure 3 illustrates a key overall observation: adjusting the screening to reflect molecular-type conditions will be more beneficial for some types of the NCI problems than for others. Figure 3 provides a statistical-performance demonstration that moving to AHBR-mRSH (from DF2-BR0 and AHBR) helps more for intra-molecular NCI than for inter-molecular NCI problems. In practice, the partly selective AHBR-mRSH improvements arise because the IDISP benchmark is often the biggest Intra-NCI challenge (see also table 1) and in that type of problems, the screening logic of the AHBR-mRSH boosts the AHBR performance.

Figure 3 (and analysis of the IDISP case) is, at the same time, encouraging. We note that the AHBR-mRSH(γ) class does not contain the exact HF-KS-DFT: it merely comprises RSH vdW-DFs for gKS-DFA. Some variation of the performance with the screening nature of specific types of problems is therefore necessary if we are to craft a quasiparticle predictor by OT of the γ value in a gKS-DFA. The important question is whether the OT process for AHBR-mRSH(γ) can be seen as primarily impacting the energy-level part of the electronic-structure problem while, generally, causing only small changes in descriptions of molecular total-energy differences?

Table S.I of the SI extends the individual-benchmark performance comparison of table 1 regarding stability towards the OT process (γ tuning). The focus is still DFT challenges [13, 22, 110] but table S.I broadens the comparison. It includes data (from [13]) for global hybrid vdW-DFs CX0 and DF2-BR0, and for the regular-functional vdW-DF2 [2] and it reports MAD values for AHBR-mRSH(γ) in the sequence $\gamma = 0.106, 0.136, 0.166, 0.181$. Also, for ease in comparison, we retain the use of bold-face and italics entries as relative indicators of functional accuracy. Specifically, in table S.I, we highlight functional entries that are at least as good as the performance leader (or compares with the viable alternatives) identified in table 1.

Table S.I shows that we can often, but not always, complete an OT without causing significant loss in accuracy in the total-energy description. The table illustrates that the per-benchmark impression from table 1 (concerning AHBR-mRSH performance) generally survives as we increase γ but also that the C60ISO benchmark (on fullerene isomerization energies [110]) is an exception. Within the AHBR-mRSH(γ) DFA subclass, the performance for C60ISO is found to rapidly deteriorate when tuning $\gamma > 0.106$. A good and safe description of C_{60} isomerization energies requires instead use of one of the global hybrids (B3LYP+D3, CX0, or DF2-BR0). Accordingly, we can at most hope for robustness or transferability towards γ tuning in a statistical sense, e.g. averaged over benchmark groups.

Figure 4 reports our cross-GMTKN55 survey of the statistical performance of AHBR-mRSH(γ). Our analysis is here resolved into each of the five major GMTKN5 benchmark groups and we also show analysis for the full GMTKN55 benchmark suite. The top panel of the figure tracks WTMAD2 values (in kcal mol^{-1}) as we compute them by AHBR-mRSH(γ) studies at increasing γ values; the fixed-γ AHBR-mRSH form is reflected in the $\gamma = 0.106$ values. Meanwhile, the set of dashed lines reflects the DF2-BR0 performance [13], and we repeat that this global hybrid constitutes the $\gamma \to 0$ limit of the general AHBR-mRSH(γ) functional class.

We find for the NCI groups that the performance is not, in general, sensitive to the choice of γ, at this statistical-performance level. We also find that performance in this functional class actually improves for the barrier problems as the γ value increases. While the large-γ value of AHBR-mRSH(γ) will still differ from the B86R (concerning the inclusion of short-range Fock exchange) the AHBR-mRSH(γ) DFAs tend to recoup some of the B86R-LRC advantages for that group. The results of this survey suggest that moving the exchange description closer, overall, to pure Fock exchange, is helping on some barrier-height problem. However, a detailed discussion of how one might use such information to make DFT more versatile at succeeding at molecular-transition-state descriptions is beyond the present scope.

With an overall focus of potential for gKS usage at both quasiparticles and total energy variations, figure 4 suggests that it is often but, alas, not always possible to retain consistency. The combined WTMAD2- and MoM-type statistical analysis (shown in the top and bottom panels, respectively) clearly reveal that there are issues with the OT for the group of large-molecule benchmarks; these

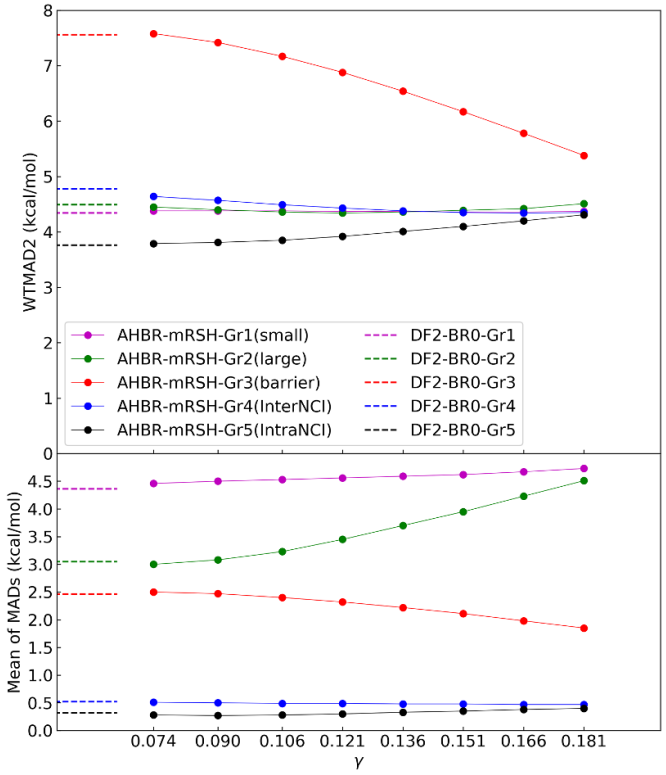

Figure 4. Dependence of averaged performance of the AHBR-mRSH-functional class on possible choices of the assumed inverse length scale γ, as asserted for the set of GMTKN55 benchmark groups. The set of long-dashed lines identifies the group-resolved performances of DF2-BR0 which corresponds to AHBR-mRSH($\gamma \to 0$). The set of solid lines are guides to the eye, connecting our set of full-GMTKN55 performance surveys (circles).

performance limitations are dominated by a γ dependence that exists for accuracy in the C60ISO benchmark. Departing from an assumption of perfect asymptotic screening in the exchange description (as in AHBR-mRSH*) impacts descriptions of both Cu_2O_2 complexes [16, 18, 20] and C_{60} isomerization [110].

Still, working with the generic DFT-energy predictor AHBR-mRSH and with the potential quasiparticle predictor AHBR-mRSH* can typically be seen as consistent, at least for smaller molecular systems. Whether that is relevant, however, hinges upon finding that the OT AHBR-mRSH* is, in fact, working as a good predictor of frontier quasiparticles orbitals in molecules, below.

4.3. AHBR-mRSH*: optimal tuning set by the thymine nucleotide ionization

Table 2 contrasts AHBR-mRSH(γ) results for the thymine IP value $I^{\gamma}_{\mathrm{DFA}}$ with those that are simultaneously reported in such QE studies for the HOMO energy position $-\epsilon_H$. The results are provided in studies where we placed the nucleobases in a cubic unit cell with a 40 Å; we shall return to a brief discussion of the the role played by size convergence. The bold-faced $\gamma = 0.166$ entry highlights that we impose a guard rail against uncontrolled changes in the XC balance, see section II.

We use the table 2 entries to implement our OT criterion, equation (4). The OT process sets an optimal $\gamma*$ value for use in an AHBR-mRSH* = AHBR-mRSH($\gamma*$) form that can describe quasiparticles in the nucleobases adenine, thymine, cytosine, and guanine, as well as in their WC pairs. The table caption describes the practical sides of the process. We find that a good choice for this specific system type is using $\gamma* = 0.166$ and setting AHBR-mRSH* = AHBR-mRSH($\gamma = 0.166$). This safe choice reflects that the mutual interactions of such systems are clearly impacted by vdW forces, as documented later.

Table 2. Summary of the OT process that defines an AHBR-mRSH* form for nucleobase characterizations. In practice, we study the thymine nucleobase in a large cubic unit cell of side length 40 Å to circumvent vacuum-level offsets in descriptions of the gKS orbital levels, see text. We next extract AHBR-mRSH(γ) results for the IP value $I^{\gamma}_{\rm DFA}$, for the HOMO level $\epsilon^{\gamma}_{\rm H}$, and for the HOMO-LUMO gap $g^{\gamma}_{\rm gKS} = \epsilon^{\gamma}_{\rm L} - \epsilon^{\gamma}_{\rm H}$, all listed in eVs. A best-possible $\gamma*$ value is then sought by optimizing compliance with equation (4). However, as indicated by the boldface $\gamma = 0.166$ value, we also impose a constraint on the γ tuning, as we are working within the vdW-DF framework, see text. The best possible pick is then this $\gamma* = 0.166$ value for thymine. The last two columns summarize a corresponding process for adenine, and show again a fair compliance with equation (4) by the thymine-guided choice $\gamma*$. We use this value and thus the OT AHBR-mRSH* = AHBR-mRSH($\gamma*$) in all characterizations of nucleobases systems.

	Thymine				Adenine	
γ	$I^{\gamma}_{\rm DFA}$	$-\epsilon^{\gamma}_{H}$	$I^{\gamma}_{\rm DFA} + \epsilon^{\gamma}_{H}$	$g^{\gamma}_{\rm gKS}$	$I^{\gamma}_{\rm DFA}$	$-\epsilon^{\gamma}_{H}$
0.106	9.149	8.432	0.72	7.974	8.417	7.449
0.121	9.164	8.573	0.59	8.229	8.432	7.884
0.136	9.181	8.706	0.47	8.466	8.449	8.012
0.151	9.198	8.833	0.37	8.685	8.466	8.131
0.166	9.217	8.952	**0.27**	8.856	8.485	8.243
0.181	9.236	9.064	0.17	8.978	8.503	8.348
0.196	9.255	9.171	0.08	9.092	8.522	8.446

In using equation (4) and setting the AHBR-mRSH* nature, above, it is relevant to discuss the small role that is played by converging QE planewave descriptions with use of an electrostatic decoupling [160] and with the inverse volume of the unit-cell chosen for the QE planewave studies we use. Planewaves are attractive for descriptions of the nucleobases because the LUMO levels are known to have the character of a so-called dipole-trapped state and these are not easily represented by a basis set that is tied to atomic positions. Still, having a planewave description also means that we rely on periodic repetitions of the molecules in their unit cells. At the system sizes we employ, there is safe electrostostatic decoupling in the molecule descriptions, see appendix of [13].

Nevertheless, a discussion of the 'vacuum' level (meaning the potential far from the nuclei) is important, given that we are using planewave DFT. In QE the 'vacuum' level is set as an average of the set of pseudopotentials. This fact does not impact the gKS-determination of the HOMO-LUMO gap in our molecular studies; we have explicitly verified that there are no changes in the computed gap values $g^{\gamma}_{\rm DFA} = \epsilon^{\gamma}_{L} - \epsilon^{\gamma}_{H}$ whether we use cubic cells of side length 30 Å or 40 Å. However, there is an impact on the apparent gKS-orbital energy positions because QE reports those orbital energies relative to the false vacuum level; for example, we find that for every γ value considered in table 2, that the HOMO energy levels ϵ^{γ}_{H} is systematically ~ 22.5 meV larger for thymine when obtained, instead, using a unit cell of 30 Å side lengths. More generally, the size impact on the apparent HOMO-level result decays slowly, with the inverse of unit-cell volume. Figure S.3 of the SI shows that the corresponding infinite-box HOMO energy levels differ by about 0.025 eV from the ϵ^{γ}_{H} entries listed in table 2.

The AHBR-mRSH* that we have now constructed has no external input. It does not rely on fitting, as all values are computed with the QE gKS description for each and every (closely related) functional that is represented in the rows of table 2. However, at this stage it is just another DFA for gKS-DFT studies of molecules.

Figure 5 suggests that use of AHBR-mRSH* as a quasiparticle predictor (as tested below) for nucleobase systems, will bring a clear consistency benefit: it remains true on the total-energy differences. The figure reflects DFA-specific performance measures (MAD values in kcal mol^{-1}) for two related WC-step benchmarks [169]; this data is listed in table S.2. Some of the performance

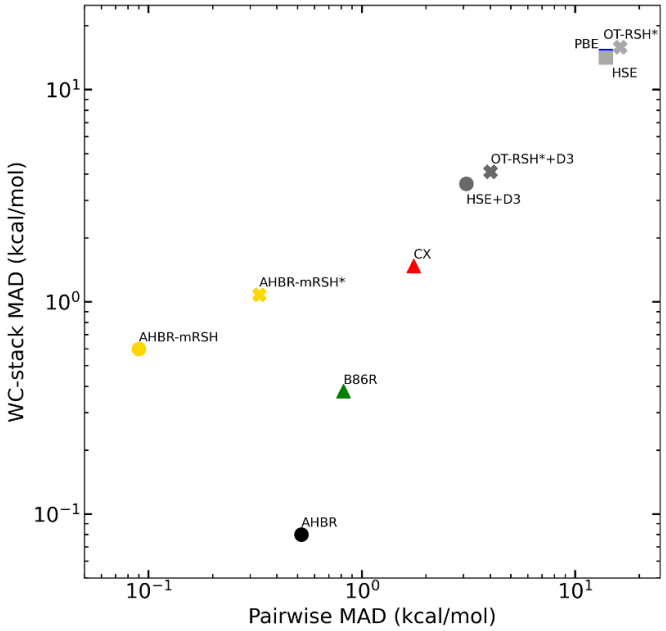

Figure 5. Accuracy on descriptions of the stacking of a DNA step of WC pairs. The assessments is given in terms of both a 'WC-stack' benchmark and the related (nucleobase) 'Pairwise' benchmark [13, 169]. The benchmarks consider all ten ways we can stack one of the two types of WC pairs on top of another WC pair. The geometries are also reference data, taken as plausible geometries that exist in DNA [169]. The nature of the benchmark are further summarized in the text. We contrast performance of the set of DFAs, as asserted in kcal mol^{-1} relative to the benchmark reference data [169]. Specifically, we report accuracy tests for both the new (fixed-$\gamma = 0.106$, general-purpose) molecular-total-energy predictor AHBR-mRSH, and the corresponding OT form, AHBR-mRSH*. For comparison we also include performance assessments of PBE, HSE, the dispersion-corrected HSE+D3, and for the PBE/HSE-based OT RSH (denoted OT-RSH* and OT-RSH+D3* depending on whether the D3 dispersion is included). The CX marker shows vdW-DF-cx performance.

numbers are repeated from [13] for ease of comparison. The figure uses double-logarithmic scales because switching from the semilocal-correlation HSE to RSH vdW-DFs leads to more than a factor-of-150 better accuracy on the WC-step benchmarks.

The figure considers accuracy of DFAs on descriptions of WC-step formation in DNA [13, 169–171]. In a model for such WC-pair steps, we need only consider two WC pairs in parallel configurations that are rotated approximately 36° relative to each other [170]. There are then ten different ways that we can stack such dimers of WC pairs as listed, for example, in table 1 of [169]. There and in further discussions below, the identification of the WC-pair step is given in a short-hand notation with 'XpY' where 'X' and 'Y' are each one of the nucleobases (we use A/T/C/G for adenine/thymine/cytosine/guanine) [169].

Figure 5 correlates the performance of the DFAs in two related benchmarks of the WC-step formation. Kruse *et al* [169] introduced these performance tests by providing reference data obtained by so-called DLPNO-CCSD(T) (complete-basis set) studies [169, 172] for all ten WC-step structures; DLPNO-CCSD(T) is also used for reference data in the GMTKN55 benchmark suite [110]. One of these benchmarks, here denoted 'WC-stack', concerns the energy gain that results when one WC pair is arranged on top of another, again in the plausible geometries [13, 169]. The second benchmark, here denoted 'Pairwise', considers instead alone the (larger) set of two-nucleobase interactions that can be seen as contributing to such WC-step energies. This focus on pairwise summation of nucleobase interactions is natural for descriptions based on dispersion corrected DFA, such HSE+D3. The vdW-DFs, including AHBR-mRSH*, provide instead a description that is based directly on the

electron-density variation and that treats all binding contributions within the same XC functional logic [3, 5, 13, 27, 106].

The figure shows that, even as just DFAs, the AHBR-mRSH and AHBR-mRSH* are both useful for characterizations of weakly interaction problems. The figure furthermore illustrates that by seeking the AHBR-mRSH* upgrade to HSE-based OT-RSH [32], we can gain considerably on consistency in total-energy and orbital characterizations for the ubiquitous class of weakly interacting molecular problems, e.g. for the nucleobase assembly in a DNA step of WC pairs.

We find that switching from the AHBR to the fixed-γ AHBR-mRSH form decreases (increases) accuracy on the WC-step (on the 'Pairwise') set, but all these MAD values stay within 1 kcal mol^{-1}. The same essentially also holds when we complete the OT process [using criterion equation (4) on thymine] to form AHBR-mRSH*. When we instead use the PBE/HSE-based gKS-DFA defined by that $\gamma*$ choice (forming the corresponding OTRSH*), the accuracy is poor for the nucleobase interaction energies.

We note in passing that one can add the semi-empirical D3 dispersion correction [166] to the PBE-bases functions and, for example, use HSE+D3 for improved-energy (and improved-structure) descriptions. In terms of quasiparticle predictions this suggests that one might also move between use of the PBE/HSE-based OT-RSH* to an OT-RSH-D3* form. The figure shows that these descriptions are still significantly less accurate on relevant total-energy differences for WC-step assembly [169]. The best (molecule-focused) gKS-DFA performers, as asserted for descriptions of energy differences, are the here-proposed non-empirical designs AHBR-mRSH and the AHBR-mRSH*. In short, AHBR-mRSH* has advantages if it actually works for quasiparticle predictions, next.

4.4. AHBR-mRSH on nucleobase quasiparticles*

Our (thymine-based) AHBR-mRSH* construction, above, faces two immediate questions in a proposed role as a new quasiparticle predictor: (1) Is it accurate also beyond the thymine HOMO level, i.e. for the frontier levels of all of the four nucleobases in DNA? and (2) Does it recoup piecewise linearity in its variation of its total-energy description with fractional changes in the electron count, $N_{\rm el} = N + \delta N$? The piecewise linearity is a property that the exact DFT must have [60] and it is the key reason that we pursued the OT process to get the AHBR-mRSH(γ) descriptions as close as possible to the exact HF-KS DFT description, section 2.

Table 3 answers the first question, on AHBR-mRSH* accuracy. The top four sections of the table report the frontier-level quasiparticle energy characterization that results when using our thymine-tuned AHBR-mRSH* as a predictor for thymine, adenine, cytosine, and guanine. There, and for the predictions of levels in the T-A and C-G WC pairs, we set the atomic configuration by the TpC-step reference data in [169]. All energy entries are listed in eVs and we compare AHBR-mRSH* results for the HOMO-1 (below abbreviated H-1), the HOMO, the LUMO, and the LUMO+1 (below abbreviated L+1) levels with G_0W_0(PBE) results [44], with experimental results (where available) [34, 42, 149–151, 154], and with results from PBE-based Koopmans-DFT [33, 113, 117, 118, 173] for the set of individual nucleobases [34]. The Koopmans-DFT results are obtained in the KIPZ algorithm [174], implemented in the KCP code extension [113, 173] to QE. These KIPZ results also reflect the atomic relaxations that result when an electron is added [34]. While lacking an account for vdW forces, this KIPZ-DFT study furthermore enforces both the piece-wise linearity constraint defined and discussed in [31, 60, 67] and a generalization of the Perdew–Zunger approach [17] to remove residual self-interaction effects [33, 174].

Table 3. Comparison of frontier-orbital energies (or 'levels') for individual bases, WC pairs, and WC-pair stacking into so-called WC steps, denoted 'ApA', 'CpC' and 'ApG'. The WC steps are formed by stacking two WC pairs in geometries that are considered relevant for understanding nucleobase assembly in DNA, [169], see also text and [13]. We take the atomic configuration as frozen and given as in the DLPNO-CCSD(T) [172] studies of mutual binding energies that Kruse *et al* used to define two benchmarks on nucleobase interactions [169]. Our AHBR-mRSH* characterizations (identified by a 'mRSH*' shorthand) are here done in a 30 Å-cubed cell. All energies are listed in eV.

		HOMO-1	HOMO	LUMO	LUMO+1
Thymine	mRSH*	−9.776	−8.910	−0.070	−0.008
	G_0W_0	−8.94[a]	−8.63[a]	0.26[a]	−0.24[a]
	Exper.	−10.14[b]	−9.2[b]	−0.06[c]	∼0.4[d]
	KIPZ	−9.77[e]	−9.02[e]	−0.06[e]	0.32[e]
Adenine	mRSH*	−9.236	−8.214	0.007	0.121
	G_0W_0	−8.80[a]	−7.99[a]	0.31[a]	0.25[a]
	Exper.	−9.45[b]	−8.47[b]	0.012[f]	∼0.5[d]
	KIPZ	−9.01[e]	−8.41[e]	0.02[e]	0.47[e]
Cytosine	mRSH*	−9.174	−8.565	−0.110	0.072
	G_0W_0	−8.5[a]	−8.18[a]	0.23[a]	0.02[a]
	Exper.	−9.55[b]	−8.89[b]	−0.23[d]	∼0.4[d]
	KIPZ	−9.12[e]	−8.70[e]	−0.11[e]	0.41[e]
Guanine	mRSH*	−9.350	−7.816	−0.165	0.050
	G_0W_0	−8.67[a]	−7.64[a]	—	0.43[a]
	Exper.	−9.90[g]	−8.30[g]	—	—
	KIPZ	−9.25[e]	−8.07[e]	−0.08[e]	0.36[e]
T-A WC-pair	mRSH*	−8.732	−8.052	0.007	0.106
C-G WC-pair	mRSH*	−9.165	−7.343	−0.108	−0.076
ApA step	mRSH*	−7.792	−7.766	−0.032	0.112
CpC step	mRSH*	−7.089	−6.834	−0.273	−0.089
ApG step	mRSH*	−7.573	−7.095	−0.109	0.082

[a] [44].
[b] [154].
[c] [150].
[d] Collected in [34, 42].
[e] [34].
[f] [151].
[g] [149].

We find that our new AHBR-mRSH* quasiparticle description is, overall, in qualitative agreements with G_0W_0(PBE) results, with experiments, and partly also with the Koopmans-DFT predictions, for all nucleobases. This holds even though we limited the OT at $\gamma \leqslant 0.166$ and seek a simple gKS-DFA framework. The experiments and KIPZ agree on finding a dipole-trapped nature (a valence-trapped nature) for the LUMO (for the LUMO+1) orbitals in nucleobases. The G_0W_0(PBE) study finds the opposite alignment between empty-state orbital nature and the energies and, following [34], we list the G_0W_0(PBE) empty-state level predictions as sorted according to the nature of the KIPZ/experimental orbitals. With AHBR-mRSH* we systematically find a dipole-trapped nature for the nucleobase LUMOs.

Figure 6 illustrates the qualitative success of our AHBR-mRSH* as a generic nucleotide quasiparticle level predictor and the options for an extraction of estimates for the energy positions of deeper lying quasiparticle levels. The set of crosses show experimental results for the LUMO+1 (abbreviated L+1), the LUMO, the HOMO and the HOMO-1 (abbreviated H-1) quasiparticle energy positions (where available). The set of dashed lines indicate either the IP value

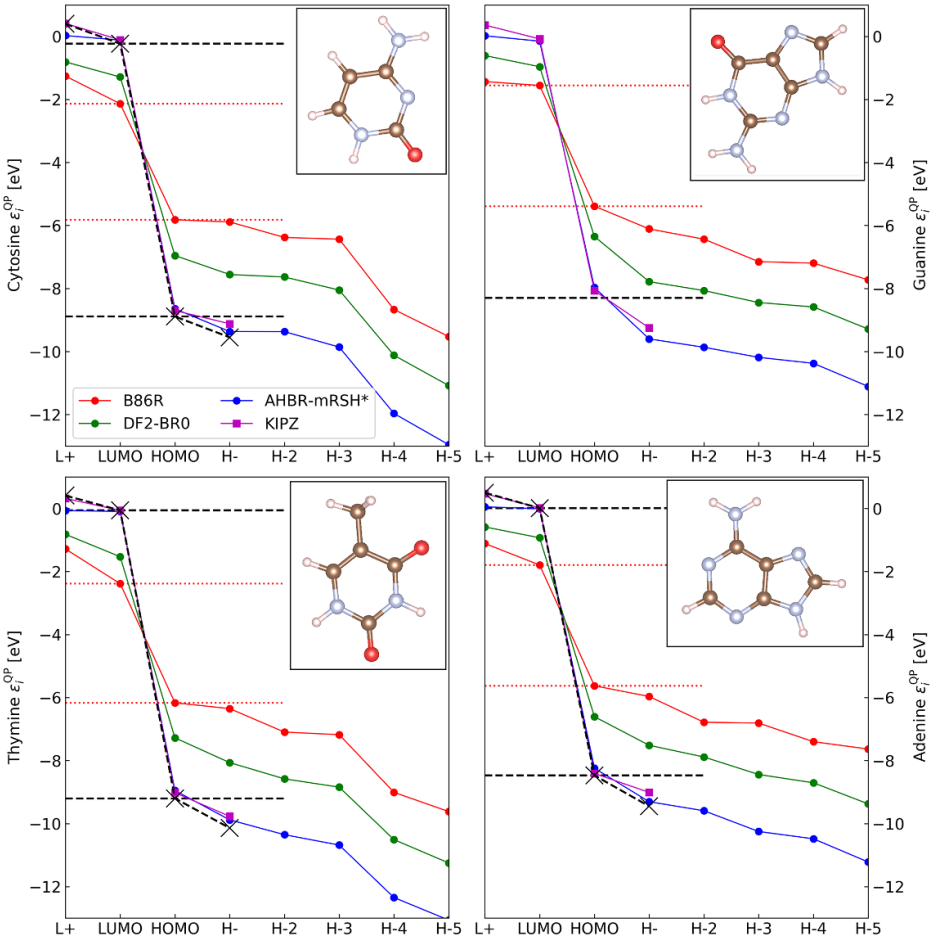

Figure 6. Comparisons of frontier-orbital energy level predictions in regular and hybrid vdW-DF functionals for cytosine and guanine (top row, left and right panels, respectively) and for thymine and adenine (bottom row, left and right panels, respectively). Tuning AHBR-mRSH to the computed thymine ionization potential (see text,) we compare the resulting AHBR-mRSH* description of quasiparticle energy levels with those obtained by B86R, by DF2-BR0, by the PBE-based KIPZ Koopmans-DFT descriptions [34, 173] as well as in experiments (where available, see table 3 and the overview provided in [34]). The pair of dotted lines identify the B86R KS-DFA description of the gap. For thymine, adenine, and cytosine, the pair of dashed lines identify the experimental values for the ionization potential and the electron affinity, i.e. the nature of the fundamental gap that a correct quasiparticle description can capture. For guanine we are only aware of data for the ionization-potential value we show.

$I^{\gamma*}_{\mathrm{DFA}}$ for guanine or the full experimental characterization of the fundamental gap (energy difference between the EA and IP) for the other nucleobases. There is an excellent agreement between measurements and the KIPZ results that also reflect atomic-relaxation effects [34]. Meanwhile, contrasting the results obtained by the B86R KS-DFT, by the global hybrid DF2-BR0, and by the AHBR-mRSH* gKS-DFT calculations (using B86R, DF2-BR0, and AHBR-mRSH*, we find that exclusively the latter has accuracy with regards to experiments that match that of a G_0W_0(PBE) study [44] and approach that of the KIPZ description. Motivated by this validation on the frontier-orbital energy levels, we also show our approximations for the energy positions of deeper levels. At the same time we note that deeper-lying quasiparticles cannot be fully described by an integer-weight gKS-DFT orbitals like those the AHBR-mRSH* gKS-DFA provide [83].

Figure 7 compares the B86R and AHBR-mRSH* descriptions of the thymine HOMO and LUMO/LUMO+ orbitals and identify the presence of qualitative changes among LUMO and LUMO+ orbitals for both thymine (top row) and for

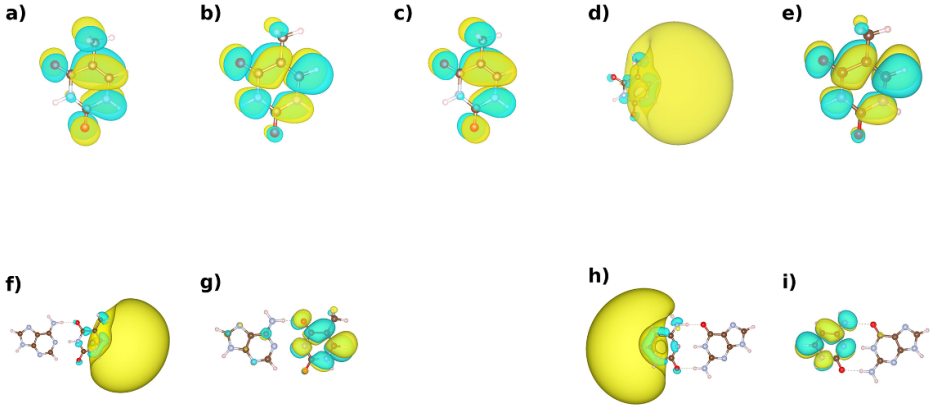

Figure 7. *Top row:* panels '(a)' through '(e)' contrast results for the frontier orbitals of thymine in a regular vdW-DF and in the new OT-RSH vdW-DF. Colors track changes in wavefunction signs. Panels '(c)' and '(d)' ('(a)' and '(b)') show contours of wavefunction-square variations, $|\phi_{\rm H}|^2$ and $|\phi_{\rm L}|^2$, as calculated for the HOMO and LUMO in AHBR-mRSH* (in B86R); a dipole-trapped nature of the thymine LUMO, panel '(d)', is corroborated by Koopmans-DFT results [34]. Panel '(e)' shows the AHBR-mRSH* result for the wavefunction-square variation of the LUMO+1 state; for general systems, we denote the lowest energy form of such 'valence-localized' (on-molecule) orbitals as $\phi_{\rm L+v}$, noting that the presence of both Φ_L and $\phi_{\rm L+v}$ orbitals is consistent with G_0W_0(PBE) results for the nucleobases [44]. *Bottom row:* panels '(f)' and '(g)' ('(h)' and '(i)') show that AHBR-mRSH also predicts a similar switch from a dipole-trapped LUMO nature to a valence-localized $\phi_{\rm L+v}$ behavior in the LUMO+2 state in the A-T WC pair (in the LUMO+1 state of the G-C WC pair). The HOMO levels (not shown) are localized on the opposite nucleobase in both of the WC pairs. The switch in the predictions of the LUMO+ nature impacts our approximations for optical matrix elements, $\bar{M}_{\rm L,H}$ and $\bar{M}_{\rm L+v,H}$, for potential exciton formation and decay, see text.

the two WC pairs (bottom row). All panels show the wavefunction-square variations, $|\phi(\mathbf{r})|^2$, using color to track wavefunction-sign changes. For the thymine HOMO level, shown in the top row panel '(a)' and '(c)', there are no discernible changes as we go from the density explicit B86R functional to the OT AHBR-mRSH* for thymine. However, AHBR-mRSH* predicts a LUMO nature (panel '(d)') that is not confined on the molecule, i.e. not having a shape like valence-bound B86R LUMO (panel '(b)'). Rather AHBR-mRSH* finds instead a LUMO that is a dipole trapped, i.e. located mostly outside the region containing the nuclei; the same applies also for the other nucleobases (adenine, cytosine, and guanine), not shown. These AHBR-mRSH* results are in qualitatively agreement with the nature of the LUMOs obtained in Koopmans-DFT [34].

Panel '(e)' highlights that AHBR-mRSH* also predicts a change from a dipole-trapped to a valence-localized nature for the (bound) LUMO+1 state of thymine (and for a higher-energy LUMO+3 state of cytosine, not shown); we label this type of states as 'L+v', in general. Koopmans-DFT (that also tracks the impact of atomic relaxations) finds that valence localization arises in LUMO+1 states of the nucleobases [34].

Panels '(f)' through '(i)' of figure 7 contrast the nature of the LUMO and 'L+v' orbitals for the two WC pairs, as predicted in AHBR-mRSH*. Here we find that the WC-pair LUMO (shown in panel '(f)') again emerges as a dipole-trapped state located outside the thymine in the A-T pair and outside the cytosine in the C-G pair (panel '(h)'). Also, AHBR-mRSH* predicts that more localized 'L+v' states exist as LUMO+1 (trapped by thymine, panel '(g)') and as LUMO+2 (trapped by cytosine, panel '(i)') for the T-A and C-G pairs, respectively. We note in passing that the HOMO levels of WC pairs (not shown) are located on the opposite nucleobase than the thymine and cytosine base that weakly (more completely) traps the LUMO ('L+v') states.

Table 4. Comparison of AHBR-mRSH* bandgap-estimates $g^{\gamma*}_{\mathrm{gKS}}$ and of (spherically averaged measures for) dipole-transition-moment elements, $\bar{M}_{\mathrm{L,H}}$ and $\bar{M}_{\mathrm{L+v,H}}$. These matrix elements underpin a simple modeling of exciton formation, see text. More broadly, these matrix elements quantify the extent of unoccupied-orbital changes that arise with a switching between a dipole-trapped LUMO state and the valence-localized $\Psi_{\mathrm{L+v}}$ nature that AHBR-mRSH* (correctly) finds for LUMO+ states in Thymine, Cytosine and for both of the WC pairs, see figure 7.

System	$g^{\gamma*}_{\mathrm{gKS}}$ (eV)	$\lvert\bar{M}_{\mathrm{L,H}}\rvert^2 \times 10^3$	$\lvert\bar{M}_{\mathrm{L+v,H}}\rvert^2$
Adenine	8.22	3.9	—
Thymine	8.46	0.9	2.4047
Cytosine	7.65	11	0.3870
Guanine	8.84	2.9	—
T-A WC pair	8.06	0.8	0.1050
C-G WC pair	7.24	0.003	0.0254

Overall, figure 7 suggests an inherent value in seeking a quasiparticle predictor, like AHBR-mRSH*, that treats nonlocal correlations directly within the XC functional description. We may set the atomic structure of a molecular system from experiments or by calculations using dispersion corrected DFAs and from there proceed to determine quasiparticles using a GGA, HSE, or the HSE-based OT-RSH. Such a process will reveal what we may call indirect signatures of nonlocal-correlation effects in the electronic structure [175]. However, figure 7 shows that the LUMO, as well as 'L+v' quasiparticles, have significant probability weight in the regions far from the nuclei, reflecting also that AHBR-mRSH* (correctly) describes these as either weakly bound or weakly trapped. That is, the LUMO and LUMO+ quasiparticle states and resonances exist in the very low-density regions where nonlocal-correlation XC contribution dominates, binding the molecular systems [5, 145, 176]. The AHBR-mRSH* goes beyond the HSE-based OT-RSH by keeping the description of vdW forces and NCIs inside the XC functional and by thus additionally capturing direct nonlocal-correlation signatures in the quasiparticles, i.e. in the molecule electronic structure.

Table 4 summarizes our quantification of the variation in wavefunction nature that we predict in AHBR-mRSH*. It does so by contrasting the pair of descriptors $|\bar{M}_{L,H}|^2$ and $|\bar{M}_{L+v,H}|^2$ of the rate of exciton creation that can be expected by photon excitation out of the HOMO level specific to the systems. The former descriptor highlights the dipole-trapped character of the LUMO, by yielding small values reflecting a limited wavefunction overlap. The latter measure is only available when we find a valence-located (higher-energy) empty orbital $\phi_{\mathrm{L+v}}$ by AHBR-mRSH*, i.e. for thymine, cytosine and for both of the WC pairs. Comparing the $|\bar{M}_{L,H}|^2$ and $|\bar{M}_{L+v,H}|^2$ descriptors we see clear differences of whether we consider HOMO overlaps with a dipole-trapped or a valence-localized LUMO or LUMO+. For the HOMO-to-'L+v' descriptor (when available) we find moderate values but the HOMO-to-LUMO descriptors are small.

In section 2 we formulated an expectation that the $|\bar{M}_{L+v,H}|^2$ descriptor is generally the more relevant qualitative indicator of the rate of exciton creation or decay (by luminescence). We expect that this holds for optical excitation out of the HOMO level of an individual nucleobase molecule, but we also note that a deeper discussion will be needed when considering a system of two or more nucleobases. For example, the AHBR-mRSH* results for the band gaps $g^{\gamma*}_{\mathrm{gKS}}$ (also reported in table 4) of WC pairs are smaller than those of the individual nucleobases that make up the pairs. This is possible because the locations of HOMO and LUMO levels are spatially separated, located on different nucleobases, as noted above. For WC pairs, a direct application of the $|\bar{M}_{L+v,H}|^2$

descriptor may not be relevant, at least not unless the photon energy is kept right at the quasiparticle gap of the WC pair. For example, we predict that cytosine has the largest rate of exciton creation within the four nucleobases. With a slightly larger photon energy, it may therefore be possible to make a (short-lived) intra-cytosine excitation in a C-G WC pair. Advanced Bethe-Salpeter equation [177–179], EoMCC [124, 128, 139, 156] or time-dependent DFT studies [180–186] are needed to understand the nature of such excitations.

We also note that the AHBR-mRSH* characterization of wavefunction overlaps, table 4, need not permit even a qualitative discussion of luminescence in nucleobases system. We are characterizing HOMO and LUMO states of the WC pairs, finding that they end up on opposite nucleobases of the pair. We find that, accordingly, the overlap measures $|\bar{M}_{L+\nu,H}|^2$ and $|\bar{M}_{L,H}|^2$ are significantly suppressed compared to what applies for an individual nucleobase. Upon a generic exciton creation, there is fast electron-electron scattering driving a thermalization of the electron and the hole and favoring landing the former (latter) in the LUMO (HOMO) of the combined system, e.g. in a WC pair. However, once there is an electron and hole in such states, there is a strong inter-molecule dipole that will drive also atomic relaxations [156, 181, 183, 184]. Again, more advanced characterizations are needed to understand whether the system still supports an electron and hole behavior in the form of an inter-molecule dipole by the time the exciton (may) give rise to luminescence.

Figure 8 answers the second question on AHBR-mRSH* usefulness: this OT vdW-inclusive DFA ensures a piecewise linearity in the description of the nucleobases. We find that AHBR-mRSH* (constructed for nucleobase studies) is consistent with the key thermodynamics constraint on DFT [60]. The figure reports the AHBR-mRSH* total-energy variation for thymine and guanine (top and bottom panels, respectively) that arises as we add or subtract a fraction of an electron δN. For ease of comparison we have set the total energy of the neutral system (in each case) to zero; that is, we report resulting AHBR-mRSH* total-energy changes, denoted $\delta E_{\mathrm{mRSH}*}$. The dashed lines in the panels identify the corresponding AHBR-mRSH* results for minus the HOMO energy.

The figure shows that the AHBR-mRSH* DFA gives us a very good but also not perfect alignment of the thymine and guanine IPs of these levels. This is expected (and evident also in table 3) because we limited the tuning to $\gamma \leqslant 0.166$ on other grounds, section 2.

Nevertheless and importantly: the figure demonstrates that AHBR-mRSH* complies with a piecewise-linearity criterion on its total-energy description. It does so to a very high and excellent degree for thymine and guanine, respectively. There are essentially no energy changes when adding a fraction of an electron; this is consistent with our finding that the LUMO levels are all pushed to the vacuum-energy position in the AHBR-mRSH* description of these nucleobases. There is also a discernible energy cost of removing a fraction δN of an electron from these bases and that cost is well approximated as δN times the IP value (and thus, approximately, $-\epsilon_H$). This is as it should be [21, 31–34, 36, 62–64, 66–68, 70, 82–84, 113–117, 121, 125–127, 130, 131, 133–136, 138, 140].

We note that piece-wise linearity is essential to enable compliance with the tuning criterion, equation (4) [140, 141]. However, we are not aware of arguments that this criterion itself, in turn, secures the piecewise linearity. In any case, we think the question concerns inherent functional quality because (1) the piecewise linearity is itself mandated by a thermodynamical criterion on charge transfer, as formulated in [60] and (2) documenting accuracy on predicting charge transfers (like δN) implies some robustness towards excessive localization or delocalization.

We finally observe that internal consistency of the RSH vdW-DFs is a key argument for suggesting broader use, e.g. via this perspective paper. Like G_0W_0

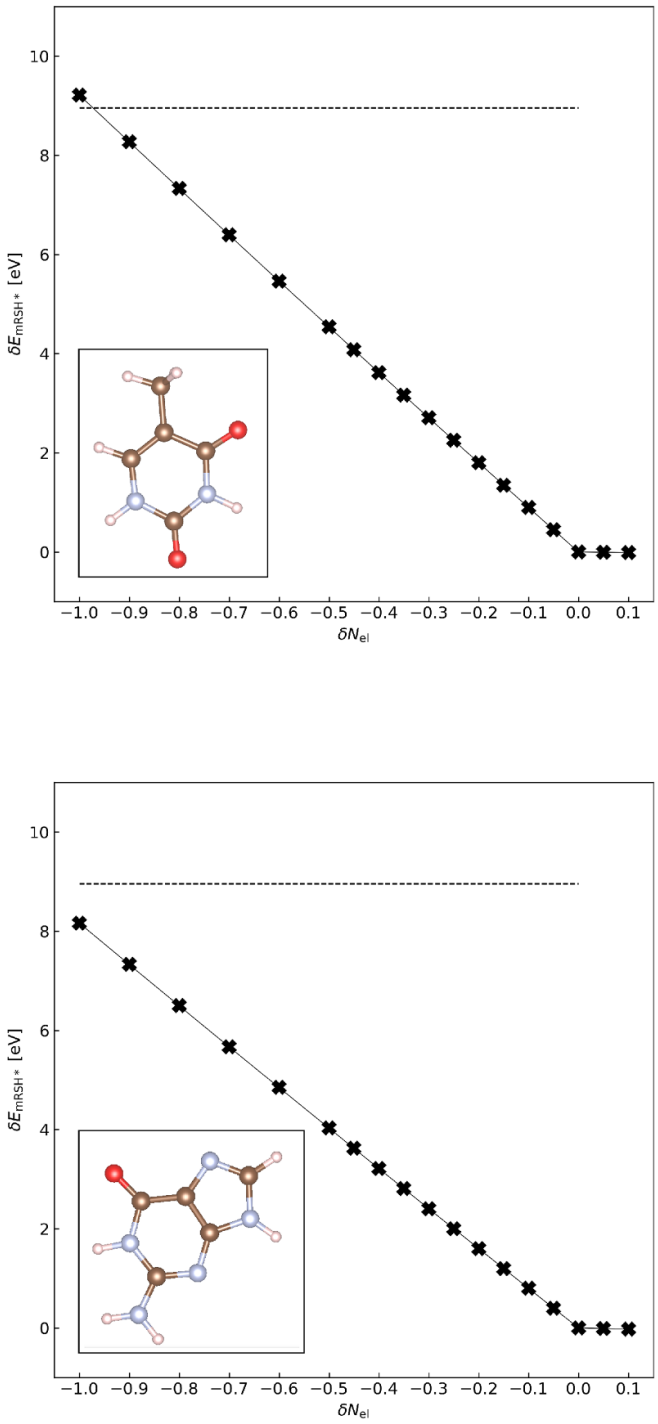

Figure 8. Piecewise linearity in the energy versus electron count for thymine (top panel) and guanine (bottom panel), for the AHBR-mRSH* that we craft to characterize nucleobase quasiparticles by tuning γ to a thymine ionization criterion.

and the random-phase approximation, our AHBR-mRSH* is defined to provide an internally consistent inclusion of correlation that is based directly on screening and other formal MBPT arguments [1, 3, 5, 49, 50, 52, 54, 59, 105–107, 187]. This internal consistency, in turn, motivates our expectation that we can take AHBR-mRSH* accuracy on molecular binding-energy descriptions, figures 4 and 5, as promise for use as an approximative predictor of quasiparticles in weakly interacting molecular systems.

5. Summary and outlook

We introduce a new RSH vdW-DF design, AHBR-mRSH*, for computing both energies and gkS-DFA approximation of frontier quasiparticle. We illustrate use by discussing optical transitions in nucleobase systems. Our approach for estimating quasiparticles is computationally fast, for example, compared to G_0W_0 studies. We find that we are able to use it to describe up to about one hundred atoms on a standard high-performance-computer cluster.

As part of the method documentation we address accuracy as a DFT-type total-energy. Specifically, we explore whether the here-defined AHBR-mRSH or the OT AHBR-mRSH* forms show accuracy loss or even gain compared the related, existing RSH-vdW-DF AHBR for molecular problem. We report the mixed answer: the improvements are not uniform but a small net effect survives when averaged over the GMTKN55 benchmark groups. Based on the analysis, we find that both AHBR and AHBR-mRSH are highly versatile functionals for general molecular problems.

The finding of small performance differences implies, in turn, that the AHBR has strength as a general purpose gKS-DFT, even if the AHBR-mRSH and AHBR-mRSH* also have advantages for small-molecule problems. That is, we conclude that the AHBR remains the best general-purpose XC functional: It works very well as a total-energy predictor across bulk, surface, and molecular problem. Accordingly, the AHBR is naturally set up to determine total-energy variations also in heterogeneous systems, for example, the energy of molecule adsorption at metallic surfaces. We must avoid the AHBR-mRSH and AHBR-mRSH* at metallic problems because the implied inclusion of long-range exchange (via the Fock exchange inclusion) produces a divergence in the density of states of the metal at the Fermi level.

Meanwhile, the finding that AHBR-mRSH and AHBR-mRSH* generally both retain the strong AHBR level of accuracy does mean that we can use them as mutually consistent gKS-DFA tools for predicting molecular total-energy differences and frontier-level quasiparticles, respectively. In fact, we find that we can often rely on the OT AHBR-mRSH* form to succeed at both tasks. These conclusions follow because the RSH vdW-DFs have the transferability to survive adverse effects of changing the inverse-length scale γ value inside the RSH vdW-DF description of exchange. In essence, we find that the here-defined OT AHBR-mRSH* form generalizes the PBE-based OT-RSH description with 1) an internally-consistent account of vdW interactions and 2) an ability to rely on the same XC functional logic for both the total-energy and quasiparticle part of the electronic-structure problem (near the frontier levels).

We end this perspective with the hope that the consistent inclusions of vdW and weak-interaction impacts on the electron-structure problem by the OT AHBR-mRSH* gKS-DFA can help in the study of quasiparticles in molecules and in interacting molecular systems.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

We thank Jeffrey Neaton and Zhenfei Liu for input on adapting code for the QE weighting of the Fock-exchange contributions to the form required for AHBR-mRSH and for discussions. We also thank Martin Rahm, Marcus Wilhelmsson, and Joakim Andreasson for discussions. We are grateful for support

by the Swedish Research Foundation, through Grants 2020-04997 and 2022-03277, the Swedish Foundation of Strategic research, through Grant IMF17-0324, the Sweden's Innovation Agency Vinnova, through Project No. 2020-05179, Chalmers Area-of-Advance (AoA) Production, and Chalmers AoA Nano. The authors also acknowledge computer allocations from the Swedish National Infrastructure for Computing (SNIC and NAISS) under Contracts SNIC2021-3-18, SNIC2022-3-16, NAISS2023-3-22, NAISS2023-6-306, NAISS2024-6-432, and NAISS2024-3-16 as well from the Chalmers Centre for Computing, Science and Engineering (C3SE).

ORCID iDs

Elsebeth Schröder https://orcid.org/0000-0003-4995-3585
Raul Quintero-Monsebaiz https://orcid.org/0000-0002-4518-5497
Yang Jiao https://orcid.org/0000-0003-4771-0088
Per Hyldgaard https://orcid.org/0000-0001-5810-8119

References

[1] Dion M, Rydberg H, Schröder E, Langreth D C and Lundqvist B I 2004 van der Waals density functional for general geometries *Phys. Rev. Lett.* **92** 246401
[2] Lee K, Murray È D, Kong L, Lundqvist B I and Langreth D C 2010 Higher-accuracy van der Waals density functional *Phys. Rev. B* **82** 081101(R)
[3] Berland K, Cooper V R, Lee K, Schröder E, Thonhauser T, Hyldgaard P and Lundqvist B I 2015 van der Waals forces in density functional theory: a review of the vdW-DF method *Rep. Prog. Phys.* **78** 066501
[4] Thonhauser T, Zuluaga S, Arter C A, Berland K, Schröder E and Hyldgaard P 2015 Spin signature of nonlocal correlation binding in metal-organic frameworks *Phys. Rev. Lett.* **115** 136402
[5] Hyldgaard P, Jiao Y and Shukla V 2020 Screening nature of the van der Waals density functional method: a review and analysis of the many-body physics foundation *J. Phys.: Condens. Matter* **32** 393001
[6] Berland K, Jiao Y, Lee J-H, Rangel T, Neaton J B and Hyldgaard P 2017 Assessment of two hybrid van der Waals density functionals for covalent and noncovalent binding of molecules *J. Chem. Phys.* **146** 234106
[7] Jiao Y, Schröder E and Hyldgaard P 2018 Extent of Fock-exchange mixing for a hybrid van der Waals density functional? *J. Chem. Phys.* **148** 194115
[8] Perrichon A, Granhed E J, Romanelli G, Piovano A, Lindman A, Hyldgaard P, Wahnström G and Karlsson M 2020 Unraveling the ground-state structure of $BaZrO_3$ by neutron scattering experiments and first-principles calculations *Chem. Mater.* **32** 2824
[9] Granhed E J, Wahnström G and Hyldgaard P 2020 $BaZrO_3$ stability under pressure: the role of non-local exchange and correlation *Phys. Rev. B* **101** 224105
[10] Shukla V, Jiao Y, Frostenson C M and Hyldgaard P 2022 vdW-DF-ahcx: a range-separated van der Waals density functional hybrid *J. Phys.: Condens. Matter* **34** 025902
[11] Frostenson C M, Granhed E J, Shukla V, Olsson P A T, Schröder E and Hyldgaard P 2022 Hard and soft materials: putting consistent van der Waals density functionals to work *Electron. Struct.* **4** 014001
[12] Lee J-H, Hyldgaard P and Neaton J B 2022 An assessment of density functionals for predicting CO_2 adsorption in diamine-functionalized metal-organic frameworks *J. Chem. Phys.* **156** 154113
[13] Shukla V, Jiao Y, Lee J-H, Schröder E, Neaton J B and Hyldgaard P 2022 Accurate nonempirical range-separated hybrid van der Waals density functional for complex molecular problems, solids and surfaces *Phys. Rev. X* **12** 041003
[14] Abidin A F Z and Hamada I 2022 Interaction of water with nitrogen-doped graphene *Phys. Rev. B* **105** 075416
[15] Seyedraoufi S and Berland K 2022 Improved proton-transfer barriers with van der Waals density functionals: role or repulsive non-local correlation *J. Chem. Phys.* **156** 244106
[16] Frostenson C, Feng Y, Hyldgaard P and Grönbeck H 2024 Range-separated hybrid van der Waals density functionals to describe Cu_2O_2-complexes *Chem. Phys. Lett.* **856** 4009
[17] Perdew J P and Zunger A 1981 Self-interaction correction to density-functional approximations for many-electron systems *Phys. Rev. B* **23** 5048
[18] Cramer C, Włoch M, Piecuch P, Puzzarini C and Gagliardi L 2006 Theoretical models on the Cu_2O_2 torture track: mechanistic implications for oxytyrosinase and small–molecule analogues *J. Phys. Chem. A* **110** 1991

[19] Gräfenstein J and Cremer D 2005 Development of a CAS-DFT method covering non-dynamical and dynamical electron correlation in a balanced way *Mol. Phys.* **103** 279
[20] Gräfenstein J, Kraka E, Filatov M and Cremer D 2002 Can unrestricted density functional theory describe open shell singlet biradicals? *Int. J. Mol. Sci.* **3** 360
[21] Kümmel S and Kronik L 2008 Orbital-dependent density functionals: theory and applications *Rev. Mod. Phys.* **80** 3
[22] Lee D, Furche F and Burke K 2010 Accuracy of electron affinities of atoms in approximate density functional theory *J. Phys. Chem. Lett.* **1** 2124
[23] Becke A D 2014 Perspective: fifty years of density-functional theory in chemical physics *J. Chem. Phys.* **140** 18A301
[24] Burke K 2012 Perspective on density functional theory *J. Chem. Phys.* **136** 150901
[25] Berland K and Hyldgaard P 2014 Exchange functional that tests the robustness of the plasmon description of the van der Waals density functional *Phys. Rev. B* **89** 035412
[26] Hamada I 2014 van der Waals density functional made accurate *Phys. Rev. B* **89** 121103(R)
[27] Berland K, Arter C A, Cooper V R, Lee K, Lundqvist B I, Schröder E, Thonhauser T and Hyldgaard P 2014 van der Waals density functionals built upon the electron-gas tradition: facing the challenge of competing interactions *J. Chem. Phys.* **140** 18A539
[28] Gharaee L, Erhart P and Hyldgaard P 2017 Finite-temperature properties of non-magnetic transition metals: comparison of the performance of constraint-based semi and nonlocal functionals *Phys. Rev. B* **95** 085147
[29] Tran F, Kalantari L, Traoré B, Rocquefelte X and Blaha P 2019 Nonlocal van der Waals functionals for solids: choosing an appropriate one *Phys. Rev. Mater.* **3** 063602
[30] Gerrits N, Smeets E W F, Vuckovic S, Powell A D, Doblhoff-Dier K and Kroes G-J 2020 Density functional theory for molecule-metal surface reactions: when does the generalized gradient approximation get it right and what to do if it does not *Phys. Chem. Lett.* **11** 10552
[31] Perdew J P and Levy M 1983 Physical content of the exact Kohn–Sham orbital energies: band gaps and derivative discontinuities *Phys. Rev. Lett.* **51** 1884
[32] Stein T, Eisenberg H, Kronik L and Baer R 2010 Fundamental gaps in finite systems from eigenvalues of a generalized Kohn–Sham method *Phys. Rev. Lett.* **105** 266802
[33] Borghi G, Ferretti A, Nguyen N L, Dabo I and Marzari N 2014 Koopmans-compliant functionals and their performance against reference molecular data *Phys. Rev. B* **90** 075135
[34] Nguyen N L, Borghi G, Ferretti A and Marzari N 2016 First-principles photoemission spectroscopy of DNA and RNA nucleobases from Koopmans-compliant functionals *J. Chem. Theory Comput.* **12** 3948
[35] Perdew J P *et al* 2017 Understanding band gaps of solids in generalized Kohn–Sham theory *Proc. Natl Acad. Sci.* **114** 2801
[36] Wing D, Ohad G, Haber J B, Filip M R, Gant S E, Neaton J B and Kronik L 2021 Band gaps of crystalline solids from Wannier-localization–based optimal tuning of a screened range-separated hybrid functional *Proc. Natl Acad. Sci.* **118** e2104556118
[37] Fetter A L and Walecka J D 1971 *Quantum Theory of Many-Particle Systems* (McGraw-Hill)
[38] Aulbur W G, Jönsson L and Wilkins J W 2000 Quasiparticle calculations in solids *Solid State Physics* vol 54, ed F Seitz, D Turnbull and H Ehrenreich (Academic) p 1
[39] Hedin L 1965 New method for calculating the one-particle green's function with application to the electron-gas problem *Phys. Rev.* **139** A796
[40] Hedin L and Lundqvist S 1969 Effects of electron-electron and electron-phonon interactions on the one-electron states of solids *Solid State Physics* vol 23, ed F Seitz, D Turnbull and H Ehrenreich (Academic) p 1
[41] Hedin L 1980 Effects of recoil on shake-up spectra in metals *Phys. Scr.* **21** 477
[42] Roca-Sanjuan D, Merchan M, Serrano-Andres L and Rubio M 2008 *Ab initio* determination of the electron affinities of DNA and RNA nucleobases *J. Chem. Phys.* **129** 095104
[43] Rangel T, Berland K, Sharifzadeh S, Brown-Altvater F, Lee K, Hyldgaard P, Kronik L and Neaton J B 2016 Structural and excited-state properties of oligoacene crystals from first principles *Phys. Rev. B* **93** 115206
[44] Qian X, Umari P and Marzari N 2011 Photoelectron properties of DNA and RNA bases from many-body perturbation theory *Phys. Rev. B* **84** 075103
[45] Kohn W and Sham L J 1965 Self-consistent equations including exchange and correlation effects *Phys. Rev.* **140** A1133
[46] Seidl A, Görling A, Vogl P, Majewski J A and Levy M 1996 Generalized Kohn–Sham schemes and the band-gap problem *Phys. Rev. B* **53** 3764
[47] Hohenberg P and Kohn W 1964 Inhomogeneous electron gas *Phys. Rev.* **136** B864
[48] Hedin L and Lundqvist B I 1971 Explicit local exchange-correlation potentials *J. Phys. C: Solid State Phys.* **4** 2064
[49] Gunnarsson O and Lundqvist B I 1976 Exchange and correlation in atoms, molecules and solids by the spin-density-functional formalism *Phys. Rev. B* **13** 4274
[50] Langreth D C and Perdew J P 1977 Exchange-correlation energy of a metallic surface: wave-vector analysis *Phys. Rev. B* **15** 2884

[51] Gunnarsson O, Jonson M and Lundqvist B I 1979 Descriptions of exchange and correlation effects in inhomogeneous electron systems *Phys. Rev. B* **20** 3136
[52] Langreth D C and Perdew J P 1980 Theory of nonuniform electronic systems. I. Analysis of the gradient approximation and a generalization that works *Phys. Rev. B* **21** 5469
[53] Langreth D C and Mehl M J 1981 Easily implementable nonlocal exchange-correlation energy functional *Phys. Rev. Lett.* **47** 446
[54] Langreth D C and Vosko S H 1987 Exact electron-gas response functions at high density *Phys. Rev. Lett.* **59** 497
[55] Becke A D 1988 Density-functional exchange-energy approximation with correct asymptotic behavior *Phys. Rev. A* **38** 3098
[56] Becke A D 1993 Density-functional thermochemistry. III. The role of exact exchange *J. Chem. Phys.* **98** 5648
[57] Heyd J, Scuseria G E and Ernzerhof M 2003 Hybrid functionals based on a screened Coulomb potential *J. Chem. Phys.* **118** 8207
[58] Becke A D and Johnson E R 2005 A density-functional model of the dispersion interaction *J. Chem. Phys.* **123** 154101
[59] Thonhauser T, Cooper V R, Li S, Puzder A, Hyldgaard P and Langreth D C 2007 van der Waals density functional: self-consistent potential and the nature of the van der Waals bond *Phys. Rev. B* **76** 125112
[60] Perdew J P, Parr R G, Levy M and Balduz Jr J L 1982 Density-functional theory for fractional particle number: derivative discontinutities of the energy *Phys. Rev. Lett.* **49** 1691
[61] Levy M, Perdew J P and Sahni V 1983 Exact differential equation for the density and ionization energy of a many-particle system *Phys. Rev. A* **30** 2745
[62] Godby R W, Schlüter M and Sham L J 1986 Accurate exchange-correlation potential for silicon and its discontinuity on addition of an electron *Phys. Rev. Lett.* **56** 2415
[63] Krieger J B, Li Y and Iafrate G J 1992 Construction and application of an accurate local spin-polarized Kohn–Sham potential with integer discontinuity: exchange-only theory *Phys. Rev. A* **45** 101
[64] Kuisma M, Ojanen J, Enkovaara J and Rantala T T 2010 Kohn–Sham potential with discontinuity band gap for materials *Phys. Rev. B* **82** 105106
[65] Racioppi S, Lolur P, Hyldgaard P and Rahm M 2023 A density functional theory for the average electron energy *J. Chem. Theory Comput.* **19** 799
[66] Gritsenko O, van Leeuwen R, van Lenthe E and Baerends E J 1995 Self-consistent approximation to the Kohn–Sham exchangepotential *Phys. Rev. A* **51** 1944
[67] Kraisler E and Kronik L 2013 Piecewise linearity of approximate density functionals revisited: implications for frontier orbital energies *Phys. Rev. Lett.* **110** 126403
[68] Rafaely-Abramson S, Sharifzadeh S, Gocvind N, Autschbach J, Neaton J B, Baer R and Kronik L 2012 Quasiparticle spectra from a nonempirical optimally tuned range-separated hybrid density functional *Phys. Rev. Lett.* **109** 226405
[69] Rafaely-Abramson S, Sharifzadeh S, Jain M, Baer R, Neaton J B and Kronik L 2013 Gap renormalization of molecular crystals from density functional theory *Phys. Rev. B* **88** 081204(R)
[70] Liu Z-F, Egger D A, Rafaely-Abramson S, Kronik L and Neaton J B 2017 Energy level alignment at molecule-metal interfaces from an optimally tuned range-separated hybrid functional *J. Chem. Phys.* **146** 092326
[71] Giannozzi P *et al* 2009 QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials *J. Phys.: Condens. Matter* **21** 395502
[72] Carnimeo I, Baroni S and Giannozzi P 2019 Fast hybrid density-functional computations using plane-wave basis sets *Electron. Struct.* **1** 015009
[73] Lin L 2016 Adaptively compressed exchange operator *J. Chem. Theory Comput.* **12** 2242
[74] Giannozzi P *et al* 2017 Advanced capabilities for materials modelling with quantum espresso *J. Phys.: Condens. Matter* **29** 465901
[75] Heyd J, Scuseria G E and Ernzerhof M 2006 Erratum: "Hybrid functionals based on a screened Coulomb potential" [J. Chem. Phys. 118, 8207 (2003)] *J. Chem. Phys.* **124** 219906
[76] Ernzerhof M and Perdew J P 1998 Generalized gradient approximation to the angle- and system-averaged exchange hole *J. Chem. Phys.* **109** 3313
[77] Henderson T M, Janesko B G and Scusseria G E 2008 Generalized gradient approximation model exchange holes for range-separated hybrids *J. Chem. Phys.* **128** 194105
[78] Löwdin P-O 1955 Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin orbitals and convergence problems in the method of configurational interaction *Phys. Rev.* **97** 1474
[79] Coleman T L 1963 Structure of fermion density matrices *Rev. Mod. Phys.* **32** 668
[80] Gilbert T L 1975 Hohenberg-Kohn theorem for nonlocal external potentials *Phys. Rev. B* **12** 2111
[81] Osburn J E and Levy M 1987 Energies, electron densities and first-order density matrices from one-body external potentials. II *Phys. Rev. A* **35** 3233

[82] Yang W, Zhang Y and Ayers P W 2000 Degenerate ground state and a fractional number of electrons in density and reduced density matrix functional theory *Phys. Rev. Lett.* **84** 5172
[83] Chong D P, Gritsenko O V and Baerends E J 2002 Interpretation of the Kohn–Sham orbital energies as approximative vertical ionization potentials *J. Chem. Phys.* **116** 1760
[84] Levy M 1995 Excitation energies from density-functional orbital energies *Phys. Rev. B* **52** R4313
[85] Piris M 2006 A new approach for the two-electron ciumulant in natural orbital functional theory *Int. J. Quantum Chem.* **106** 1093
[86] Fuchs F, Furthmüller J, Bechstedt F, Shishkin M and Kresse G 2007 Quasiparticle band structure based on a generalized Kohn–Sham scheme *Phys. Rev. B* **76** 115109
[87] Quintero-Monsebaiz R, Perea-Ramirez L I, Piris M and Vela A 2021 Spectroscopic properties of open shell diatomic molecules using Piris natural orbital functionals *Phys. Chem. Chem. Phys.* **23** 2953
[88] Perdew J P and Wang Y 1992 Pair-distribution function and its coupling-constant average for the spin-polarized electron gas *Phys. Rev. B* **46** 12947
[89] Perdew J P, Burke K and Ernzerhof M 1996 Generalized gradient approximation made simple *Phys. Rev. Lett.* **77** 3865
[90] Ou-Yang H and Levy M 1990 Theorem for exact local exchange potential *Phys. Rev. Lett.* **65** 1036
[91] Savin A, Umrigar C J and Gonze X 1998 Relationship of Kohn–Sham eigenvalues to excitation energies *Chem. Phys. Lett* **288** 391
[92] Perdew J P and Yue W 1986 Accurate and simple density functional for the electronic exchange energy: generalized gradient approximation *Phys. Rev. B* **33** 8800
[93] Perdew J P, Burke K and Wang Y 1996 Generalized gradient approximation for the exchange-correlation hole of a many-electron system *Phys. Rev. B* **54** 16533
[94] Perdew J P, Ernzerhof M and Burke K 1996 Rationale for mixing exact exchange with density functional approximations *J. Chem. Phys.* **105** 9982
[95] Burke K, Ernzerhof M and Perdew J P 1997 The adiabatic connection method: a non-empirical hybrid *Chem. Phys. Lett.* **265** 115
[96] Langreth D C and Perdew J P 1975 The exchange-correlation energy of a metallic surface *Solid State Commun.* **17** 1425
[97] Vydrov O A and Van Voorhis T 2010 Nonlocal van der Waals density functional: the simpler the better *J. Chem. Phys.* **133** 244103
[98] Adamo C and Barone V 1999 Towards reliable density functional methods without adjustable parameters: the PBE0 model *J. Chem. Phys.* **110** 6158
[99] Skone J H, Govoni M and Galli G 2014 Self-consistent hybrid functional for condensed systems *Phys. Rev. B* **89** 195112
[100] Bischoff T, Wiktor J, Chen W and Pasquarello A 2019 Nonemperical hybrid functionals for band gaps of inorganic metal-halide perovskites *Phys. Rev. Mater.* **3** 123802
[101] Bischoff T, Reshetnyak I and Pasquarello A 2019 Adjustable potential probes for band-gap predictions of extended systems through nonempirical hybrid functionals *Phys. Rev. B* **99** 201114(R)
[102] Cramer C, Kinal A, Włoch M, Piecuch P and Gagliardi L 2006 Theoretical characterization of end-on and side-on peroxide coordination in ligated Cu_2O_2 models *J. Phys. Chem. A* **110** 11557
[103] Cramer C, Smith B and Tolman W 1996 *Ab initio* characterization of the isomerism between the μ–η^2:η^2–peroxo– and bis(μ–oxo)dicopper Cores *J. Am. Chem. Soc.* **118** 11283
[104] Gräfenstein J and Cremer D 2000 Can density functional theory describe multi-reference systems? Investigation of carbenes and organic biradicals *Phys. Chem. Chem. Phys.* **2** 2091
[105] Maggs A C and Ashcroft N W 1987 Electronic fluctuation and cohesion in metals *Phys. Rev. Lett.* **59** 113
[106] Hyldgaard P, Berland K and Schröder E 2014 Interpretation of van der Waals density functionals *Phys. Rev. B* **90** 075148
[107] Rapcewicz K and Ashcroft N W 1991 Fluctuation attraction in condensed matter: a nonlocal functional approach *Phys. Rev. B* **44** 4032
[108] Andersson Y, Langreth D C and Lundqvist B I 1996 van der Waals interactions in density-functional theory *Phys. Rev. Lett.* **76** 102
[109] Langreth D C, Dion M, Rydberg H, Schröder E, Hyldgaard P and Lundqvist B I 2005 van der Waals density functional theory with applications *Int. J. Quantum Chem.* **101** 599
[110] Goerigk L, Hansen A, Bayer C, Ehrlich S, Najibi A and Grimme S 2017 A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions *Phys. Chem. Chem. Phys.* **19** 32184
[111] See supplemental material for tables and figures summarizing convergence studies as well as performance benchmarks.
[112] Janak J F 1978 Proof that $\partial E/\partial n_i = \epsilon_i$ in density-functional theory *Phys. Rev. B* **18** 7165
[113] Dabo I, Ferretti A, Poilvert N, Li Y, Marzari N and Cococcioni M 2010 Koopmans' condition for density functional theory *Phys. Rev. B* **82** 115121

[114] Zheng X, Cohen A J, Mori-Sanchez P, Hu X and Yang W 2011 Improving band gap prediction in density functional theory from molecules to solids *Phys. Rev. Lett.* **107** 026403
[115] Görling A 2015 Exchange-correlation potentials with proper discontinuities for physical meaningful Kohn–Sham eigenvalues and band structures *Phys. Rev. B* **91** 245120
[116] Ma J and Wang L-W 2016 Using Wannier functions to improve solid band gap predictioons in density functional theory *Sci. Rep.* **6** 24924
[117] Nguyen N L, Colonna N, Ferrettie A and Marzari N 2018 Koopmans-compliant spectral functionals for extended systems *Phys. Rev. X* **8** 021051
[118] Colonna N, Gennaro R D, Linscot E and Marzari N 2022 Koopmans spectral functionals in periodic boundary conditions *J. Chem. Theory Comput.* **18** 5435
[119] Slater J C 1974 *Quantum Theory of Molecules and Solids* vol 14 (McGraw-Hill)
[120] Sham L J and Schlüter M 1983 Density-functional theory of the energy gap *Phys. Rev. Lett.* **51** 1888
[121] Levine Z and Allan D C 1989 Linear optical response in silicon and germanium including self-energy effects *Phys. Rev. Lett.* **63** 1719
[122] Levy M, Pathak R K, Perdew J P and Wei S 1987 Indirect-path methods for atomic and molecular energies and new Koopmans theorems *Phys. Rev. A* **36** 2491
[123] Anisimov V I, Zaanen J A and Andersen O K 1991 Band theory and Mott insulators: Hubbard U instead of Stoner I *Phys. Rev. B* **44** 943
[124] Stanton J F and Gauss J 1994 Analytic energy derivatives for ionized states described by the equation-of-motion coupled cluster method *J. Chem. Phys.* **101** 8938
[125] Casida M E 1995 Generalization of the optimized-effective-potential model to include electron correlation: a variational derivation of the Sham-Schlüter equation for the exact exchange-correlation potential *Phys. Rev. A* **51** 2005
[126] Grabo T and Gross E K U 1995 Density-functional theory using an optimized exchange-correlation potential *Chem. Phys. Lett.* **240** 141
[127] Görling A 1996 Density-functional theory for excited states *Phys. Rev. A* **54** 3912
[128] Bartlett R J, Del Bene J E, Perera S A and Mattie R P 1997 Ammonia: the prototypical lone pair molecule *J. Mol. Struct.: Theochem* **400** 157
[129] Städele M, Majewski J A, Vogl P and Göeling A 1997 Exact Kohn–Sham exchange potential in semiconductors *Phys. Rev. Lett.* **79** 2089
[130] Cococcioni M and de Gironcoli S 2005 Linear response approach to the calculation of the effective interaction parameters in the LDA+U method *Phys. Rev. B* **71** 035105
[131] Kulik H J, Cococcioni M, Scherlis D A and Marzari N 2005 Density functional theory in transition-metal chemistry: a self-consistent Hubbard *U* approach *Phys. Rev. Lett.* **97** 103001
[132] Anisimov V I and Kozhevnikov A V 2005 Transition state method and Wannier functions *Phys. Rev. B* **72** 075125
[133] Becke A D and Johnson E R 2006 A simple effective potential for exchange *J. Chem. Phys.* **124** 221101
[134] Grüning M, Marini A and Rubio A 2006 Effect of spatial nonlocality on the density functional band gap *Phys. Rev. B* **74** 161103
[135] Cohen A J, Mori-Sanchez P and Yang W 2008 Fractional charge perspective on the band gap in density-functional theory *Phys. Rev. B* **77** 115123
[136] Mori-Sanchez P, Cohen A J and Yang W 2008 Localization and delocalization errors in density functional theory and implications for band-gap predictions *Phys. Rev. Lett.* **100** 146401
[137] Li C, Zheng X, Cohen A J, Mori-Sanchez P and Yang W 2015 Local scaling correction for reducing delocalization error in density functional approximations *Phys. Rev. Lett.* **114** 053001
[138] Weng M, Li S, Kie Ma J Z, Pan F and Wang L-W 2017 Wannier Koopman method calculations of the band gaps of alkali halides *Appl. Phys. Lett.* **111** 054101
[139] Quintero-Monsebaiz R and Loos P-F 2023 Equation generator for equation-of-motion coupled cluster assisted by computer algebra system *AIP Adv.* **13** 085035
[140] Tozer D J and Handy N C 1998 Improving virtual Kohn–Sham orbitals and eigenvalues: application to excitation energies and static polarizabilities *J. Chem. Phys.* **109** 10180
[141] Koch W and Holthausen M C 2002 *A Chemist's Guide to Density Functional Theory* (Wiley-VCH)
[142] Skone J H, Govoni M and Galli G 2016 Nonempirical range-separated hybrid functionals for solids and molecules *Phys. Rev. B* **93** 235106
[143] Miceli G, Chen W, Reshetnyak I and Pasquarello A 2018 Nonempirical hybrid functionals for band gaps and polaronic distortions in solids *Phys. Rev. B* **97** 121112(R)
[144] Görling A and Levy M 1993 Correlation-energy functional and its high-density limit obtained from a coupling-constant perturbation expansion *Phys. Rev. B* **47** 13105
[145] Jiao Y, Schröder E and Hyldgaard P 2018 Signatures of van der Waals binding: a coupling-constant scaling analysis *Phys. Rev. B* **97** 085115
[146] Langreth D C 1970 Singularities in the x-ray spectra of metals *Phys. Rev. B* **1** 471
[147] We use atomic units through out except when otherwise noted.

[148] Rydberg H, Dion M, Jacobson N, Schröder E, Hyldgaard P, Simak S I, Langreth D C and Lundqvist B I 2003 van der Waals density functional for layered structures *Phys. Rev. Lett.* **91** 126402
[149] Dougherty D, Younathan E, Voll R, Abdulnur S and McGlynn S 1978 Photoelectron spectroscopy of some biological molecules *J. Electron Spectrosc. Relat. Phenom.* **13** 379
[150] Hendricks J H, Lyapustina S A, de Clercq H L, Snodgrass J T and Bowen K H 1996 Dipole bound, nucleic acid base anions studied via negative ion photoelectron spectroscopy *J. Chem. Phys.* **104** 7788
[151] Desfrancois C, Abdoul-Carima H and Schermann J P 1996 Electron attachment to isolated nucleic acid bases *J. Chem. Phys.* **104** 7792
[152] Schiedt J, Weinkauf R, Neumark D M and Schlag E W 1998 Anion spectroscopy of uracil, thymine and the amino-oxo and amino-hydroxy tautomers of cytosine and their water clusters *Chem. Phys.* **239** 511
[153] Svozil D, Jungwirth P and Havlas Z 2004 Electron binding to nucleic acid bases. Experimental and theoretical studies. A review *Collect. Czech. Chem. Commun.* **69** 1395
[154] Trofimov A B, Schirmer J, Kobychev V B, Potts A W, Holland D M and Karlsson L 2006 Photoelectron spectra of the nucleobases cytosine, thymine and adenine *J. Phys. B: At. Mol. Opt. Phys.* **39** 305
[155] Zaytseva D I L, Trofinov A B, Schirmer J, Plekan O, Feyer V, Richter R, Coreno M and Prince K C 2009 Theoretical and experimental study of valence-shell ionization spectra of guanine *J. Phys. Chem. A* **113** 15142
[156] Tripathi D and Dutta A K 2019 Electron attachment to DNA base pairs: an interplay of dipole- and valence-bound states *J. Phys. Chem. A* **123** 10131
[157] Mahan G D 1990 *Many-Particle Physics* 2nd edn (Plenum Press)
[158] Prendergast D and Galli G 2006 X-ray absorption of water from first-principle calculations *Phys. Rev. Lett.* **96** 215502
[159] Patra A, Jana S, Constantin L A and Samal P 2020 Efficient yet accurate dispersion-corrected semilocal exchange–correlation functionals for non-covalent interactions *J. Chem. Phys.* **153** 084117
[160] Makov G and Payne M C 1995 Periodic boundary conditions in *ab initio* calculations *Phys. Rev. B* **51** 4014
[161] Dabo I, Kozinsky B, Singh-Miller N E and Marzari N 2008 Electrostatics in periodic boundary conditions and real-space corrections *Phys. Rev. B* **77** 115139
[162] Andreussi O, Dabo I and Marzari N 2012 Revised self-consistent continuum solvation in electronic-structure calculations *J. Chem. Phys.* **136** 064102
[163] Nattino F, Dupont C, Marzari N and Andraussi O 2019 Functional extrapolations to tame unbound anions in density-functional theory calculations *J. Chem. Theory Comput.* **15** 6313
[164] Lee C, Yang W and Parr R G 1988 Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density *Phys. Rev. B* **37** 785
[165] Stephens P J, Devlin F J, Chabalowski C F and Frisch M J 1994 *Ab initio* calculation of vibrational absorption and circular dichroism spectra using density functional force fields *J. Phys. Chem.* **98** 11623
[166] Grimme S, Antony J, Ehrlich S and Krieg H 2010 A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu *J. Chem. Phys.* **132** 154104
[167] Mehta N, Casanova-Paez M and Goerigk L 2018 Semi-empirical or non-empirical double-hybrid density functionals: which are more robust? *Phys. Chem. Chem. Phys.* **20** 23175
[168] Najibi A, Casanova-Paez M and Goerigk L 2021 Analysis of recent BLYP- and PBE-based range-separated double-hybrid density functional approximations for main-group thermochemistry, kinetics and noncovalent interactions *J. Phys. Chem. A* **125** 4026
[169] Kruse H, Banas P and Sponer J 2019 Investigations of stacked DNA base-pair steps: highly accurate stacking interaction energies, energy decomposition and many-body stacking effects *J. Chem. Theory Comput.* **15** 95
[170] Cooper V R, Thonhauser T, Puzder A, Schröder E, Lundqvist B I and Langreth D C 2008 Stacking interactions and the twist of DNA *J. Am. Chem. Soc.* **130** 1304
[171] Li S, Cooper V R, Thonhauser T, Lundqvist B I and Langreth D C 2009 Stacking interactions and DNA intercalation *J. Phys. Chem. B* **113** 11166
[172] Riplinger C, Pinski P, Becker U, Valeev E F and Neese F 2016 Sparse maps—a systematic infrastructure for reduced-scaling electronic structure methods. II. Linear-scaling domain based pair natural orbital coupled cluster theory *J. Chem. Phys.* **144** 024109
[173] Linscott E B, Colonna N, De Gennaro R, Nguyen N L, Borghi G, Ferretti A, Dabo I and Marzari N 2023 Koopmans: an open source package for accurately and efficiently predicting spectral properties with Koopmans functionals *J. Chem. Theory Comput.* **19** 7097
[174] Borghi G, Park C-H, Nguyen N L, Ferretti A and Marzari N 2015 Variational minimization of orbital-density-dependent functionals *Phys. Rev. B* **91** 115121

[175] Rohrer J and Hyldgaard P 2011 Stacking and band structure of van der Waals bonded graphane multilayers *Phys. Rev. B* **83** 165423
[176] Tao J, Jiao Y, Mo Y, Yang Z-H, Zhu J-X, Hyldgaard P and Perdew J P 2018 First-principles study of the binding energy between nanostructures and its scaling with system size *Phys. Rev. B* **97** 155143
[177] Onida G, Reining L, Godby R W, Sole R D and Andreoni W 1995 *Ab initio* calculations of the quasiparticle and absorption spectra of clusters: the sodium tetramer *Phys. Rev. Lett.* **75** 818
[178] Rohlfing M and Louie S G 2000 Electron-hole excitations and optical spectra from first principles *Phys. Rev. B* **62** 4927
[179] Onida G, Reining L and Rubio A 2002 Electronic excitations: density-functional versus many-body Green's-function approaches *Rev. Mod. Phys.* **74** 601
[180] Runge E and Gross E K U 1984 Density-functional theory for time-dependent systems *Phys. Rev. Lett.* **52** 997
[181] Berman O and Mukamel S 2003 Quasiparticle density-matrix representation of nonlinear time-dependent density functional response functions *Phys. Rev. A* **67** 042503
[182] Hellgren M and Gross E K U 2012 Effect of discontinuities in Kohn–Sham-based chemical reactivity theory *J. Chem. Phys.* **136** 114102
[183] Hellgren M, Räsänen E and Gross E K U 2013 Optimal control of strong-field ionization with time-dependent density-functional theory *Phys. Rev. A* **88** 013414
[184] Tavernelli I 2014 Nonadiabatic molecular dynamics simulations: synergies between theory and experiments *Acc. Chem. Res.* **48** 792
[185] Kowalewski M, Bennett K and Mukamel S 2016 Non-adiabatic dynamics of molecules in optical cavities *J. Chem. Phys.* **144** 054309
[186] Cho D, Rouxel J R, Kowalewski M, Saurabh P, Lee J Y and Mukamel S 2018 Phase cycling RT-TDDFT simulation protocol for nonlinear XUV and x-ray molecular spectroscopy *J. Phys. Chem. Lett.* **9** 1072
[187] Ceperley D M and Alder B J 1980 Ground state of the electron gas by a stochastic method *Phys. Rev. Lett.* **45** 566