PCCP

ROYAL SOCIETY OF CHEMISTRY

REVIEW

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2026, **28**, 3779

Contemporary DFT: learning from traditional and recent trends for the development and assessment of accurate exchange–correlation functionals

E. Brémond, [a] A. J. Pérez-Jiménez, [b] C. Adamo [c] and J. C. Sancho-García *[b]

Density functional theory (DFT) is the most widely and accepted model for calculating the electronic structure of physical systems, but practical applications rely on approximations to the (still) unknown exact exchange–correlation functional. We revisit here the main advances, families, and tipping points for the development of accurate expressions for the exchange–correlation functional, focusing on the historical evolution followed by DFT but also on the underlying reasons for that, while emphasizing both the theoretical foundations and the last methodological and technical developments (exemplified by deep-learned models). The latter are built by training a neural network on a very extended collection of molecular data, with the DM21, the aPBE0-ML, and the Skala functionals now available using such a strategy. This degree of development could have not been possible without the knowledge achieved so far, after more than half a century of investigations and applications of DFT to all kind of systems, and it is rooted on the traditional local (LDA), semi-local (GGA or meta-GGA) and non-local (hybrid and double-hybrid) functionals, which are still important pillars of DFT and are expected to coexist with deep-learned models, as well as on the creation of large and diverse datasets of nearly-exact reference results, which are also needed to train any of the deep-learned models envisioned.

Received 2nd September 2025,
Accepted 26th December 2025

DOI: 10.1039/d5cp03373j

rsc.li/pccp

1 Introduction

It is no exaggeration to say that density functional theory (DFT) forms part now of the core of quantum and computational chemistry, electronic structure theories, atomic and molecular physics, theoretical spectroscopy and light-matter interactions, finite-size or periodic materials science, and condensed matter physics, to name just a few of scientific fields more proned to use this theory widely and routinely. Evidently, given the historical and current importance of these fields for the whole scientific knowledge achieved so far in Chemistry, Physics, and Materials Science, this statement can be scaled up to imagine applications and developments comprising systems and properties from all the elements of the periodic table and from all states of matter. Since the main ingredient of any DFT calculation is the exchange–correlation functional, and actually the key for its impactful success and universal adoption by the scientific community, we will focus on its evolution in this perspective, which is summarized in Table 1: the first models were proposed in the late 50s, while the most modern machine-learned models date from the early 2020s, thus covering more than half a century (seven decades in fact) of continuous improvements, studies, and validations.

Therefore, considering a context focused on the development and assessment of exchange–correlation functionals, DFT could be now experiencing a ''new age'', with the developers of modern functionals based on Machine Learning (ML) advances invoking a ''game change'', after a set of deep-learned exchange–correlation functionals have recently appeared with the hope of reaching a competitive accuracy, and at a computational cost similar or even lower, compared with the best among the more traditional models. To better put these new models into the due perspective, we will first summarize here the main theoretical frameworks successfully employed so far for the development of exchange–correlation functionals, with the most modern expression being sufficiently accurate and of general applicability to drive such an immense success of DFT in recent times. We will restrict ourselves to ground-state formulations, not only for simplicity but also because deep-learned expressions are not extensively applied to excited states

[a] *Université Paris Cité, ITODYS, CNRS, F-75006 Paris, France*
[b] *Department of Physical Chemistry, University of Alicante, E-03080 Alicante, Spain. E-mail: jc.sancho@ua.es*
[c] *Chimie ParisTech, PSL Research University, CNRS, Institute of Chemistry for Life and Health Sciences (i-CLeHS), FRE 2027, F-75005 Paris, France*

Table 1 Summary of the developments and families of functionals (with some examples) of DFT

Jacob's ladder rung	Scaling	Examples	Availability
LDA	$O(N^3)$	SVWN	>1950s
GGA	$O(N^3)$	BLYP (semiempirical), PBE (non-empirical),	>1980s
meta-GGA	$O(N^3)$	M06-L (semiempirical), r^2SCAN (non-empirical)	>1990s
Hybrid	$O(N^4)$	B3LYP (semiempirical), PBE0 (non-empirical), ωB97X (semiempirical), RSX-PBE0 (non-empirical)	>1990s
Double-hybrid	$O(N^5)$	B2-PLYP (semiempirical), PBE-QIDH (non-empirical), ωB97X-2 (semiempirical), RSX-PBE-QIDH (non-empirical)	>2000s
Deep-learned		DM21, CF22D, aPBE0-ML, Skala, *etc.*	>2020s

yet. We refer the readers to literature reviews[1–7] on excited-state methods such as TD-DFT, GW, or DFT+DMFT for a deeper discussion. We will also follow a chronological order to also inform the readers about the timeline followed by DFT to reach its current state-of-the-art.

2 A route to DFT through its bibliographic and historical importance

We will first illustrate the importance of DFT along the history by resorting first to bibliographic arguments. In 2014, a literature survey was conducted to disclose the top-10 most cited scientific publications for modern Science,[8] among a larger collection of top-100 studies which was also recently updated.[9] Entering into that top-100 list required more than 12 000 citations in 2014. However, that threshold was increased to 30 000 in 2025, which also evidences this increase as a byproduct of the large volume of bibliographic records occurred in the last decade. The top-10 list from 2014 included two of the expressions most used for the exchange–correlation functionals: the correlation functional developed[10] by Lee–Yang–Parr (LYP) in 1988 and the first hybrid functional[11] (named later as B3LYP) developed by Becke in 1993. It is important to remark that about half of the papers in the former (2014) top-100 list have changed due to the irruption of newer papers in the updated (2025) list. However, the update made in 2025 of the top-10 still contains two expressions for exchange–correlation functionals: the LYP correlation functional again, together with the Perdew–Burke–Erzenhof (PBE) exchange and correlation functionals developed[12] in 1996 (the previous B3LYP expression ranks now 13). Interestingly, comparing both LYP and PBE models one can already observe another difference which has always permeated the field: while both expressions respond to a solid and deep mathematical formulation, the LYP expression (or the B3LYP one) includes a very reduced set of to-be-fitted parameters, while the PBE expression does not, with the values of the constants of the latter adjusted to reproduce known conditions and numerical limits. Actually, those DFT-based publications are gaining citations at an accelerating rate, and are accompanied into that top-10 by another manuscript dealing with the extension of DFT calculations from molecular to solid-state situations.[13] Furthermore, the foundational manuscripts for DFT are also included into both (2014 and 2025) top-100 lists: the Hohenberg–Kohn theorems[14] and the Kohn–Sham (KS) orbitalic (or one-electron) model[15] in positions 39 and 34 (2014) and 57 and 41 (2025), which also shows the importance of the theoretical framework general to any DFT modern calculation (and not only the importance of the particular model chosen for practical calculations). To conclude this bibliographic analysis, we also notice that the exchange functional developed[16] by Becke in 1988 is also listed in positions 25 (2014) and 48 (2025) and that an extension of DFT to mimic non-covalent interactions[17] is also part of the last top-100 list at position 92. Those numbers are certainly an impressive achievement for DFT from the bibliometric point of view.

Another milestone for DFT was the awarding in 1998 of the Nobel Prize in Chemistry to Walter Kohn[18] for ''his development of DFT'', shared with John A. Pople ''for his development of computational methods in quantum chemistry'', recognizing in such a way how DFT was a perfect companion for wave-function-based theories. To bracket the importance of DFT from a different point of view, it is pertinent to comment that the *libxc* library[19–21] of density functional approximations, which is also an example of the state-of-the-art of the field, includes over 600 expressions for the exchange–correlation functional and is being used or interfaced with more than 30 codes. Note that this holds independently of the type of software (proprietary or commercial) or the basis set nature (Slater-type, Gaussian-type, plane waves, or grid-free) used, to name just a pair of the existing differences between the set of massively adopted codes nowadays. Indeed, this proliferation of expressions, implementations, and codes also explains the rise in scientific publications during the last decades:[22] the yearly publications citing DFT in the 90s was approximately proportional to $O(10^3)$ but increased an order of magnitude the next decade, to $O(10^4)$, and is now doubling its number every 5 years.[23] This slope is only comparable to the rate of growth of publications for *e.g.* climate change, which is somehow surprising for a theory, with the due approximations, having the goal of understanding the electronic structure of matter after calculating the electronic energy and its changes according to the microscopic environment. The growing number of DFT expressions is also accompanied by a higher (formal) computational cost, but this higher scaling with the system size has not precluded the large number of applications leading to this impressive situation. The reasons for that can be roughly traced back to better algorithms and techniques[24–27] together with the exponential improvement in scientific computing along the last decades.

3 The form of the exchange–correlation functional as a timeline of DFT advances

A general, but unfortunately unknown, expression for any exchange–correlation functional, $E_{xc}[\rho(\mathbf{r})]$, arises from the contribution of the following terms:

$$E_{xc}[\rho(\mathbf{r})] = (V_{ee}[\rho(\mathbf{r})] - J[\rho(\mathbf{r})]) + (T[\rho(\mathbf{r})] - T_s[\rho(\mathbf{r})]), \quad (1)$$

with the density functionals entering the above equation being, respectively, the exact kinetic energy, $T[\rho(\mathbf{r})]$, the kinetic energy of the non-interacting particle system taken as reference, $T_s[\rho(\mathbf{r})]$, the electron–electron interaction energy, $V_{ee}[\rho(\mathbf{r})]$, and the classical charge density interaction, $J[\rho(\mathbf{r})]$. In other words, the exchange–correlation functional incorporates the non-classical effects of the type: (i) the (Fermi) correlation between same-spin electrons (or exchange energy), (ii) the (Coulomb) correlation between opposite-spin electrons (or correlation energy), and (iii) the difference between the kinetic energy of the interacting and the (fictitious or KS) non-interacting particle system. The explicit calculation of the latter term is omitted in practice and thus assumed to be part of the $E_{xc}[\rho(\mathbf{r})]$ total values. Note that the explicit (closed-shell, for simplicity) expressions of all the terms are known, since:

$$T_s[\rho(\mathbf{r})] = \sum_i^{N/2} \left\langle \phi_i \left| -\frac{1}{2}\nabla_i^2 \right| \phi_i \right\rangle, \quad (2)$$

$$V_{ee}[\rho(\mathbf{r})] = \frac{1}{2}\int d\mathbf{r} \int \frac{\rho(\mathbf{r})\rho_{xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}', \quad (3)$$

$$J[\rho(\mathbf{r})] = \frac{1}{2}\int d\mathbf{r} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}', \quad (4)$$

but not the corresponding (explicit) form of $E_{xc}[\rho(\mathbf{r})]$ which needs to be thus approximated. Note also that a set of self-consistently obtained orbitals ($\{\phi_i\}$) is also needed as part of the calculations, with $T_s[\rho(\mathbf{r})]$ using the KS orbitals for that, and from which the corresponding density is obtained as $\rho(\mathbf{r}) = \sum_i^{N/2} |\phi_i(\mathbf{r})|^2$. On the other hand, $\rho_{xc}(\mathbf{r},\mathbf{r}')$ is the exchange–correlation hole of an electron at $\mathbf{r}$, or the reduction in probability around each of the electrons to find another electron at $\mathbf{r}'$. One can thus assume that the possible (general) ingredients of the $E_{xc}[\rho(\mathbf{r})]$ expression can be the density at the reference (or local) point, $\rho(\mathbf{r})$, the density at another (non-local) spatially separated point, $\rho(\mathbf{r}')$, the mathematical form adopted or simplified for the exchange–correlation hole, $\rho_{xc}(\mathbf{r}, \mathbf{r}')$, and the occupied (or even virtual) orbitals, $\{\phi_i\}$.

We should also comment at this point some of the specific errors made in practical DFT calculations, arising from approximating this formal (and exact up to now) theoretical framework. First of all, it should be noticed that for the exact $E_{xc}[\rho(\mathbf{r})]$ expression, the $V_{ee}[\rho(\mathbf{r})]$ term is not affected by the so-called Self-Interaction Error (SIE). However, when approximated expressions are introduced for the exchange–correlation functional, SIE is not zero anymore, thus leading to one of the intrinsic problems of DFT[28,29] and its multiple analysis and approaches to understand and then reduce the impact on practical calculations.[30–35] On the other hand, a density-driven error has been recently identified and examined, providing a simple procedure for improving the results of DFT calculations in many situations upon using the Hartree–Fock density in place of the exact density.[36,37]

Of course, an outcome such as eqn (1) obeys to a solid theoretical framework, rooted first on the Hohenberg–Kohn theorems, proving the existence of an energy density functional for all the electronic systems which is also variationally well-defined, and on the Kohn–Sham approximation, simplifying the treatment after introducing a non-interacting particle systems whose associated density matches that of the exact system formed by interacting particles.[15] These KS orbitals, $\{\phi_i\}$, and their self-energies, ϵ_i, are self-consistently obtained as part of any DFT calculation after solving iteratively the corresponding one-particle equations:

$$\left\{ -\frac{1}{2}\nabla_i^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}) \right\} \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \quad (5)$$

with $v(\mathbf{r})$ the external potential of the electronic system, and $v_{xc}(\mathbf{r})$ the exchange–correlation potential or $v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})}$. However, the exchange–correlation functional also enters into this equation through the corresponding potential, explaining why its initial choice is so critical to the performance and accuracy of any DFT calculation, still deserving theoretical and computational efforts after more than half a century of developments. Finally, within this formalism, the exact electronic energy is therefore written as:

$$E[\rho(\mathbf{r})] = \int v(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} + T_s[\rho(\mathbf{r})] + J[\rho(\mathbf{r})] + E_{xc}[\rho(\mathbf{r})], \quad (6)$$

for which any prediction and comparison can be done further.

3.1 Former but pedagogical models

Arguably, the first time (1927) in which an electronic energy term could be written as an explicit functional of the density was part of the Thomas-Fermi (TF) model,[38,39] a semiclassical theory developed for the electronic structure of many-body systems arising shortly after the Schrödinger equation, circumventing for the first time in history the use of the wavefunction as the primary variable to obtain the energy and associated properties of an N-electron system. The total kinetic energy of the electrons in this model is given by $T[\rho(\mathbf{r})] = \int t(\mathbf{r}) d\mathbf{r} = C_{TF}\int \rho(\mathbf{r})^{5/3} d\mathbf{r}$, with C_{TF} a constant, which can be thus added to the potential energy of the electrons. However, that model neglected any difference for electron–electron interactions with respect to spin, which was later added thanks to the work[40] of Dirac (1930) and again in the simple form $E_x[\rho(\mathbf{r})] = -C_x\int \rho(\mathbf{r})^{4/3} d\mathbf{r}$, with C_x a constant, which also connects with the X_α method developed later (1951) by Slater. Despite the

limitations of these early theories, it might bring the tipping point needed to understand the origin of all the later efforts. Simply speaking, $E_{xc}[\rho(\mathbf{r})]$ is thought to be modelled as $\int e_{xc}[\rho(\mathbf{r})]d\mathbf{r}$, with e_{xc} the exchange–correlation energy per unit volume, with the possible separability of terms $e_{xc} = e_x + e_c$ for convenience. Later on, the successful treatment of the system[41] known as the "Uniform Electron Gas" corroborated this framework, leading to the same functional form for $T[\rho(\mathbf{r})]$ and $E_x[\rho(\mathbf{r})]$, but with $C_{TF} = \frac{3}{10}(3\pi^2)^{2/3}$ and $C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}$. The correlation energy of this model was also numerically modelled[42] by Vosko–Wilk–Nusair (VWN), giving rise to a functional form for $E_c[\rho(\mathbf{r})]$ too. Since both exchange and correlation functionals only depend on $\rho(\mathbf{r})$, the approach was consequently called Local Density Approximation (LDA) as it is known today.

3.2 GGA and meta-GGA models

The form of the LDA exchange and kinetic energies can also be derived from the coordinate scaling of the density, $\rho_\lambda(\mathbf{r}) = \lambda^3\rho(\lambda\mathbf{r})$, which is one of the known conditions to be satisfied by any model, and where the λ^3 factor ensures that the scaled density still integrates to N electrons as it should be. However, the poor predictions of LDA in molecular systems motivated to go a step further considering the inhomogeneity of matter around any spatial reference point. This was done by introducing the dimensionless variable $x(\mathbf{r}) = \frac{\nabla\rho(\mathbf{r})}{\rho^{4/3}(\mathbf{r})}$ into the integrand of the *e.g.* exchange functional, which also guarantees the correct coordinate scaling for this kind of expressions:

$$E_x^{GGA}[\rho(\mathbf{r})] = \int \rho(\mathbf{r})^{4/3} F_x(x(\mathbf{r}))d\mathbf{r}, \quad (7)$$

giving thus rise to the generalized gradient approximation (GGA). The enhancement factor F_x (or F_{xc} if the total exchange–correlation expression is instead considered) can take different mathematical forms depending on the route followed by the developers.[43–46] Note that if $\rho(\mathbf{r})$ tends to a constant value, the enhancement factor should vanish approaching the LDA limit, which can be thus viewed as another known condition to be imposed for developing models. If that factor is intended to bring some inhomogeneity, the simplest gradient expansion would thus be $F_x(x(\mathbf{r})) = 1 + x^2 + O(x^4)$ but it unfortunately diverges. Therefore, one of the most famous solutions[16] was proposed in 1988 by A. Becke as $F_x(x(\mathbf{r})) = 1 + \beta\frac{x^2}{1 + 6\beta x \sinh^{-1}(x)}$, which also fits to the long-range behaviour of the energy density, $\lim_{|r|\to\infty} F_x(x(\mathbf{r})) = -\frac{1}{2|\mathbf{r}|}$, another known condition to be fulfilled by approximate expressions. The introduction of β also allows for some numerical flexibility of the model, and it is actually fitted to reproduce a small set of atomic exchange energies as close as possible. Overall, GGA is not only numerically more accurate than LDA but it also permits to introduce some of the mathematical conditions and exact constraints known for the exact functional. Considering higher-order derivatives of the density, as $\nabla^2\rho(\mathbf{r})$ by itself or through the kinetic energy density $\tau(\mathbf{r}) = \frac{1}{2}\sum_i^{N/2}|\nabla\phi_i(\mathbf{r})|^2$, has also been explored successfully[47,48] leading to the next level of functionals known as meta-GGA. Additionally, this set of elements (*i.e.*, the density, its gradient, and higher-order derivatives) introduced step-by-step into the exchange–correlation integrand motivated their hierarchization, with LDA, GGA, and meta-GGA forming the three (lowest) ladders of a hierarchy known as Jacob's ladder[49] (not only symbolically but also considering the timeline of DFT).

On the other hand, there are also examples of GGA exchange–correlation functionals free of adjustable parameters, such as the prominent case of PBE, which introduced in 1996 an enhancement factor for the exchange functional of the form $F_x(s(\mathbf{r})) = 1 + \kappa - \frac{\kappa}{1 + \frac{\mu}{\kappa}x^2}$, with $s = x/x_0$ and $x_0 = 2(3\pi^2)^{1/3}$ able to satisfy the correct uniform gas limit, $F_x(0) = 1$, the spin-scaling relationship of the exact exchange energy, $E_x[\rho_\uparrow, \rho_\downarrow] = \frac{1}{2}(E_x[2\rho_\uparrow] + E_x[2\rho_\downarrow])$, the expected behaviour at small density variations, $F_x(x(\mathbf{r})) = 1 + \mu x^2$ if $x \to 0$, and the Lieb-Oxford bound, $E_x[\rho_\uparrow, \rho_\downarrow] \geq E_{xc}[\rho_\uparrow, \rho_\downarrow] \geq -1.679\int\rho(r)^{4/3}d\mathbf{r}$. Note that the values of κ and μ are not longer fitted to any external or reference data but obtained during the own derivation of the model, thus motivating the non-empirical adjective added to these models. The extension of this philosophy to the meta-GGA case is also possible, with the Tao–Perdew–Staroverov–Scuseria (TPSS) functional one of the step further in the hierarchy of models[50] thanks to the progressive introduction of more exact constraints without empirical parameters. The most recent example of meta-GGA models is probably the SCAN (strongly constrained and appropriately normed) family of functionals,[51,52] with the latest regularized-restored SCAN (or r^2SCAN) functional[53] including most of the exact constraints known for (meta-)GGAs. It is important to remark that the division between semiempirical and non-empirical functionals still impregnates the field.

3.3 Hybrid and double-hybrid models

All the LDA, GGA, and meta-GGA functionals share the formal (asymptotic) scaling of $O(N^3)$ with respect to the system size N, although differing in their numerical quadrature and possible algorithmic speedups, but also share some of the limitations imposed by the functional form chosen as starting point, see eqn (7). GGA and meta-GGA are also considered as semi-local functionals, in the sense that some degree of non-locality is introduced thanks to that dependence on $\nabla\rho(\mathbf{r})$ and $\nabla^2\rho(\mathbf{r})$, but the importance of non-locality has prompted other developments. For instance, it is well-known that local or semi-local functionals fail to deliver accurate band gaps of extended systems and key materials (*i.e.*, insulators or semiconductors) due to the self-interaction error,[54] and that the incorporation of

a fraction of non-local exchange reduces the systematic underestimation of band gaps.[55,56] On the other hand, the introduction of non-local exchange is also known to benefit thermochemistry and thermochemical kinetics calculations.[57,58] Finally, we also mention their importance for accurate predictions for electron transport, photocurrent generation, and photovoltaics applications.[59–61] Overall, global hybrid functionals add a relatively low-to-medium fraction of exact-exchange energy to a GGA or meta-GGA exchange functional, improving considerably their accuracy, which have therefore prompted its wide use in Chemistry, Materials Science, or Condensed Matter Physics.[62]

Actually, the adiabatic connection method (ACM) allows to go a step further with respect to GGA or meta-GGA models, representing the majority of the calculations performed today at least at the molecular and supramolecular scales, thus briefly revisiting here the foundations and formal framework. The motivation is to find a mathematically admissible connection[63–65] between the non-interacting (fictitious) and interacting (real) particle system, thanks to a λ-dependent electron–electron interaction term written as $\hat{V}_{ee}^{\lambda} = \sum_{i<j}^{N} \frac{\lambda}{|\mathbf{r}-\mathbf{r}'|}$. If $\lambda \in [0,1]$, both systems are defined at the limits $\lambda = 0$ (fictitious) and $\lambda = 1$ (real). The exchange–correlation functional therefore represents the interpolation path between them, expressed as:

$$E_{xc}[\rho(\mathbf{r})] = \int_{\lambda=0}^{\lambda=1} \mathcal{W}_{\lambda}[\rho(\mathbf{r})]\mathrm{d}\lambda. \quad (8)$$

with $\mathcal{W}_{\lambda}[\rho(\mathbf{r})] = \langle \Psi_{\lambda}[\rho(\mathbf{r})]|\hat{V}_{ee}^{\lambda}|\Psi_{\lambda}[\rho(\mathbf{r})]\rangle - J[\rho(\mathbf{r})]$,, with Ψ_{λ} the wavefunction minimizing $\langle \hat{T} + \lambda\hat{V}_{ee}\rangle$ whilst giving the exact density. The pending task is thus to guess the form of $\mathcal{W}_{\lambda}[\rho(\mathbf{r})]$, with the simplest (linear[66] and quadratic[67]) expressions being $\mathcal{W}_{\lambda}[\rho(\mathbf{r})] = \alpha[\rho] + \lambda\beta[\rho]$ and $\mathcal{W}_{\lambda}[\rho(\mathbf{r})] = \alpha[\rho] + \lambda\beta[\rho] + \lambda^2\gamma[\rho]$, respectively. The information needed to infer the final functional form of $E_{xc}[\rho(\mathbf{r})]$ is again based on a set of known conditions, such as:

$$\mathcal{W}_{\lambda=0} = E_x^{EXX}[\{\phi_i\}], \quad (9)$$

$$\mathcal{W}_{\lambda=1} = E_{xc}^{exact}[\rho(\mathbf{r})], \quad (10)$$

$$\mathcal{W}'_{\lambda=0} = \left(\frac{\partial \mathcal{W}_{\lambda}}{\partial\lambda}\right)_{\lambda=0} = 2E_c^{PT2}[\{\phi_i,\phi_a\}], \quad (11)$$

with the EXact-Exchange (EXX) energy, $E_x^{EXX}[\{\phi_i\}]$, the exact exchange–correlation energy, $E_{xc}^{exact}[\rho(\mathbf{r})]$, and the 2nd-order Perturbation Theory, $E_c^{PT2}[\{\phi_i, \phi_a\}]$, depending now on the set of occupied ($\{\phi_i\}$) and virtual ($\{\phi_a\}$) KS orbitals, respectively. The implications of this dependence on wavefunction-based energy terms cannot be underemphasized: any exchange–correlation functional might include exact-exchange and/or correlation energies additionally to the exchange and correlation energies given by a density functional, thus representing the next ladders of the hierarchy of models, fourth and fifth respectively, at the price of increasing the formal computational scaling from $O(N^3)$ (LDA, GGA, and mega-GGA) to $O(N^4)$ (hybrid), and $O(N^5)$ (double-hybrid).

Interestingly, the combination of $E_x^{EXX}[\{\phi_i\}]$, $E_c^{PT2}[\{\phi_i,\phi_a\}]$, and $E_{xc}[\rho(\mathbf{r})]$ energy terms can be easily controlled and tuned step-by-step by introducing a linear combination of them:

$$E_{xc}[\rho(\mathbf{r})] = \lambda_x E_x^{EXX}[\{\phi_i\}] + (1-\lambda_x)E_x[\rho(\mathbf{r})] + E_c[\rho(\mathbf{r})], \quad (12)$$

$$E_{xc}[\rho(\mathbf{r})] = \lambda_x E_x^{EXX}[\{\phi_i\}] + (1-\lambda_x)E_x[\rho(\mathbf{r})] + \lambda_c E_c^{PT2}[\{\phi_i,\phi_a\}] + (1-\lambda_c)E_c[\rho(\mathbf{r})], \quad (13)$$

leading to the hybrid and double-hybrid approximations, depending on λ_c being zero or not, respectively. Note the universality of this particular yet simple combination of energy terms, in the sense that in principle any form for the $E_x[\rho(\mathbf{r})]$ and the $E_c[\rho(\mathbf{r})]$ functionals can be chosen for that, and the flexibility, in the sense that the values of λ_x and λ_c are not initially restricted. We also note other non-linear forms[68,69] as alternative to eqn (13) also motivated by the ACM.

The preference for the fitting procedure of those values has led to the proliferation of expressions such as the original B3PW91[11] or the later popularized B3LYP,[70] for which the particular combination of energy terms is:

$$E_{xc}[\rho(\mathbf{r})] = aE_x^{EXX}[\{\phi_i\}] + (1-a-b)E_x^{LDA}[\rho(\mathbf{r})] + bE_x^{GGA}[\rho(\mathbf{r})] + (1-c)E_c^{LDA}[\rho(\mathbf{r})] + cE_c^{GGA}[\rho(\mathbf{r})], \quad (14)$$

with $a = 0.20$, $b = 0.72$, and $c = 0.81$ values obtained from their fitting to a set of thermochemical data taken as reference. It is worth to mention that since the first applications of the B3LYP functional,[71] the used coefficients are those proposed by Becke in its original work, obtained for the B3PW91 and thus not specifically optimized for B3LYP. Their optimization lead to a functional close to the so-called B1LYP.[72] However, it is also known that those values are functional-dependent and/or training-set dependent; for instance, using the OPTX exchange functional[73] instead, the coefficients of the corresponding O3LYP model become $a = 0.12$, $b = 0.93$, and $c = 0.81$. Another of the most widely applied hybrid functionals is M06-2X,[74] incorporating more than 30 parameters optimized simultaneously and including λ_x in the fit. The corresponding extension to doubly-hybrid functionals was pioneered by S. Grimme in 2006, with the B2-PLYP model,[75] for which the optimal fit to thermochemical data led to $\lambda_x = 0.53$ and $\lambda_c = 0.27$, without reoptimizing the parameters entering into the $E_x[\rho(\mathbf{r})]$ o $E_c[\rho(\mathbf{r})]$ functionals in line with the B3LYP model. One can easily observe how BLYP, B3LYP, and B2-PLYP would form a hierarchy of models (*i.e.*, any step up would ideally lead to more accurate values) according to the Jacob's ladder (see Fig. 1) as it has been thoroughly illustrated along the decades[76–79] with some exceptions too.[80–82]

The advantages of developing and working with non-empirical expressions (*i.e.*, PBE) can also be extended to these (double-)hybrid expressions too,[83] as it is exemplified by the parameter-free PBE0-n models:

$$E_{xc}^{PBE0-n}[\rho(\mathbf{r})] = E_{xc}^{PBE}[\rho] + \frac{1}{n}\left(E_x^{EXX}[\{\phi_i\}] - E_x^{PBE}[\rho(\mathbf{r})]\right), \quad (15)$$

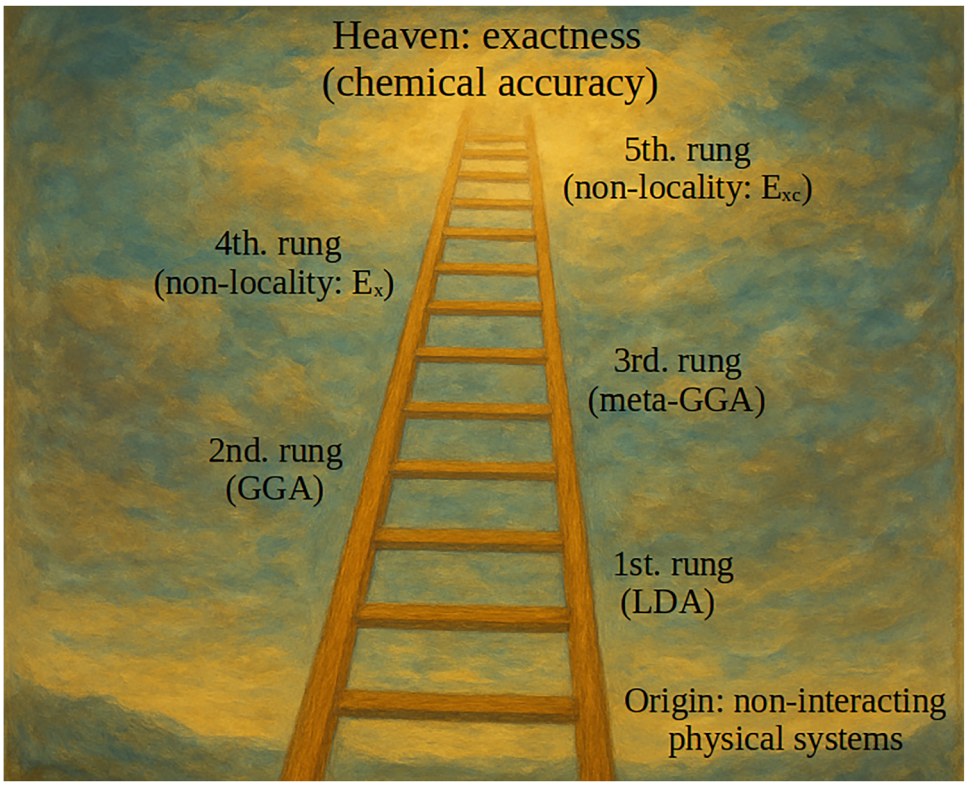

Fig. 1 Illustration of the Jacob's ladder in DFT, where any upper rung of the hierarchy adds more physical information and complexity, thus allowing to reach a more accurate representation of chemical systems.

with $n = 4$ corresponding to the PBE0-1/4 (better known as PBE0 or PBE1PBE[84]) model,[85] which has also been updated more recently to the PBE0-1/3 ($n = 3$) functional.[86] Furthermore, a set of non-empirical double-hybrid expressions is also available, with an additional constraint imposed[87,88] such as $\lambda_x = \lambda_c^2$ (1-DH models) or $\lambda_x = \lambda_c^3$ (LS1-DH models) with examples of the latter such as PBE0-DH,[89] or PBE-QIDH[67] if the ACM is concomitantly employed to derive the final expression. These non-empirical hybridization schemes are shown to be very robust, as it has also been systematically demonstrated,[90] since the results are not largely modified depending on the underlying expressions used for $E_{xc}[\rho(\mathbf{r})]$, or in other words, the linear combination of terms given by eqn (13) plays a more important role for the final accuracy of the results than the functional itself. The hierarchy constituted by *e.g.* PBE, PBE0, and PBE-QIDH could thus serve as an indication of the errors provided by the lower levels, and the physical reasons for that, bracketing conveniently the influence of each of the terms in the final results according again to the Jacob's ladder typical performance.

Actually, the large number of exchange–correlation functionals available, belonging to the categories briefly discussed above, motivated the compilation of the *libxc* library,[19] encompassing all the developments along the last 50 years and rapidly being interfaced to main codes.[20]

3.4 Other variants and corrections

Complementarily to these LDA, GGA, meta-GGA, hybrid or double-hybrid approximations, there have also been transversal developments transferable to all the models, such as the range-separation of the two-electron operator, or the use of a spatial-dependent EXX weight, in the hope of correcting some of the identified deficiencies of DFT. As a matter of illustration, hybrid and double-hybrid functionals still formally miss the correct long-range behaviour of the exchange potential, which should asymptotically go as $\lim_{|\mathbf{r}|\to\infty} v_x = -\frac{1}{|\mathbf{r}|}$. A possible solution to the problem, since $\lambda_x \to 1$ is known to deteriorate the accuracy of the corresponding expressions for *e.g.* thermochemical calculations, is to force that the amount of exact exchange increases as the interaction becomes more long-ranged, by splitting the electron–electron interaction term:

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{\mathrm{erf}(\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} + \frac{1-\mathrm{erf}(\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}, \tag{16}$$

with erf being the error function, and where the first (second) term gives the short-range (long-range) contribution.[91] Another strategy[92] consists on a partition respecting the weight of the hybrid functional at short-range:

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{[\alpha+\beta\,\mathrm{erf}(\omega|\mathbf{r}-\mathbf{r}'|)]}{|\mathbf{r}-\mathbf{r}'|} + \frac{1-[\alpha+\beta\,\mathrm{erf}(\omega|\mathbf{r}-\mathbf{r}'|)]}{|\mathbf{r}-\mathbf{r}'|}, \tag{17}$$

with erf being the error function and α, β, and ω to be defined and/or fitted to reproduce the chosen reference data (highly employed range-separated hybrid functionals are LC-ωPBE,[93]

CAM-B3LYP[92] or the ωB97X family[94,95]). These range-separated functionals are now considered a practical remedy for dealing with the (total or partial) missing of long-range interactions of (mostly) semi-local and hybrid functionals, which is known to largely impact on charge-transfer processes and excitations, among other applications.[96–100]

Note that any range-separated hybrid functional can also be made non-empirical thanks to a recent development disclosing[101,102] the relationship between λ_x and ω, and that range-separated semiempirical (*e.g.* ωB97X-2[103]) and non-empirical (*e.g.* RSX-PBE-QIDH[104]) double-hybrid functionals are also available. Another degree of empiricism is introduced if the value of ω is adapted to any of the system and/or geometries considered, which defines the optimally-tuned range-separated functionals.[105] On the other hand, local hybrid functionals[106] involve position-dependent ratio of EXX and an exchange (GGA or meta-GGA) functional, with the latter models[107] employing $\tau_w(\mathbf{r})$ and the conventional KS kinetic-energy density $\tau(\mathbf{r})$ in the form $g(\mathbf{r}) = \dfrac{\tau_w(\mathbf{r})}{\tau(\mathbf{r})}$.

Particular to the case of double-hybrid functionals, there is also the possibility of allowing the splitting of the PT2 term in Eq. (13) into the same-spin and opposite-spin correlation contributions,[108] scaled by the corresponding coefficients c_{ss} and c_{os}:

$$E_c^{PT2} = c_{ss}E_c^{ss\text{-}PT2} + c_{os}E_c^{os\text{-}PT2}, \quad (18)$$

with again coexisting the two approaches (semiempirical and non-empirical) to find the c_{ss} and c_{os} values: by fitting the c_{ss} and c_{os} values to reference data (*e.g.* the DSD-based family of expressions[109]) or by imposing some universal-like constraints (*e.g.* the SOS1-PBE-QIDH functional[110]) to obtain the best possible values ($c_{ss} = 0$, $c_{os} = 4/3$).

Finally, it is also known that non-covalent (long-range) electronic effects are not completely introduced by any rung of the Jacob's ladder, although this omission is attenuated by the correlation energy introduced by the PT2 term of double-hybrid functionals but it is also largely dependent on the λ_c weight. However, the flexibility (and additivity) of DFT energy terms also allows to add *a posteriori* this energy, once the densities $\rho(\mathbf{r})$ and $\rho(\mathbf{r}')$ are available, thanks to the *e.g.* VV10[111] correlation energy functional, taken here as example of a modern and widely used non-local correlation functional:

$$E_c^{VV10}[\rho(\mathbf{r}),\rho(\mathbf{r}')] = \int d\mathbf{r}\,\rho(\mathbf{r})\left[\beta(b) + \frac{1}{2}\int d\mathbf{r}'\Phi(\mathbf{r},\mathbf{r}')\rho(\mathbf{r}')\right], \quad (19)$$

which correctly decays thanks to the form of the function $\Phi \propto \dfrac{1}{|\mathbf{r}-\mathbf{r}'|^6}$. The parameter b can be optimized to merge this correction to any of the available GGA, meta-GGA, hybrid, and double-hybrid functionals.[112,113] The VV10 expression is part of a broader family of density functional methods for van der Waals interactions (vdW-DF), which can be found as part of the *libvdwxc* library[114] (vdW-DF1, vdW-DF2, vdW-DF-C09, vdW-DF-cx, vdW-DF-optb86, rev-vdW-DF2, vdW-DF-optPBE, vdW-DF-optB88, BEEF-vdW, and mBEEF-vdW) where the non-local correlation forces are captured through a formal analysis of screened response in the electron gas.

Other cost-efficient alternatives also exist, such as the widely used D*n* methods[17,115] (D2, D3, and D4) of S. Grimme *et al.*, the eXchange-hole Dipole Moment (XDM) formalism,[116] the density-dependent energy correction[117] (dDsC), or the tailored DH-SVPD basis set[118] for double-hybrid functionals such as PBE-QIDH. Due to its widespread use to efficiently (but more empirically) incorporate the non-covalent interactions at a negligible computational cost, the expression for the modern versions of those D*n* methods is:

$$E^{D4(BJ)}(R_{AB}) = -\sum_{n=6,8} s_n \sum_{B>A}^{atom\text{-}pairs} \frac{C_n^{AB}}{R_{AB}^n} f^n(R_{AB}), \quad (20)$$

which is built on the pairwise (two-body) interactions between atom centers A and B (R_{AB} is the distance between atoms pairs), C_n^{AB} are the corresponding nth-order dispersion coefficients, and s_6 and s_8 are fitting parameters, with $f^n(R_{AB})$ the (Becke–Johnson) damping function:

$$f^n(R_{AB}) = \frac{R_{AB}^n}{R_{AB}^n + \left(a_1 R_0^{AB} + a_2\right)^n}, \quad (21)$$

based on R_0^{AB} (an atom-pair-specific cutoff radii) and introducing two new parameters a_1 and a_2. A three-body term from the geometrical orientation of all atom triples A, B, C is also added. As it happens with other variants, the set of parameters is fitted for each functional using high-level reference interaction energies of weakly bound systems.

4 Optimal fitting and datasets in DFT

All the advances in DFT over the last decades have been fostered by the extensive benchmarking done in parallel for all the models, either by the own developers or as a part of larger and comparative studies, using extended datasets for that and providing guidance to any user about what functionals are recommended in general or for some specific or tailored application. Computational protocols,[119] tutorial reviews,[120] and best practices guidelines[121,122] for a wise use of DFT are rooted on the robust and general performance of the models on these datasets after the adequate and step-by-step calibration studies done. On the other hand, in the case of semiempirical functionals, these datasets are also very often used for the fitting of the parameters entering into the model. Therefore, the importance of designing and using datasets for past, current, and future studies and developments cannot be underrated. The design of the reference datasets (not specifically for DFT) has reached a high degree of maturity,[123] which also means to include chemically relevant systems of varied sizes and composition together with a wide variety of properties. We here mention some of the most general-purpose employed datasets, such as the GMTKN55,[124] the W4,[125] the MGCDB84,[121] or the ACCDB,[126] together with some other more suited to a particular electronic structure problem, such as the treatment of non-covalent interactions (*e.g.* S66 × 8[127,128]), the treatment of very large systems

(*e.g.* CiM13[129]), the extension to excited-states (*e.g.* QUEST[130]), *etc.* to name just a few of the most known examples nowadays. Regarding the Materials Science field, and thus the large-scale datasets available, we can mention Materials Project,[131,132] JARVIS,[133,134] and Materials Cloud.[135]

Of particular importance for DFT was also the derivation of the Minnesota datasets,[136] created along several years of efforts and refinements and used to develop some of the most used exchange–correlation functionals, and the generalized use of the GMTKN55 dataset in comparison studies, basically covering basic properties and reaction energies for small systems, reaction energies for large systems and isomerization reactions, reaction barrier heights, and intra- and intermolecular non-covalent interactions. Another effort to create meta-data is the NCIAtlas collection[137–141] of non-covalent interactions of several types. Finally, we also emphasize the large-scale dataset Open Molecules 2025 (OMol25),[142] recently launched as an example of the culmination of the work around compilation of datasets. OMol25 is composed of small molecules, biomolecules, metal complexes, and electrolytes, covering 83 elements of the periodic table and systems up to a size of 350 atoms, and includes more than 100 million of DFT calculations at the ωB97M-V/def2-TZVPD level of theory. The motivation behind that huge dataset was to create a comprehensive data for training of machine learning models performing accurately for molecular chemistry.

Note also that these larger and modern datasets are rooted on historical developments mostly done for wavefunction-based theories, such as W4,[143] HEAT,[144] FPA,[145] FPD,[146] Gn theories,[147] ccCA,[148] to name just a few of them, also learning about the minimum level of theory needed to deliver sub-kcal mol^{-1} or even sub-kJ mol^{-1} accuracy for *e.g.* thermochemical calculations. That means that the reference (or nearly-exact) results needed for modern datasets are usually obtained at the (DLPNO-)CCSD(T)/CBS level of theory; *i.e.*, using CCSD(T) (Coupled-Cluster Singles, Doubles, and perturbatively estimated Triples) possibly assisted by DLPNO (Domain Local Pair Natural Orbitals[149]), a technique applied in recent years to extend the size of the system tackled,[150] and with basis sets large enough to reach the Complete Basis Set (CBS) limit. This model chemistry represents the best trade-off between accuracy and computational cost, although higher levels of CC theory can also be employed if those results are believed to be not sufficiently converged.[151]

Despite the emergence of lower cost methods able to deliver reference results, such as those provided by DLPNO, databases of reactions remain poorly diverse and are mostly focused on small molecular systems, not often representative of systems in use for chemical (real-world) applications. Following this line of research, some efforts were recently done to develop chemical databases gathering reactions more and more representative of current trend in chemistry, and containing new systems not yet used to parameterized semi empirical approaches, opening thus the road to unbiased benchmarking.[152] As as example of this, we mention the BH9 database[153] gathering a total of 449 real-life organic chemistry reactions developed with the purpose of filling this gap. It is also worth noting that machine-learning techniques are an open road to improve the diversity of these databases,[154] with some new datasets (*e.g.* CYCLO70) already available built under this principle, even if yet marginally applied for that purpose.

Delving into the use of datasets for the optimization of the parameters entering into widely used models, also recognizing that the following division is necessarily arbitrary, users have qualitatively divided the functionals into marginally parameterized (if the exchange and correlation functionals contains very few parameters, between 5–10, and the reference data were limited but chemically meaningful, as *e.g.* BLYP[10,16] or B3LYP[11]), moderately parameterized (if more parameters are introduced into the models, but not exceeding 10–20, using extended datasets for that, as *e.g.* B97M-V[155]), and heavily or highly parameterized (if a relatively large number of parameters is used, in the range of 30–50, and intensively trained, as *e.g.* M06-L and M06-2X[74]). Note that this rough classification also basically obeys to the temporal timeline followed for the development of DFT models, since the number and quality of the reference data and datasets in the 80s and 90s were necessarily more modest and smaller than in the next decades. Note also that the final number of parameters of an exchange–correlation functional, specially true for those heavily parameterized, is not critically determined by the range-separation (*i.e.*, only ω is added), or the hybrid (*i.e.*, only λ_x is needed) or double-hybrid (*i.e.*, only λ_c is required, and possibly but not always c_{ss} and c_{os}) schemes, although their ultimate performance it is. On the other hand, non-empirical (or minimally empirical) functionals are obviously not parameterized, neither being proned to properties or systems covered as part of the training sets, so that a genuine use of the datasets for assessment purposes is done. A different consideration deserves the coupling of any correction for non-covalent interactions with the exchange–correlation functional, independently of semiempirical or non-empirical models, since the corresponding parameterization cannot be neglected.

Motivated by the difficulties to express the exchange–correlation functional in a compact and mathematical meaningful analytical form, some authors[156–158] searched in the late 90s and early 00s for a more systematic yet optimal functional form based on extensive parameterization of polynomial or other continuum expansions. We show just an example of this approach, which additionally avoided the splitting of the exchange and correlation expressions:

$$E_{xc}[\rho(\mathbf{r})] = \int F_{xc}\left(\rho_\uparrow, \rho_\downarrow, |\nabla\rho_\uparrow|, |\nabla\rho_\downarrow|, \nabla\rho_\uparrow \cdot \nabla\rho_\downarrow\right) d\mathbf{r}, \qquad (22)$$

with $F_{xc} = \sum_{abcd} \omega_{abcd} A^a B^b C^c D^d$, and where ω_{abcd} are parameters, with the remaining functions being $A^a = \rho_\uparrow^a + \rho_\downarrow^a$, $B^b = \left(\frac{\rho_\uparrow - \rho_\downarrow}{\rho_\uparrow + \rho_\downarrow}\right)^{2b}$, $C^c = \frac{|\nabla\rho_\uparrow|^c + |\nabla\rho_\downarrow|^c}{2\left(\rho_\uparrow + \rho_\downarrow\right)^{4c/3}}$, and $D^d = \left(\frac{|\nabla\rho_\uparrow|^2 + |\nabla\rho_\downarrow|^2 - 2\nabla\rho_\uparrow \cdot \nabla\rho_\downarrow}{\left(\rho_\uparrow + \rho_\downarrow\right)^{8/3}}\right)^d$, with in principle no limits imposed for the set of (a, b, c, d) exponents and corresponding parameters. However, this line of

research was more an exploration of the limit of accuracy (and redundancy of parameters and/or training data) that is ultimately possible with GGA or meta-GGA inspired expressions, but it constitutes a first step towards more systematic and modern approaches. Overall, the intense parameterization done in recent times has also disclosed some bottlenecks of semiempirical approaches, such as the bias-variance tradeoff,[159] the numerical instabilities possibly caused by an overfitting,[160] and the importance of the quality of the data used for the fitting.[161]

Another recent approach[195] circumvents the need to develop new exchange–correlation functionals, and their optimization, and simply weights the total (ground-state) energy by functional-specific weights, in the form $E = \sum_i^X \hat{\omega}_i E_i$, with $\hat{\omega}_i$ the normalized weights and E_i the total energies calculated with the X density functionals entering the ensemble. This procedure guarantees the size-consistency of the results and the easy calculation of forces and potentials, and it could be also extended to densities with $\rho(\mathbf{r}) = \sum_i^X \hat{\omega}_i \rho_i(\mathbf{r})$. The GMTKN55 data are used to train the ensemble with up to 77 functionals belonging to 2nd–5th rungs of Jacob's ladder included into the regression procedure, to reach an optimal functional (DENS24) achieving an error as small as 1.6 kcal mol^{-1} for that GMTKN55 dataset. The transferability of the ensemble and the basis set dependence of the model was also tested, as well as ways to reduce the associated computational cost to get each of the E_i energies.

5 Deep-learned DFT models

We have summarized in previous sections how the conscious and meticulous work for many decades is behind the current success of DFT, regardless some generalized flaws of the theory[162] such as the (many-electron) self-interaction error, or some limits imposed by the search and optimization of the parameter space. On the other hand, the recent rise of machine learning for the design of exchange–correlation functionals[163] has consequently led to a new paradigm of work, which we will also try to summarize here with three examples of functionals fully machine-learned, but rooted on the prescriptions and knowledge so far achieved and belonging to the 3rd–4th rungs of the Jacob's ladder: the DeepMind (DM21),[164] the "adaptative" PBE0 (aPBE0-ML)[165] and the recently presented Skala model.[166] Interestingly, two of these models (DM21 and Skala) are developed by private companies, Google Deepmind and Microsoft (Microsoft Research – AI for Science, and Microsoft Quantum), respectively. A common point of all these ML-based approaches is the interest to replace traditional, empirically or physically motivated approximations, with a neural-network mapping from electron density (and related magnitudes) to exchange–correlation energies. Note also that prior to DM21, there was some attempts to develop machine-learned exchange–correlation functionals (*e.g.* ML-ωPBE[167]) but not for such a general purpose as those developed later. Additionally, the use of neural network had not demonstrated any comparable or superior performance with respect to traditional (or analytical, from a mathematical point of view) functionals, as those summarized before, before the level of accuracy achieved by DM21, aPBE0-ML, or Skala functionals. For other pioneering approaches, we also mention OrbNet,[168] NeuralXC,[169] DeePHF,[170] or CF22D.[171]

5.1 The deepmind functional

In the DM21 model[172] the proportion of exact-exchange varies locally in space, thus belonging to the family of local hybrid functionals, and also incorporates their range-separated values, which has the advantage of (formally) removing the delocalization error.[173] The explicit form of the DM21 functional integrates local energies calculated by a MultiLayer Perceptron (MLP) which took as input both local and non-local features of the occupied KS orbitals:

$$E_{\mathrm{xc}}^{\mathrm{MLP}}[\rho(\mathbf{r})] = \int f_\theta(x(\mathbf{r})) \cdot \begin{bmatrix} e_x^{\mathrm{LDA}}(\mathbf{r}) \\ e_x^{\mathrm{EXX}}(\mathbf{r}) \\ e_x^{\omega\mathrm{EXX}}(\mathbf{r}) \end{bmatrix} \mathrm{d}\mathbf{r}, \tag{23}$$

where $f_\theta(x(\mathbf{r}))$ represents a single set of weights, θ, accepting a vector, $x(\mathbf{r})$, at each grid point $\mathbf{r}$, and returning a vector of 3 enhancement factors before computation of the integral. The MLP part is complemented by the typical D3(BJ) term, and the DM21 expression is just made $E_{\mathrm{xc}}^{\mathrm{DM21}}[\rho(\mathbf{r})] = E_{\mathrm{xc}}^{\mathrm{MLP}}[\rho(\mathbf{r})] + E_{\mathrm{D3(BJ)}}$. The $x(\mathbf{r})$ values supplied at each point of the grid are computed from a spin-indexed ($\sigma \in \{\uparrow,\downarrow\}$) set of typical DFT variables, based on the density matrix, D_{ij}^σ, the density itself, $\rho^\sigma(\mathbf{r}) = \sum_{ij} D_{ij}^\sigma \phi_i(\mathbf{r})\phi_j(\mathbf{r})$, the square norm of the gradient of the spin densities and of the total density, $|\nabla\rho^\sigma(\mathbf{r})|^2$ and $|\nabla\rho(\mathbf{r})|^2$, and the kinetic energy density, $\tau(\mathbf{r}) = \frac{1}{2} D_{ij}^\sigma |\nabla\phi_i(\mathbf{r})|^2$. The local energies are:

$$e_{\mathrm{x}}^{\mathrm{LDA}}(\mathbf{r}) = -C_{\mathrm{x}}[\rho_\uparrow(\mathbf{r}) + \rho_\downarrow(\mathbf{r})]^{4/3} \tag{24}$$

$$e_x^{\omega\mathrm{EXX}}(\mathbf{r}) = -\frac{1}{2}\sum_{ijkl} D_{ik}^\sigma D_{jl}^\sigma \int \phi_i^\star(\mathbf{r})\phi_j^\star(\mathbf{r}') \frac{\mathrm{erf}(\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} \phi_j(\mathbf{r})\phi_i(\mathbf{r}')\mathrm{d}\mathbf{r}', \tag{25}$$

with ω = 0.4 chosen based on validation studies according to the own developers; $e_{\mathrm{x}}^{\mathrm{EXX}}(\mathbf{r})$ is calculated just setting $\omega = 0$ in the above equation. Note that all the ingredients entering into the model are part of the main features of the DFT advances gained in the last decades. The model was trained on extensive atomic and molecular datasets (more than 1000 points), but also considering exact (integer-charge) mathematical constraints, and fractional-charge and fractional-spin information (more than 1000 points), complemented by a D3(BJ) two-body correction as it is usually done in traditional models. Note also that the hard-to-control piecewise linear variation of the total energy with respect to fractional electron number was forced to be part of those constraints. Despite some initial criticism of the model,[174,175] DM21 was found to be more accurate than a set of (more standard) functionals such as revPBE-D3(BJ), MN15-L-D3(0), SCAN-D3(0), B3LYP-D3(BJ), ωB97X-V, M06-2X-D3(0), and PW6B95-D3(0) by looking at the WTMAD-2 error of the GMTKN55 dataset, and only inferior to the DSD-PBEP86-

D3(BJ) double-hybrid functional, among the traditional functionals considered. One notable performance of the DM21 model was the correct dissociation limit of challenging systems such as H_2, H_2^+, and N_2, the good performance for the SIE4x4 dataset, as well as for other complex situations like a charged adenine–thymine base pair, a compressed hydrogen H_n chain, and diradical transition states. The developers of DM21 caution against its application to atoms heavier than Kr because the training set does not include heavy elements or pseudopotentials.

DM21 has recently been applied to transition metal chemistry,[176] which can be seen as a proof-of-concept for its transferability since DM21 was exclusively trained on main-group chemistry as said before; more particularly, it was applied to the TMC151 dataset, featuring transition metal dimer dissociation energies, metal–organic reaction energies, and barriers of complexes of second- and third-row transition metals. The authors concluded that despite a similar performance to B3LYP, the model faced convergence issues most of the times, thus complicating its self-consistent implementation; however, the use of B3LYP densities instead (DM21@B3LYP) improved the results, although at the price of a higher computational cost comparable even to *i.e.* B2-PLYP, a double-hybrid exchange–correlation functional. One possible solution for this, already pointed out by the authors of the study, would be to incorporate transition metal reactions into the machine-learned functional, with the due care to not deteriorate the performance of main-group chemistry, borrowing again one of the strategies followed historically, that is, extending and extending the size and diversity of the datasets used for the development and assessment of the models.

DM21 has also been applied to neutral and charged water clusters,[177] with a varying degree of accuracy as a function of the water cluster size: $(H_2O)_n$ (n = 2–10), $H_3O^+\cdot(H_2O)_n$ and $OH^-\cdot(H_2O)_n$ clusters with n = 2–6. The functional performed similarly to one the best range-separated meta-GGA functionals (ωB97M-V) but the errors for the normalized interaction energies (*i.e.*, interaction energies divided by n) increased with the cluster size, thus being size-dependent. The authors also introduced a many-body potential (MB-DM21) to correct for missing many-body non-covalent interactions, and applied the corrected model to calculate some properties of liquid water with partial success, recommending to improve the functional form and physical content of machine-learned density functionals. Besides energy calculations, DM21 has also been extended to geometry optimizations and potential energy surfaces (PES). As another example of application of DM21, we mention the DM21-based calculations[178] of some PES upon variation of bond angles (*e.g.*, H–O–H in water) and bond lengths (*e.g.*, C–H or C–C bond length in common organic molecules such as C_4H_8). The authors found a very good agreement between CCSD(T) and DM21-based results for all the PES tackled, including the stretched H_2 and H_2^+ PES at long H–H distances, comparable or better than results obtained at the hybrid PW6B95/cc-pVnZ level.

5.2 The Skala functional

The first step of the developers of the Skala functional[166] was the compilation of an unprecedented dataset of physically meaningful values (149 079) to train the model, including a set of Total Atomization Energies (TAE) as the largest subset, which is a property known to be difficult to calculate and has always been part of both early and modern datasets. The TAE dataset is composed by 78 650 values with an expected accuracy of ± 1 kcal mol^{-1} relative to experiments, and achieved by a suited wavefunction-based protocol such as CCSD(T)/CBS. The TAE were calculated for molecules containing up to five non-H atoms (Li–F, Na–Cl) and the total number of points was complemented by other datasets. The latter included atomic (total energies, electron affinities, and ionization potentials), molecular (conformational energies, proton affinities, and ionization potentials), reaction kinetics, and both intra- and intermolecular non-covalent interactions from the NCIAtlas collection (D442x10, SH250x10, R739x5, and HB300SPXx10).

The chosen form of the Skala functional was also motivated by satisfying energetically relevant constraints such as the high-density uniform coordinate scaling, the Lieb-Oxford lower bound, and the LDA limit if the enhancement factor is made $f_\theta(x(\mathbf{r})) = 1$, which can be achieved by the form:

$$E_{\mathrm{xc}}^{\theta}[\rho(\mathbf{r})] = -\frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3}\int f_\theta(x(\mathbf{r}))\left[\rho_\uparrow(\mathbf{r}) + \rho_\downarrow(\mathbf{r})\right]^{4/3}\mathrm{d}\mathbf{r}, \quad (26)$$

with $f_\theta(x(\mathbf{r}))$ a learnable enhancement factor, whose modelling is done by a deep neural network taken as input a set of semi-local, density-dependent features $x(\mathbf{r})$, from a meta-GGA model. Actually, those features considered are again the density, the square norm of the gradient of the density and of the total density, and the kinetic energy density. The model was also complemented by a standard D3(BJ) correction. The authors chosen to work with discretized models, by evaluating the expression on a set of points $\{r_i \in \mathbf{R}^3,\ i = 1,\ldots,M\}$:

$$E_{\mathrm{xc}}^{\theta}[\rho(\mathbf{r})] \approx -\frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3}\sum_i^M f_\theta(x(r_i))\left[\rho_\uparrow(r_i) + \rho_\downarrow(r_i)\right]^{4/3} w_i, \quad (27)$$

with the associated weights $w_i \in \mathbf{R}$ for those points r_i. The set of $x(r_i)$ features are $\rho^\uparrow(r_i)$, $\rho^\downarrow(r_i)$, $|\nabla\rho^\uparrow(r_i)|^2$, $|\nabla\rho^\downarrow(r_i)|^2$, $|\nabla\rho^\uparrow(r_i) + \nabla\rho^\downarrow(r_i)|^2$, $\tau^\uparrow(r_i)$, $\tau^\downarrow(r_i)$, with the input thus of the neural network being a tensor $\mathbf{R}^{M+7}$. These spatial fields are projected onto grids in a way compatible with typical DFT integration schemes. The model learns non-local correlation effects directly from these fields (*i.e.*, message-passing mechanisms to encode long-range electron correlation) thus improving the description given by semi-local functionals. Among other interesting features of the Skala model, its accuracy was demonstrated to improve systematically with adding more training data by including step-by-step more datasets into the model. Concerning the computational cost of Skala, a pair of implementations were tested: a GPU-based version integrated into Accelerated DFT,[179] and a CPU-based version implemented in PySCF.[180] Skala showed to be very competitive particularly for the GPU-based version, with a cost comparable to that of r^2SCAN and ten

times lower than that of M06-2X. The CPU-based implementation was slightly less successful but still competitive with those of r^2SCAN, B97M-V, M06-2X, and ωB97M-V.

5.2.1 Comparison between DM21 and Skala for some established datasets and subsets. Compared to the WTMAD-2 value dropped by DM21 for the GMTKN55 dataset, 3.97 kcal mol^{-1}, Skala gave a very close value of 3.89 kcal mol^{-1} despite an almost two order of magnitude increase in the number of data points used to train the model. To put these numbers in a due context, the lowest achieved WTMAD-2 value nowadays is obtained by the DH23 model,[181,182] a 12-parameter double-hybrid density functional leading to an error or 1.7 kcal mol^{-1}, with some non-empirical double-hybrid density functionals[183,184] also giving errors as low as 3–5 kcal mol^{-1}. Skala performed remarkably accurate on the W4–17 dataset[125] of TAE, with a mean absolute error of 1 kcal mol^{-1} for the 200 values included, roughly speaking, situating the model as the best performer for this dataset among all the functionals considered (revPBE, r^2SCAN, B97M-V, B3LYP, M06-2X, ωB97X-V, and ωB97M-V). The W4–17 was part of the training of DM21, and it was thus not analyzed separately in the corresponding publication.

Looking more in detail at the performance of Skala for the different subsets of the families of the GMTKN55 dataset, the largest error was found for the SIE4x4 (self-interaction-error related systems, with 13.6 kcal mol^{-1}) subset, the DIE60 (relative energies between C_{60} isomers, 8.54 kcal mol^{-1}), BHPERI (barrier heights of pericyclic reactions, 3.13 kcal mol^{-1}), WATER27 (Binding energies in $(H_2O)_n$, $H^+\cdot(H_2O)_n$ and $OH^-\cdot(H_2O)_n$, 2.47 kcal mol^{-1}), and IDISP (intramolecular dispersion interactions, 3.46 kcal mol^{-1}). On the other hand, the errors given by DM21 for those subsets are 4.92, 11.35, 0.77, 1.96, and 1.37 kcal mol^{-1}, respectively, evidencing the large discrepancies between both models when looking at their performances for the individual subsets. Actually, the largest error dropped by DM21 were found for the SIE4x4 (4.92 kcal mol^{-1}) and DIE60 (11.35 kcal mol^{-1}) for the families of basic properties and reaction energies for small systems, and reaction energies for large systems and isomerization reactions, respectively. For the family of reaction barrier heights, the DM21 largest error was found for the BH76 dataset (barrier heights of various reaction types, 2.08 kcal mol^{-1}). Considering the intra- and intermolecular families of non-covalent interactions, the DM21 largest errors are again found for the WATER27 (1.96 kcal mol^{-1}) and IDISP (1.37 kcal mol^{-1}) datasets.

Additionally, Skala was also tested[166] for equilibrium geometries with a deviation of 0.012 Å for the bond lengths of the CCse21 dataset,[185] a typical deviation by semi-local functionals[186] and larger than the values (mostly around 0.04–0.05 Å) provided by the rest of the functionals (r^2SCAN, B97M-V, B3LYP, M06-2X, ωB97X-V, and ωB97M-V) chosen with the exception of revPBE which also leads to a deviation of 0.012 Å. By perusing now[187] the LMGB35 dataset[17] of light atoms main group bond lengths, Skala gave a mean absolute error of 0.014 Å, compared with 0.006 Å by DM21 and thus competitive with PBE0 results. However, except BO, F_2^+, N_2^+, NF, NH^+, O_2^+, and OH^+, those molecules belonging to LMGB35 are also included into the W4–17 used for training the DM21. Skala was also applied to the HMGB11 dataset of heavy atoms main group bond length, providing larger deviations (of 0.03 Å) than for light atoms, but comparable to those achieved by revPBE or ωB97X-V functionals.

Finally, taken from the original study,[166] the authors also compared Skala and DM21 cost. We remind that the former deep-learned form (Skala) avoids a position-dependent fraction of exact-exchange (DM21) with a clear benefit for the (CPU-based) computational cost which is substantially lower by two orders of magnitude. For this comparison, the authors used a sample of molecules of varying sizes, from 50 to 900 atoms and from 1000 to 20 000 orbitals, roughly speaking, taken also from established datasets such as S30L[188] HS13L[189] and NCI16L.[190]

5.3 The adaptative PBE0 functional

The "adaptative" PBE0 functional (aPBE0) was also recently developed[165] by von Lilienfeld *et al.*, based on the hypothesis that a system-dependent optimal λ_x^{opt} value exists, which is actually a system-specific scalar label leading to vanishing error for the model if compared with reference of high-level electronic structures methods, in the form:

$$E_{xc}[\rho(\mathbf{r})] = \lambda_x^{opt}\{(Z_I, \mathbf{R}_I)_I, S\}E_x^{EXX}[\{\phi_i\}] + (1 - \lambda_x^{opt}\{(Z_I, \mathbf{R}_I)_I, S\})E_x[\rho(\mathbf{r})] + E_c[\rho(\mathbf{r})], \quad (28)$$

with Z_I and $\mathbf{R}_I$ denoting the nuclear charge and coordinates, respectively, with S specifying the electronic spin state; or, in other words, the optimal value depends on both the external potential and the electronic spin state. In a further step, by setting $\lambda_x^{opt} = 1$ the correlation term can also be made system-dependent to find the corresponding γ_c^{opt} optimal values, in the form $\gamma_c^{opt}\{(Z_I, \mathbf{R}_I)_I, S\}E_c[\rho(\mathbf{r})]$. The authors first calculated at the CCSD(T) level the TAE of 1169 small molecular fragments (containing 1–5 heavy atoms) reported in ref. 191 to find the optimal λ_x^{opt} values for those systems. The normal distribution plot of the individual 1169 values was centered at $\lambda_x = 0.42$ approximately. The mean absolute error was impressively reduced from 3.46 (PBE0) to 0.02 (aPBE0) kcal mol^{-1} upon this adaptation of values. To assess the transferability of the model, they calculate TAE for a set of 50 new molecules, largely reducing again the mean absolute error with respect to nearly-exact reference results from 4.68 (PBE0) to 0.02 (aPBE0) kcal mol^{-1}. The procedure was also applied to the spin gap of a large set of carbenes ($\approx$3000) in their singlet and triplet states and taken from the QMSpin dataset,[192] with excellent results too going from a mean absolute error of 10.2 (PBE0) to 0.1 (aPBE0) kcal mol^{-1}. Note also that $\lambda_x^{opt,triplet} = 1.102\lambda_x^{opt,singlet} - 0.118$. The results by aPBE0, be for TAE or for the singlet–triplet spin gaps, were always better than those provided by semi-local (PBE, BLYP or r^2SCAN), hybrid (PBE0, B3LYP, and M06-2X), and ωB97XD functionals.

After having confirmed the hypothesis that a significant improvement of the results can be obtained by a system-dependent mixing of semi-local and exact-exchange, the authors explored the possibility of deep-learned λ_x^{opt} values: using the generated λ_x^{opt} values as training data, they use a predictive model to infer values for any other system of interest,

naming this variant as aPBE0-ML consequently. They predicted on-the-fly the λ_x^{opt} values for the revised QM9 dataset[193] (revQM9) consisting on 130 000 molecules, roughly speaking. The deep-learned λ_x^{opt} values followed again a Gaussian-like distribution, peaking at $\lambda_x = 0.44$, with a standard deviation of 0.037, and thus very close to the $\lambda_x = 0.42$ value disclosed before. The TAE of the reduced set of 50 molecules showed now an error of 1.32 kcal mol^{-1} by aPBE0-ML, and thus slightly larger than that by aPBE0 itself. HOMO–LUMO gaps errors, with respect to nearly-exact GW values for 100 organic molecules from the QM7b dataset,[194] are also reduced from 3.52 (PBE0) to 0.86 (aPBE0-ML) eV, with aPBE0-ML again on top-tier of the models tested (PBE, BLYP, r^2SCAN, PBE0, B3LYP, M06-2X, LC-ωPBE, and ωB97XD) with the exception of M06-2X for which a slightly lower error (0.67 eV) was found.

6 Conclusions

The field of DFT has been living a golden age for decades, as it is clearly evidenced by the large volume of scientific publications, advances, and exchange–correlation functionals expressions implemented in codes and exploited in endless applications. Generally speaking, the main advances and developments of DFT has gone in parallel with those of modern Science. For instance, the deployment of large computing facilities and hardware/software improvements has allowed the application of DFT to systems of all sizes and chemical composition, thus going a step further with respect to earlier studies. However, other issues have also arised. Among them, we mention some of the most critical ones, such as the increasing cost of the calculations with rungs of Jacob's ladder and how to alleviate it, the dependence and nature of the errors with respect to the system size and composition, the different performance (and the reasons for that) between the class of exchange–correlation functionals employed, *etc.* That continuous validation and assessment of models was also possible by preserving the accuracy of wavefunction-based results, which are very often taken as reference, at all size scales, thanks to the introduction of extended and chemically diverse datasets, from both the composition of the systems and the properties included in them.

This rich and very fruitful context is now also experiencing another step forward due to deep-learned expressions for the exchange–correlation functional. However, the underlying hypothesis and expressions are deeply rooted on the lessons achieved so far, as well as on the key role played by the ingredients defining the exchange–correlation integrand. Furthermore, in a kind of virtuous circle, the size and availability of datasets has also been invigorated by the need to train those deep-learned expressions, which can also be seen as a source of opportunities for assessing and comparing more traditional expressions. With the due efforts in both directions, the field of DFT is expected to still grow strongly in the next decades, full of opportunities and success.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data that supports the findings of this study are available within the article or are available from the corresponding authors upon reasonable request.

Acknowledgements

The work in Alicante is supported by grant PID2023-152372NB-I00 funded by MICIU/AEI/10.13039/501100011033. E. B. thanks ANR (Agence Nationale de la Recherche) and CGI (Commissariat à l'Investissement d'Avenir) for their financial support to this work through Labex SEAM (Science and Engineering for Advanced Materials and devices), Grant no. ANR-10-LABX-096, ANR-18-IDEX-0001, and ANR-21-CE29-0003. Funded by the European Union (ERC, MaMa, no. 101097351). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

1 K. Giesbertz, K. Pernal, O. Gritsenko and E. Baerends, Excitation energies with time-dependent density matrix functional theory: Singlet two-electron systems, *J. Chem. Phys.*, 2009, **130**, 114104.
2 M. E. Casida and M. Huix-Rotllant, Progress in time-dependent density-functional theory, *Annu. Rev. Phys. Chem.*, 2012, **63**, 287–323.
3 A. D. Laurent and D. Jacquemin, TD-DFT benchmarks: a review, *Int. J. Quantum Chem.*, 2013, **113**, 2019–2039.
4 C. Adamo and D. Jacquemin, The calculations of excited-state properties with Time-Dependent Density Functional Theory, *Chem. Soc. Rev.*, 2013, **42**, 845–856.
5 X. Blase, I. Duchemin and D. Jacquemin, The Bethe-Salpeter equation in chemistry: relations with TD-DFT, applications and challenges, *Chem. Soc. Rev.*, 2018, **47**, 1022–1043.
6 E. Bremond, M. Savarese, C. Adamo and D. Jacquemin, Accuracy of TD-DFT geometries: a fresh look, *J. Chem. Theory Comput.*, 2018, **14**, 3715–3727.
7 S. Ghosh, P. Verma, C. J. Cramer, L. Gagliardi and D. G. Truhlar, Combining wave function methods with density functional theory for excited states, *Chem. Rev.*, 2018, **118**, 7249–7292.
8 R. Van Noorden, B. Maher and R. Nuzzo, The top 100 papers, *Nat. News*, 2014, **514**, 550.
9 R. Van Noorden, These are the most-cited research papers of all time, *Nature*, 2025, **640**, 591.

10 C. Lee, W. Yang and R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785.
11 A. D. Becke, Density-Functional Thermochemistry. III. The Role of Exact Exchange, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
12 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, 1996, **77**, 3865.
13 G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 11169.
14 P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Phys. Rev.*, 1964, **136**, B864.
15 W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.*, 1965, **140**, A1133.
16 A. D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098.
17 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 2010, **132**, 154104.
18 W. Kohn, Nobel Lecture: Electronic structure of matter—wave functions and density functionals, *Rev. Mod. Phys.*, 1999, **71**, 1253.
19 M. A. Marques, M. J. Oliveira and T. Burnus, Libxc: A library of exchange and correlation functionals for density functional theory, *Comput. Phys. Commun.*, 2012, **183**, 2272–2281.
20 S. Lehtola, C. Steigemann, M. J. Oliveira and M. A. Marques, Recent developments in libxc – A comprehensive library of functionals for density functional theory, *SoftwareX*, 2018, **7**, 1–5.
21 S. Lehtola and A. J. Karttunen, Free and open source software for computational chemistry education, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1610.
22 M. Dumaz, R. Boucher, M. A. Marques and A. H. Romero, Authorship and citation cultural nature in Density Functional Theory from solid state computational packages, *Scientometrics*, 2021, **126**, 6681–6695.
23 R. Haunschild, A. Barth and W. Marx, Evolution of DFT studies in view of a scientometric perspective, *J. Cheminf.*, 2016, **8**, 1–12.
24 T.-S. Lee, J. P. Lewis and W. Yang, Linear-scaling quantum mechanical calculations of biological molecules: The divide-and-conquer approach, *Comput. Mater. Sci.*, 1998, **12**, 259–277.
25 E. Artacho, D. Sánchez-Portal, P. Ordejón, A. Garcia and J. M. Soler, Linear-scaling ab initio calculations for large and complex systems, *Phys. Status Solidi B*, 1999, **215**, 809–817.
26 F. Neese, An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix, *J. Comput. Chem.*, 2003, **24**, 1740–1747.
27 H. Neugebauer, P. Pinski, S. Grimme, F. Neese and M. Bursch, Assessment of DLPNO-MP2 approximations in double-hybrid DFT, *J. Chem. Theory Comput.*, 2023, **19**, 7695–7703.
28 Y. Zhang and W. Yang, A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons, *J. Chem. Phys.*, 1998, **109**, 2604–2608.
29 P. Mori-Sánchez, A. J. Cohen and W. Yang, Many-electron self-interaction error in approximate density functionals, *J. Chem. Phys.*, 2006, **125**, 201102.
30 J. P. Perdew and A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1981, **23**, 5048.
31 V. Polo, E. Kraka and D. Cremer, Electron correlation and the self-interaction error of density functional theory, *Mol. Phys.*, 2002, **100**, 1771–1790.
32 M. Lundberg and P. E. Siegbahn, Quantifying the effects of the self-interaction error in DFT: When do the delocalized states appear?, *J. Chem. Phys.*, 2005, **122**, 224103.
33 J. Gräfenstein and D. Cremer, The self-interaction error and the description of non-dynamic electron correlation in density functional theory, *Theor. Chem. Acc.*, 2009, **123**, 171–182.
34 T. Schmidt and S. Kümmel, One-and many-electron self-interaction error in local and global hybrid functionals, *Phys. Rev. B*, 2016, **93**, 165120.
35 D. R. Lonsdale and L. Goerigk, The one-electron self-interaction error in 74 density functional approximations: a case study on hydrogenic mono-and dinuclear systems, *Phys. Chem. Chem. Phys.*, 2020, **22**, 15805–15830.
36 S. Vuckovic, S. Song, J. Kozlowski, E. Sim and K. Burke, Density functional analysis: The theory of density-corrected DFT, *J. Chem. Theory Comput.*, 2019, **15**, 6636–6646.
37 E. Sim, S. Song, S. Vuckovic and K. Burke, Improving results by improving densities: Density-corrected density functional theory, *J. Am. Chem. Soc.*, 2022, **144**, 6625–6639.
38 L. H. Thomas, The calculation of atomic fields, *Math. Proc. Cambridge Philos. Soc.*, 1927, **23**, 542–548.
39 E. Fermi, Un metodo statistico per la determinazione di alcune priorieta dell'atome, *Rend. Accad. Naz. Lincei*, 1927, **6**, 32.
40 P. A. Dirac, Note on exchange phenomena in the Thomas atom, *Math. Proc. Cambridge Philos. Soc.*, 1930, **26**, 376–385.
41 E. Wigner, Effects of the electron interaction on the energy levels of electrons in metals, *Trans. Faraday Soc.*, 1938, **34**, 678–685.
42 S. H. Vosko, L. Wilk and M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, *Can. J. Phys.*, 1980, **58**, 1200–1211.
43 J. P. Perdew, Generalized gradient approximations for exchange and correlation: A look backward and forward, *Phys. B*, 1991, **172**, 1–6.
44 Y. Zhao and D. G. Truhlar, Construction of a generalized gradient approximation by restoring the density-gradient expansion and enforcing a tight Lieb-Oxford bound, *J. Chem. Phys.*, 2008, **128**, 184109.
45 A. Ruzsinszky, G. I. Csonka and G. E. Scuseria, Regularized gradient expansion for atoms, molecules, and solids, *J. Chem. Theory Comput.*, 2009, **5**, 763–769.

46 R. Peverati, Y. Zhao and D. G. Truhlar, Generalized gradient approximation that recovers the second-order density-gradient expansion with optimized across-the-board performance, *J. Phys. Chem. Lett.*, 2011, **2**, 1991–1997.
47 E. Engel and S. Vosko, Fourth-order gradient corrections to the exchange-only energy functional: Importance of $\nabla^2\rho$ contributions, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **50**, 10498.
48 J. P. Perdew, S. Kurth, A. Zupan and P. Blaha, Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation, *Phys. Rev. Lett.*, 1999, **82**, 2544.
49 J. P. Perdew and K. Schmidt, Jacob's ladder of density functional approximations for the exchange-correlation energy, *AIP Conf. Proc.*, 2001, **577**, 1–20.
50 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids, *Phys. Rev. Lett.*, 2003, **91**, 146401.
51 J. Sun, A. Ruzsinszky and J. P. Perdew, Strongly constrained and appropriately normed semilocal density functional, *Phys. Rev. Lett.*, 2015, **115**, 036402.
52 A. P. Bartók and J. R. Yates, Regularized SCAN functional, *J. Chem. Phys.*, 2019, **150**, 161101.
53 J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew and J. Sun, Accurate and numerically efficient r^2SCAN meta-generalized gradient approximation, *J. Phys. Chem. Lett.*, 2020, **11**, 8208–8215.
54 A. J. Garza and G. E. Scuseria, Predicting band gaps with hybrid density functionals, *J. Phys. Chem. Lett.*, 2016, **7**, 4165–4170.
55 J. Muscat, A. Wander and N. Harrison, On the prediction of band gaps from hybrid functional theory, *Chem. Phys. Lett.*, 2001, **342**, 397–401.
56 H. Xiao, J. Tahir-Kheli and W. A. Goddard III, Accurate band gaps for semiconductors from density functional theory, *J. Phys. Chem. Lett.*, 2011, **2**, 212–217.
57 Y. Zhao and D. G. Truhlar, Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: the MPW1B95 and MPWB1K models and comparative assessments for hydrogen bonding and van der Waals interactions, *J. Phys. Chem. A*, 2004, **108**, 6908–6918.
58 L. Goerigk and S. Grimme, A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6670–6688.
59 S. Mandal and R. Pati, Mechanism behind the switching of current induced by a gate field in a semiconducting nanowire junction, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **84**, 115306.
60 A. Tal, P. Liu, G. Kresse and A. Pasquarello, Accurate optical spectra through time-dependent density functional theory based on screening-dependent hybrid functionals, *Phys. Rev. Res.*, 2020, **2**, 032019.
61 W. Sun, T. van der Heide, V.-Q. Vuong, T. Frauenheim, M. A. Sentef, B. Aradi and C. R. Lien-Medrano, Hybrid Functional DFTB Parametrizations for Modeling Organic Photovoltaic Systems, *J. Chem. Theory Comput.*, 2025, **21**, 5103–5117.
62 G. I. Csonka, J. P. Perdew and A. Ruzsinszky, Global hybrid functionals: A look at the engine under the hood, *J. Chem. Theory Comput.*, 2010, **6**, 3688–3703.
63 W. Yang, Generalized adiabatic connection in density functional theory, *J. Chem. Phys.*, 1998, **109**, 10107–10110.
64 M. Levy, Density-functional exchange correlation through coordinate scaling in adiabatic connection and correlation hole, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1991, **43**, 4637.
65 M. Ernzerhof, Construction of the adiabatic connection, *Chem. Phys. Lett.*, 1996, **263**, 499–506.
66 A. D. Becke, A new mixing of Hartree-Fock and local density-functional theories, *J. Chem. Phys.*, 1993, **98**, 1372–1377.
67 É. Brémond, J. C. Sancho-Garca, Á. J. Pérez-Jiménez and C. Adamo, Communication: double-hybrid functionals from adiabatic-connection: the QIDH model, *J. Chem. Phys.*, 2014, **141**, 031101.
68 A. J. Cohen, P. Mori-Sánchez and W. Yang, Assessment and formal properties of exchange-correlation functionals constructed from the adiabatic connection, *J. Chem. Phys.*, 2007, **127**, 034101.
69 D. Khan, SCAN based non-linear double hybrid density functional, *J. Chem. Phys.*, 2025, **163**, 144115.
70 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
71 V. Barone and C. Adamo, Theoretical study of direct and water-assisted isomerization of formaldehyde radical cation. A comparison between density functional and post-Hartree-Fock approaches, *Chem. Phys. Lett.*, 1994, **224**, 432–438.
72 C. Adamo and V. Barone, Toward reliable adiabatic connection models free from adjustable parameters, *Chem. Phys. Lett.*, 1997, **274**, 242–250.
73 N. C. Handy and A. J. Cohen, Left-right correlation energy, *Mol. Phys.*, 2001, **99**, 403–412.
74 Y. Zhao and D. G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
75 S. Grimme, Semiempirical hybrid density functional with perturbative second-order correlation, *J. Chem. Phys.*, 2006, **124**, 034108.
76 F. Tran, J. Stelzl and P. Blaha, Rungs 1 to 4 of DFT Jacob's ladder: Extensive test on the lattice constant, bulk modulus, and cohesive energy of solids, *J. Chem. Phys.*, 2016, **144**, 204120.
77 A. Karton, How reliable is DFT in predicting relative energies of polycyclic aromatic hydrocarbon isomers? comparison of functionals from different rungs of jacob's ladder, *J. Comput. Chem.*, 2017, **38**, 370–382.

78 I. Y. Zhang and X. Xu, On the top rung of Jacob's ladder of density functional theory: Toward resolving the dilemma of SIE and NCE, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2021, **11**, e1490.
79 A. Karton, Big data benchmarking: how do DFT methods across the rungs of Jacob's ladder perform for a dataset of 122k CCSD(T) total atomization energies?, *Phys. Chem. Chem. Phys.*, 2024, **26**, 14594–14606.
80 J. Simmie and K. Somers, Snakes on the rungs of Jacob's ladder: anomalous vibrational spectra from double-hybrid DFT methods, *J. Phys. Chem. A*, 2020, **124**, 6899–6902.
81 A. Karton, S. L. Waite and A. J. Page, Performance of DFT for C_{60} isomerization energies: a noticeable exception to Jacob's ladder, *J. Phys. Chem. A*, 2018, **123**, 257–266.
82 S. Mandal, K. Haule, K. M. Rabe and D. Vanderbilt, Electronic correlation in nearly free electron metals with beyond-DFT methods, *npj Comput. Mater.*, 2022, **8**, 181.
83 J. P. Perdew, M. Ernzerhof and K. Burke, Rationale for mixing exact exchange with density functional approximations, *J. Chem. Phys.*, 1996, **105**, 9982–9985.
84 M. Ernzerhof and G. E. Scuseria, Assessment of the Perdew-Burke-Ernzerhof exchange-correlation functional, *J. Chem. Phys.*, 1999, **110**, 5029–5036.
85 C. Adamo and V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
86 C. A. Guido, E. Brémond, C. Adamo and P. Cortona, Communication: One third: A new recipe for the PBE0 paradigm, *J. Chem. Phys.*, 2013, **138**, 021104.
87 J. Toulouse, K. Sharkas, E. Brémond and C. Adamo, Communication: Rationale for a new class of double-hybrid approximations in density-functional theory, *J. Chem. Phys.*, 2011, **135**, 101102.
88 K. Sharkas, J. Toulouse and A. Savin, Double-hybrid density-functional theory made rigorous, *J. Chem. Phys.*, 2011, **134**, 064113.
89 E. Brémond and C. Adamo, Seeking for parameter-free double-hybrid functionals: the PBE0-DH model, *J. Chem. Phys.*, 2011, **135**, 024106.
90 E. Brémond, M. Savarese, Á. J. Pérez-Jiménez, J. C. Sancho-Garca and C. Adamo, Systematic improvement of density functionals through parameter-free hybridization schemes, *J. Phys. Chem. Lett.*, 2015, **6**, 3540–3545.
91 T. Tsuneda and K. Hirao, Long-range correction for density functional theory, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 375–390.
92 T. Yanai, D. P. Tew and N. C. Handy, A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP), *Chem. Phys. Lett.*, 2004, **393**, 51–57.
93 H. Iikura, T. Tsuneda, T. Yanai and K. Hirao, A long-range correction scheme for generalized-gradient-approximation exchange functionals, *J. Chem. Phys.*, 2001, **115**, 3540–3544.
94 J.-D. Chai and M. Head-Gordon, Systematic optimization of long-range corrected hybrid density functionals, *J. Chem. Phys.*, 2008, **128**, 084106.
95 J.-D. Chai and M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.
96 D. Tozer and N. Handy, On the determination of excitation energies using density functional theory, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2117–2121.
97 D. Tozer, Relationship between long-range charge-transfer excitation energy error and integer discontinuity in Kohn–Sham theory, *J. Chem. Phys.*, 2003, **119**, 12697–12699.
98 T. M. Henderson, B. G. Janesko and G. E. Scuseria, Range separation and local hybridization in density functional theory, *J. Phys. Chem. A*, 2008, **112**, 12530–12542.
99 T. Stein, L. Kronik and R. Baer, Reliable prediction of charge transfer excitations in molecular complexes using time-dependent density functional theory, *J. Am. Chem. Soc.*, 2009, **131**, 2818–2820.
100 T. Körzdörfer and J.-L. Brédas, Organic electronic materials: recent advances in the DFT description of the ground and excited states using tuned range-separated hybrid functionals, *Acc. Chem. Res.*, 2014, **47**, 3284–3291.
101 É. Brémond, Á. J. Pérez-Jiménez, J. C. Sancho-Garca and C. Adamo, Range-separated hybrid density functionals made simple, *J. Chem. Phys.*, 2019, **150**, 201102.
102 É. Brémond, Á. J. Pérez-Jiménez, J. C. Sancho-Garca and C. Adamo, Range-separated hybrid and double-hybrid density functionals: A quest for the determination of the range-separation parameter, *J. Chem. Phys.*, 2020, **152**, 244124.
103 N. Mardirossian and M. Head-Gordon, Survival of the most transferable at the top of Jacob's ladder: Defining and testing the ωB97M(2) double hybrid density functional, *J. Chem. Phys.*, 2018, **148**, 241736.
104 E. Brémond, M. Savarese, Á. J. Pérez-Jiménez, J. C. Sancho-Garca and C. Adamo, Range-Separated Double-Hybrid Functional from Nonempirical Constraints, *J. Chem. Theory Comput.*, 2018, **14**, 4052–4062.
105 A. Karolewski, L. Kronik and S. Kümmel, Using optimally tuned range separated hybrid functionals in ground-state calculations: Consequences and caveats, *J. Chem. Phys.*, 2013, **138**, 204115.
106 J. Jaramillo, G. E. Scuseria and M. Ernzerhof, Local hybrid functionals, *J. Chem. Phys.*, 2003, **118**, 1068–1073.
107 T. M. Maier, A. V. Arbuznikov and M. Kaupp, Local hybrid functionals: Theory, implementation, and performance of an emerging new tool in quantum chemistry and beyond, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2019, **9**, e1378.
108 S. Grimme, Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel-and antiparallel-spin pair correlation energies, *J. Chem. Phys.*, 2003, **118**, 9095–9102.
109 S. Kozuch and J. M. Martin, Spin-component-scaled double hybrids: an extensive search for the best fifth-rung functionals blending DFT and perturbation theory, *J. Comput. Chem.*, 2013, **34**, 2327–2344.
110 É. Brémond, M. Savarese, J. C. Sancho-Garca, Á. J. Pérez-Jiménez and C. Adamo, Quadratic integrand double-hybrid

made spin-component-scaled, *J. Chem. Phys.*, 2016, **144**, 124104.
111 O. A. Vydrov and T. Van Voorhis, Nonlocal van der Waals density functional: The simpler the better, *J. Chem. Phys.*, 2010, **133**, 244103.
112 W. Hujo and S. Grimme, Performance of the van der Waals Density Functional VV10 and (hybrid) GGA Variants for Thermochemistry and Noncovalent Interactions, *J. Chem. Theory Comput.*, 2011, **7**, 3866–3871.
113 J. Aragó, E. Ort and J. C. Sancho-Garca, Nonlocal van der Waals approach merged with double-hybrid density functionals: Toward the accurate treatment of noncovalent interactions, *J. Chem. Theory Comput.*, 2013, **9**, 3437–3443.
114 A. H. Larsen, M. Kuisma, J. Löfgren, Y. Pouillon, P. Erhart and P. Hyldgaard, libvdwxc: A library for exchange-correlation functionals in the vdW-DF family, *Modell. Simul. Mater. Sci. Eng.*, 2017, **25**, 065004.
115 E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, A generally applicable atomic-charge dependent London dispersion correction, *J. Chem. Phys.*, 2019, **150**, 154122.
116 A. Otero-De-La-Roza and E. R. Johnson, Non-covalent interactions and thermochemistry using XDM-corrected hybrid and range-separated hybrid density functionals, *J. Chem. Phys.*, 2013, **138**, 204109.
117 S. N. Steinmann and C. Corminboeuf, Comprehensive benchmarking of a density-dependent dispersion correction, *J. Chem. Theory Comput.*, 2011, **7**, 3567–3577.
118 J. S. Garcia, E. Brémond, M. Campetella, I. Ciofini and C. Adamo, Small basis set allowing the recovery of dispersion interactions with double-hybrid functionals, *J. Chem. Theory Comput.*, 2019, **15**, 2944–2953.
119 A. Karton, A computational chemist's guide to accurate thermochemistry for organic molecules, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2016, **6**, 292–310.
120 P. Morgante and R. Peverati, The devil in the details: a tutorial review on some undervalued aspects of density functional theory calculations, *Int. J. Quantum Chem.*, 2020, **120**, e26332.
121 N. Mardirossian and M. Head-Gordon, Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals, *Mol. Phys.*, 2017, **115**, 2315–2372.
122 M. Bursch, J.-M. Mewes, A. Hansen, S. Grimme and D. F. T. Best-practice, protocols for basic molecular computational chemistry, *Angew. Chem., Int. Ed.*, 2022, **61**, e202205735.
123 A. Karton and M. T. De Oliveira, Good Practices in Database Generation for Benchmarking Density Functional Theory, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2025, **15**, e1737.
124 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, 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.*, 2017, **19**, 32184–32215.
125 A. Karton, N. Sylvetsky and J. M. Martin, W4-17: a diverse and high-confidence dataset of atomization energies for benchmarking high-level electronic structure methods, *J. Comput. Chem.*, 2017, **38**, 2063–2075.
126 P. Morgante and R. Peverati, ACCDB: a collection of chemistry databases for broad computational purposes, *J. Comput. Chem.*, 2019, **40**, 839–848.
127 J. Rezác, K. E. Riley and P. Hobza, S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures, *J. Chem. Theory Comput.*, 2011, **7**, 2427–2438.
128 B. Brauer, M. K. Kesharwani, S. Kozuch and J. M. L. Martin, The S66x8 benchmark for noncovalent interactions revisited: explicitly correlated ab initio methods and density functional theory, *Phys. Chem. Chem. Phys.*, 2016, **18**, 20905–20925.
129 Z. Ni, Y. Guo, F. Neese, W. Li and S. Li, Cluster-in-molecule local correlation method with an accurate distant pair correction for large systems, *J. Chem. Theory Comput.*, 2021, **17**, 756–766.
130 M. Veril, A. Scemama, M. Caffarel, F. Lipparini, M. Boggio-Pasqua, D. Jacquemin and P.-F. Loos, QUESTDB: A database of highly accurate excitation energies for the electronic structure community, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2021, **11**, e1517.
131 A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner and G. Ceder, *et al.*, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Mater.*, 2013, **1**, 011002.
132 M. K. Horton, P. Huck, R. X. Yang, J. M. Munro, S. Dwaraknath, A. M. Ganose, R. S. Kingsbury, M. Wen, J. X. Shen and T. S. Mathis, *et al.*, Accelerated data-driven materials science with the Materials Project, *Nat. Mater.*, 2025, **24**, 1522–1532.
133 K. Choudhary, K. F. Garrity, A. C. Reid, B. DeCost, A. J. Biacchi, A. R. Hight Walker, Z. Trautt, J. Hattrick-Simpers, A. G. Kusne and A. Centrone, *et al.*, The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design, *npj Comput. Mater.*, 2020, **6**, 173.
134 K. Choudhary, D. Wines, K. Li, K. F. Garrity, V. Gupta, A. H. Romero, J. T. Krogel, K. Saritas, A. Fuhr and P. Ganesh, *et al.*, JARVIS-Leaderboard: a large scale benchmark of materials design methods, *npj Comput. Mater.*, 2024, **10**, 93.
135 L. Talirz, S. Kumbhar, E. Passaro, A. V. Yakutovich, V. Granata, F. Gargiulo, M. Borelli, M. Uhrin, S. P. Huber and S. Zoupanos, *et al.*, Materials Cloud, a platform for open computational science, *Sci. Data*, 2020, **7**, 299.
136 Y. Shu, Z. Zhu, S. Kanchanakungwankul and D. G. Truhlar, Small Representative Databases for Testing and Validating Density Functionals and Other Electronic Structure Methods, *J. Phys. Chem. A*, 2024, **128**, 6412–6422.
137 J. Rezac, Non-covalent interactions atlas benchmark data sets: Hydrogen bonding, *J. Chem. Theory Comput.*, 2020, **16**, 2355–2368.

138 J. Řezáč, Non-covalent interactions atlas benchmark data sets 2: Hydrogen bonding in an extended chemical space, *J. Chem. Theory Comput.*, 2020, **16**, 6305–6316.
139 K. Kříž, M. Novacek and J. Rezac, Non-covalent interactions atlas benchmark data sets 3: Repulsive contacts, *J. Chem. Theory Comput.*, 2021, **17**, 1548–1561.
140 K. Kříž and J. Řezáč, Non-covalent interactions atlas benchmark data sets 4: σ-hole interactions, *Phys. Chem. Chem. Phys.*, 2022, **24**, 14794–14804.
141 J. Řezáč, Non-Covalent Interactions Atlas benchmark data sets 5: London dispersion in an extended chemical space, *Phys. Chem. Chem. Phys.*, 2022, **24**, 14780–14793.
142 D. S. Levine, M. Shuaibi, E. W. C. Spotte-Smith, M. G. Taylor, M. R. Hasyim, K. Michel, I. Batatia, G. Csanyi, M. Dzamba and P. Eastman, *et al.*, The open molecules 2025 (omol25) dataset, evaluations, and models, arXiv, 2025, preprint, arXiv:2505.08762, DOI: 10.48550/arXiv2505.08762.
143 A. Karton, E. Rabinovich, J. M. Martin and B. Ruscic, W4 theory for computational thermochemistry: In pursuit of confident sub-kJ mol^{-1} predictions, *J. Chem. Phys.*, 2006, **125**, 144108.
144 A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez and J. F. Stanton, HEAT: High accuracy extrapolated ab initio thermochemistry, *J. Chem. Phys.*, 2004, **121**, 11599–11613.
145 P. M. Nelson, Z. L. Glick and C. D. Sherrill, Approximating large-basis coupled-cluster theory vibrational frequencies using focal-point approximations, *J. Chem. Phys.*, 2023, **159**, 094104.
146 D. Feller, K. A. Peterson and D. A. Dixon, Further benchmarks of a composite, convergent, statistically calibrated coupled-cluster-based approach for thermochemical and spectroscopic studies, *Mol. Phys.*, 2012, **110**, 2381–2399.
147 L. A. Curtiss, P. C. Redfern and K. Raghavachari, Gn theory, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 810–825.
148 N. J. DeYonker, T. R. Cundari and A. K. Wilson, The correlation consistent composite approach (ccCA): An alternative to the Gaussian-n methods, *J. Chem. Phys.*, 2006, **124**, 114104.
149 D. G. Liakos and F. Neese, Is it possible to obtain coupled cluster quality energies at near density functional theory cost? Domain-based local pair natural orbital coupled cluster vs modern density functional theory, *J. Chem. Theory Comput.*, 2015, **11**, 4054–4063.
150 I. Sandler, J. Chen, M. Taylor, S. Sharma and J. Ho, Accuracy of DLPNO-CCSD(T): Effect of basis set and system size, *J. Phys. Chem. A*, 2021, **125**, 1553–1563.
151 A. Karton, Highly accurate CCSDT(Q)/CBS reaction barrier heights for a diverse set of transition structures: Basis set convergence and cost-effective approaches for estimating post-CCSD(T) contributions, *J. Phys. Chem. A*, 2019, **123**, 6720–6732.
152 É. Brémond, H. Li, Á. J. Pérez-Jiménez, J. C. Sancho-Garca and C. Adamo, Tackling an accurate description of molecular reactivity with double-hybrid density functionals, *J. Chem. Phys.*, 2022, **156**, 161101.
153 V. K. Prasad, Z. Pei, S. Edelmann, A. Otero-de-la Roza and G. A. DiLabio, BH9, a New Comprehensive Benchmark Data Set for Barrier Heights and Reaction Energies: Assessment of Density Functional Approximations and Basis Set Incompleteness Potentials, *J. Chem. Theory Comput.*, 2022, **18**, 151–166.
154 J. E. Alfonso-Ramos, C. Adamo, É. Brémond and T. Stuyver, Improving the reliability of, and confidence in, DFT functional benchmarking through active learning, *J. Chem. Theory Comput.*, 2025, **21**, 1752–1761.
155 N. Mardirossian and M. Head-Gordon, Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V, *J. Chem. Phys.*, 2015, **142**, 074111.
156 F. A. Hamprecht, A. J. Cohen, D. J. Tozer and N. C. Handy, Development and assessment of new exchange-correlation functionals, *J. Chem. Phys.*, 1998, **109**, 6264–6271.
157 D. J. Tozer and N. C. Handy, Development of new exchange-correlation functionals. 2, *J. Phys. Chem. A*, 1998, **102**, 3162–3168.
158 A. D. Becke, Exploring the limits of gradient corrections in density functional theory, *J. Comput. Chem.*, 1999, **20**, 63–69.
159 R. Peverati, Fitting elephants in the density functionals zoo: Statistical criteria for the evaluation of density functional theory methods as a suitable replacement for counting parameters, *Int. J. Quantum Chem.*, 2021, **121**, e26379.
160 S. E. Wheeler and K. Houk, Integration grid errors for meta-GGA-predicted reaction energies: Origin of grid errors for the M06 suite of functionals, *J. Chem. Theory Comput.*, 2010, **6**, 395–404.
161 B. Chan, W. Dawson and T. Nakajima, Data Quality in the Fitting of Approximate Models: A Computational Chemistry Perspective, *J. Chem. Theory Comput.*, 2024, **20**, 10468–10476.
162 A. J. Cohen, P. Mori-Sánchez and W. Yang, Challenges for density functional theory, *Chem. Rev.*, 2012, **112**, 289–320.
163 R. Pederson, B. Kalita and K. Burke, Machine learning and density functional theory, *Nat. Rev. Phys.*, 2022, **4**, 357–358.
164 J. Kirkpatrick, B. McMorrow, D. H. Turban, A. L. Gaunt, J. S. Spencer, A. G. Matthews, A. Obika, L. Thiry, M. Fortunato and D. Pfau, *et al.*, Pushing the frontiers of density functionals by solving the fractional electron problem, *Science*, 2021, **374**, 1385–1389.
165 D. Khan, A. J. Price, B. Huang, M. L. Ach and O. A. von Lilienfeld, Adapting hybrid density functionals with machine learning, *Sci. Adv.*, 2025, **11**, eadt7769.
166 G. Luise, C.-W. Huang, T. Vogels, D. P. Kooi, S. Ehlert, S. Lanius, K. J. Giesbertz, A. Karton, D. Gunceler and M. Stanley, *et al.*, Accurate and scalable exchange-correlation with deep learning, *arXiv*, 2025, preprint, arXiv:2506.14665, DOI: 10.48550/arXiv.2506.14665.
167 C.-W. Ju, E. J. French, N. Geva, A. W. Kohn and Z. Lin, Stacked ensemble machine learning for range-separation parameters, *J. Phys. Chem. Lett.*, 2021, **12**, 9516–9524.
168 Z. Qiao, M. Welborn, A. Anandkumar, F. R. Manby and T. F. Miller, OrbNet: Deep learning for quantum chemistry using symmetry-adapted atomic-orbital features, *J. Chem. Phys.*, 2020, **153**, 124111.

169 S. Dick and M. Fernandez-Serra, Machine learning accurate exchange and correlation functionals of the electronic density, *Nat. Commun.*, 2020, **11**, 3509.
170 Y. Chen, L. Zhang, H. Wang and W. E, Ground state energy functional with Hartree-Fock efficiency and chemical accuracy, *J. Phys. Chem. A*, 2020, **124**, 7155–7165.
171 Y. Liu, C. Zhang, Z. Liu, D. G. Truhlar, Y. Wang and X. He, Supervised learning of a chemistry functional with damped dispersion, *Nat. Comput. Sci.*, 2023, **3**, 48–58.
172 J. P. Perdew, Artificial intelligence ''sees'' split electrons, *Science*, 2021, **374**, 1322–1323.
173 K. R. Bryenton, A. A. Adeleke, S. G. Dale and E. R. Johnson, Delocalization error: The greatest outstanding challenge in density-functional theory, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2023, **13**, e1631.
174 I. S. Gerasimov, T. V. Losev, E. Y. Epifanov, I. Rudenko, I. S. Bushmarinov, A. A. Ryabov, P. A. Zhilyaev and M. G. Medvedev, Comment on ''Pushing the frontiers of density functionals by solving the fractional electron problem'', *Science*, 2022, **377**, eabq3385.
175 J. Kirkpatrick, B. McMorrow, D. H. Turban, A. L. Gaunt, J. S. Spencer, A. G. Matthews, A. Obika, L. Thiry, M. Fortunato and D. Pfau, *et al.*, Response to Comment on ''Pushing the frontiers of density functionals by solving the fractional electron problem'', *Science*, 2022, **377**, eabq4282.
176 H. Zhao, T. Gould and S. Vuckovic, Deep Mind 21 functional does not extrapolate to transition metal chemistry, *Phys. Chem. Chem. Phys.*, 2024, **26**, 12289–12298.
177 E. Palos, E. Lambros, S. Dasgupta and F. Paesani, Density functional theory of water with the machine-learned DM21 functional, *J. Chem. Phys.*, 2022, **156**, 161103.
178 B. Jijila, V. Nirmala, P. Selvarengan, D. Kavitha, V. Arun Muthuraj and A. Rajagopal, Employing neural density functionals to generate potential energy surfaces, *J. Mol. Model.*, 2024, **30**, 65.
179 F. Ju, X. Wei, L. Huang, A. J. Jenkins, L. Xia, J. Zhang, J. Zhu, H. Yang, B. Shao and P. Dai, *et al.*, Acceleration without disruption: DFT software as a service, *J. Chem. Theory Comput.*, 2024, **20**, 10838–10851.
180 Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova and S. Sharma, *et al.*, PySCF: the Python-based simulations of chemistry framework, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2018, **8**, e1340.
181 A. D. Becke, G. Santra and J. M. Martin, A double-hybrid density functional based on good local physics with outstanding performance on the GMTKN55 database, *J. Chem. Phys.*, 2023, **158**, 151103.
182 A. D. Becke, A remarkably simple dispersion damping scheme and the DH24 double hybrid density functional, *J. Chem. Phys.*, 2024, **160**, 204118.
183 E. Brémond, M. Rodríguez-Mayorga, A.-J. Pérez-Jiménez, C. Adamo and J.-C. Sancho-García, Assessment of the nonempirical r^2SCAN-QIDH double-hybrid density functional against large and diverse datasets, *J. Chem. Phys.*, 2023, **159**, 141101.
184 É. Brémond, Á. J. Pérez-Jiménez, J. C. Sancho-Garca and C. Adamo, SOS1-RSX-QIDH: A spin-opposite-scaled range-separated-exchange quadratic-integrand double-hybrid density functional, *J. Chem. Phys.*, 2023, **159**, 234104.
185 M. Piccardo, E. Penocchio, C. Puzzarini, M. Biczysko and V. Barone, Semi-experimental equilibrium structure determinations by employing B3LYP/SNSD anharmonic force fields: Validation and application to semirigid organic molecules, *J. Phys. Chem. A*, 2015, **119**, 2058–2082.
186 É. Brémond, M. Savarese, N. Q. Su, Á. J. Pérez-Jiménez, X. Xu, J. C. Sancho-Garca and C. Adamo, Benchmarking density functionals on structural parameters of small-/medium-sized organic molecules, *J. Chem. Theory Comput.*, 2016, **12**, 459–465.
187 K. Kulaev, A. Ryabov, M. Medvedev, E. Burnaev and V. Vanovskiy, On the practical applicability of modern DFT functionals for chemical computations. Case study of DM21 applicability for geometry optimization, *arXiv*, 2025, preprint, arXiv:2501.12149, DOI: 10.48550/arXiv.2501.12149.
188 R. Sure and S. Grimme, Comprehensive benchmark of association (free) energies of realistic host-guest complexes, *J. Chem. Theory Comput.*, 2015, **11**, 3785–3801.
189 J. Gorges, S. Grimme and A. Hansen, Reliable prediction of association (free) energies of supramolecular complexes with heavy main group elements-the HS13L benchmark set, *Phys. Chem. Chem. Phys.*, 2022, **24**, 28831–28843.
190 J. Gorges, B. Baedorf, S. Grimme and A. Hansen, Efficient computation of the interaction energies of very large non-covalently bound complexes, *Synlett*, 2023, 1135–1146.
191 B. Huang, O. A. von Lilienfeld, J. T. Krogel, A. Benali and D. M. C. Toward, accuracy across chemical space with scalable Δ-QML, *J. Chem. Theory Comput.*, 2023, **19**, 1711–1721.
192 M. Schwilk, D. N. Tahchieva and O. A. von Lilienfeld, Large yet bounded: Spin gap ranges in carbenes, *arXiv*, 2020, preprint, arXiv:2004.10600, DOI: 10.48550/arXiv.2004.10600.
193 R. Ramakrishnan, P. O. Dral, M. Rupp and O. A. Von Lilienfeld, Quantum chemistry structures and properties of 134 kilo molecules, *Sci. Data*, 2014, **1**, 1–7.
194 B. Huang and O. A. Von Lilienfeld, Ab initio machine learning in chemical compound space, *Chem. Rev.*, 2021, **121**, 10001–10036.
195 Y. Rui, Y. Chen, E. Ivanova, V. B. Kumar, S. Smiga, I. Grabowski and P. O. Dral, The best DFT functional is the ensemble of functionals, *Adv. Sci.*, 2024, **11**, 2408239.

