Optimization of a Range-Separated UW12 Hybrid Functional

Zack M. Williams and Frederick R. Manby[*]

Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, UK

E-mail: fred.manby@bristol.ac.uk

Abstract

In a previous paper we presented a new hybrid functional B-LYP-osUW12-D3(BJ) containing the Unsöld-w_{12} (UW12) hybrid correlation model. In this paper we present a new 15-parameter range-separated hybrid density functional using a power series expansion together with UW12 correlation. This functional is optimized using the survival of the fittest strategy developed for the ωB97X-V functional, fitted to data from the Main Group Chemistry Database (MGCDB84). In addition we optimize a standard hybrid and double hybrid using the same method. We show that our fully self-consistent UW12 hybrid functional WM21-D3(BJ) outperforms both of these functionals and other range-separated hybrid functionals.

1 Introduction

Density-functional theory (DFT) has in recent years become the most widely used electronic structure method.[1,2] Much of this success derives from the ability of DFT to achieve reasonable accuracy at a relatively low cost.

While exact in theory, Kohn-Sham density-functional theory contains an unknown energy component – the exchange-correlation (XC) functional – which must be approximated. Thus the accuracy of practical DFT calculations depends on the choice of exchange-correlation functional.[3–6]

There is no systematic method to improve the exchange-correlation functional, but strategies have evolved to improve accuracy through including additional terms in the exchange-correlation functional. The simplest density-functional is the local-density approximation (LDA) in which the energy density depends only on the electron density at each point. Accuracy may be improved by including terms that depend on the gradient of the electron density, forming the generalized-gradient approximation (GGA). There are several widely-used functionals of this type, including BLYP and PBE.[7–9]

Further improvements in accuracy can be attained by including higher order density gradients and the kinetic-energy density.[10] However there appears to be a limit to the accuracy achievable through such approximations. Problems with these semi-local functionals can be related to the self-interaction error (SIE), arising from the incomplete cancellation between the self-interacting portions of Coulomb and exchange energies.[11] In semi-local functionals, the potential felt by a removed electron from a neutral molecule exhibits exponential decay in the long range, rather the the physically correct $-r^{-1}$ decay behavior.[12,13]

Hybrid density-functional methods attempt to reduce these errors by replacing a fraction of approximate density-functional exchange with exact Hartree-Fock exchange.[14,15] Hybrid functionals remain some the most popular density functionals in use today, though they are not free from problems. Most popular hybrid functionals such as B3LYP and PBE0 contain a relatively small fraction of exact exchange,[16,17] so they still contain a large amount of self-interaction error. Also, while Hartree-Fock exchange has the correct long-range behavior, global hybrid functionals with a fixed fraction c_{x} do not, since the long-range exchange potential behaves as $-c_{\mathrm{x}}r^{-1}$ instead of the correct $-r^{-1}$.

One solution to this problem is through range-separation, where the Coulomb operator is

split into short- and long-range components, such that the amount of exact exchange varies as a function of inter-electronic distance.[18,19] Such functionals are called range-separated hybrids (RSHs), and can be constructed to ensure the correct long-range behavior. In addition, these functionals exhibit reduced self-interaction error according to typical measures.[11]

It is also possible to further increase the accuracy by including some amount of non-local correlation by introducing a dependence on the virtual (unoccupied) orbitals.[20,21] Double hybrid functionals achieve this by including a fraction of MP2 correlation; they have received much interest in recent years.[22–24]

The UW12 approximation is a non-local correlation approximation given by[25]

$$E_{\mathrm{c}}^{\mathrm{UW12}} = \frac{1}{2} \sum_{ijab} [\langle ij|w_{12}|ab\rangle - \langle ij|w_{12}|ba\rangle] \langle ab|r_{12}^{-1}|ij\rangle \tag{1}$$

for occupied spin orbitals $|i\rangle, |j\rangle$, virtual spin orbitals projector $|a\rangle, |b\rangle$, and Coulomb operator r_{12}^{-1}, and two-electron geminal function w_{12}. This approximation may be formulated in a way which includes no virtual orbital dependence.[25] Global hybrid functionals which include UW12 correlation such as B-LYP-osUW12 have been shown to give accurate results without a virtual orbital dependence.[26]

In this work we wish to create a range-separated UW12 hybrid functional with correct long-range behavior using the ωB97 exchange-correlation functional as a starting point. Of particular interest are the ωB97X-V, ωB97M-V and ωB97M(2) functionals optimized using the survival of the fittest strategy.[27–29] These range-separated hybrid functionals are among the most accurate functionals currently available.

The range-separated GGA hybrid ωB97X-V is a 10 parameter functional containing range-corrected B97 exchange, B97 correlation, and the non-local VV10 functional,[30–32] with the parameters chosen to minimize the error in both the training and validation datasets considered. This method was then extended to produce the range-separated meta-GGA hybrid ωB97M-V, utilizing the large MGCDB84 database for optimization, and including meta-

3

GGA components in the optimization,[5] as well as an MP2 energy correction in ωB97M(2).[33]

Following this approach we set out to create a range-separated hybrid functional of the form ωB97X-osUW12 using the ωB97 functional together with opposite-spin UW12 correlation and a dispersion correction. We do not include any meta-GGA contribution in this work, but note that this could be included in the future.

2 Theory

The functional we aim to create contains multiple components and we begin with a brief discussion of each of them.

The Hartree-Fock exchange energy for a molecular system is given by:

$$E_{\mathrm{x}}^{\mathrm{HF}} = -\frac{1}{2} \sum_{\sigma} \sum_{ij} \left\langle i_\sigma j_\sigma \left| r_{12}^{-1} \right| j_\sigma i_\sigma \right\rangle \tag{2}$$

for $r_{12} = |\vec{r}_1 - \vec{r}_2|$ and occupied spin orbitals $|i_\sigma\rangle$, $|j_\sigma\rangle$. In the local spin density approximation (LSDA), the exchange functional is given by

$$E_{\mathrm{x}}^{\mathrm{LSDA}} = \sum_{\sigma} \int e_{\mathrm{x}}^{\mathrm{LSDA}}(\rho_\sigma) \, \mathrm{d}\vec{r}, \tag{3}$$

where

$$e_{\mathrm{x}}^{\mathrm{LSDA}}(\rho_\sigma) = -\frac{3}{4} \left(\frac{6}{4\pi} \right)^{1/3} \rho_\sigma^{4/3}(\vec{r}) \tag{4}$$

is the exchange energy per unit volume.

In the range-separation scheme, the Coulomb operator is split into short- and long-range parts using the error function such that[18,19]

$$\frac{1}{r_{12}} = \frac{\mathrm{erfc}(\omega r_{12})}{r_{12}} + \frac{\mathrm{erf}(\omega r_{12})}{r_{12}} \tag{5}$$

4

for range-separation parameter ω. The long-range exchange interaction is then given by

$$E_{\mathrm{x}}^{\text{lr-HF}} = -\frac{1}{2} \sum_{\sigma} \sum_{ij} \left\langle i_\sigma j_\sigma \left| \operatorname{erf}(\omega r_{12}) r_{12}^{-1} \right| j_\sigma i_\sigma \right\rangle, \tag{6}$$

where the kernel $\operatorname{erf}(\omega r_{12}) r_{12}^{-1}$ tends to r_{12}^{-1} as $r_{12} \to \infty$. By replacing the r_{12}^{-1} operator with $\operatorname{erfc}(\omega r_{12}) r_{12}^{-1}$ in the derivation of the LSDA exchange energy, the short range LSDA functional may be written as

$$E_{\mathrm{x}}^{\text{sr-LSDA}} = \sum_{\sigma} \int e_{\mathrm{x}}^{\text{LSDA}}(\rho_\sigma) F_\sigma(a_\sigma) \, \mathrm{d}\vec{r}, \tag{7}$$

for

$$F(a_\sigma) = 1 - \frac{2}{3} a_\sigma \left[2\sqrt{\pi} \operatorname{erf}\left(\frac{1}{a_\sigma}\right) - 3a_\sigma + a_\sigma^3 + (2a_\sigma - a_\sigma^3) \exp\left(-\frac{1}{a_\sigma^2}\right) \right] \tag{8}$$

where $a_\sigma = \omega/k_{F,\sigma}$, and $k_{F,\sigma} = [6\pi^2 \rho_\sigma]^{1/3}$ is the spin-polarized Fermi wave vector.[18]

The original range-separation scheme consists of short-range density-functional exchange and long-range HF exchange, with no HF exchange at short range. It has been shown that including a small fraction $c_{\mathrm{x}}^{\text{sr}}$ of short-range HF exchange in addition to the short-range DFT exchange is beneficial without affecting the correct long-range behavior.[34]

While the LSDA exchange functional is helpful to illustrate range-separation in exchange functionals, it is not suitable for most applications due to the homogeneous nature of the functional. A better approach is to use a GGA functional for the density functional exchange component. The B97 functional uses a power series expansion involving the density and its first derivative:[30]

$$E_{\mathrm{x}}^{\text{B97}} = \sum_{\sigma} \int e_{\mathrm{x}}^{\text{LSDA}}(\rho_\sigma) g_{\mathrm{x}}(s_\sigma) \, \mathrm{d}\vec{r} \tag{9}$$

where g_x is a power series given by

$$g_x(s_\sigma) = \sum_{i=0}^{m} c_x^i \left[\frac{\gamma_x s_\sigma^2}{1 + \gamma_x s_\sigma^2} \right]^i, \tag{10}$$

with $s_\sigma = |\nabla \rho_\sigma| / \rho_\sigma^{4/3}$, and where c_x^i is the ith coefficient. This expansion includes the non-linear exchange parameter $\gamma_x = 0.004$,[35] while the linear parameters $\{c_x^i\}$ may be determined.

In a similar manner to the LSDA exchange functional, the B97 exchange functional may be modified to produce the short-range B97 exchange functional given by:[31]

$$E_x^{\text{sr-B97}} = \sum_\sigma \int e_x^{\text{LSDA}}(\rho_\sigma) F(a_\sigma) g_x(s_\sigma) \, d\vec{r}. \tag{11}$$

The B97 correlation functional is constructed in a similar way, with separate expansions for the same and opposite spin components such that

$$E_{c,\text{ss}}^{B97} = \sum_\sigma^{\alpha,\beta} \int e_{c,\sigma\sigma}^{\text{LSDA}}(\rho_\sigma) g_{c,\text{ss}}(s_\sigma) \, d\vec{r} \tag{12}$$

$$E_{c,\text{os}}^{B97} = \int e_{c,\alpha\beta}^{\text{LSDA}}(\rho_\alpha, \rho_\beta) g_{c,\text{os}}(s_{\alpha\beta}) \, d\vec{r}, \tag{13}$$

for $s_{\alpha\beta}^2 = (s_\alpha + s_\beta)/2$. The power series expansions in these expressions are like equation (10), with different linear parameters $\{c_{c,\text{ss}}^i\}$ and $\{c_{c,\text{ss}}^i\}$, and non-linear parameters $\gamma_{c,\text{ss}} = 0.02$ and $\gamma_{c,\text{os}} = 0.006$ respectively. These were fitted to the correlation energies of He and Ne.[35]

The same- and opposite-spin LSDA correlation energy densities per unit volume $e_{c,\sigma\sigma}^{\text{LSDA}}, e_{c,\alpha\beta}^{\text{LSDA}}$ are extracted from the total LSDA correlation energy density using:

$$e_{c,\sigma\sigma}^{\text{LSDA}}(\rho_\sigma) = e_c^{\text{LSDA}}(\rho_\sigma, 0) \tag{14}$$

$$e_{c,\alpha\beta}^{\text{LSDA}}(\rho_\alpha, \rho_\beta) = e_c^{\text{LSDA}}(\rho_\alpha, \rho_\beta) - e_c^{\text{LSDA}}(\rho_\alpha, 0) - e_c^{\text{LSDA}}(0, \rho_\beta) \tag{15}$$

The linear coefficients $\{c_{c,\text{ss}}^i\}$ and $\{c_{c,\text{os}}^i\}$ may be optimized systematically in the same manner as the exchange coefficients.

The UW12 correlation energy can also be split into the same and opposite spin components such that

$$E_{\mathrm{c,os}}^{\mathrm{UW12}} = \sum_{ij} \langle i_\alpha j_\beta | w_{12}^{\mathrm{os}} \hat{Q}_{12} r_{12}^{-1} | i_\alpha j_\beta \rangle \tag{16}$$

$$E_{\mathrm{c,ss}}^{\mathrm{UW12}} = \frac{1}{2} \sum_\sigma \sum_{ij} \left[\langle i_\sigma j_\sigma | w_{12}^{\mathrm{ss}} \hat{Q}_{12} r_{12}^{-1} | i_\sigma j_\sigma \rangle - \langle j_\sigma i_\sigma | w_{12}^{\mathrm{ss}} \hat{Q}_{12} r_{12}^{-1} | i_\sigma j_\sigma \rangle \right], \tag{17}$$

where w_{12}^{ss} and w_{12}^{os} are the same- and opposite-spin geminal functions respectively, with projector $\hat{Q}_{12}$ given by

$$\hat{Q}_{12} = \hat{v}_1 \hat{v}_2 = (\hat{1} - \hat{o}_1)(\hat{1} - \hat{o}_2) = \hat{1} - \hat{o}_1 - \hat{o}_2 + \hat{o}_1 \hat{o}_2, \tag{18}$$

for which $\hat{o}$ and $\hat{v}$ are projectors onto the occupied and virtual spaces respectively. Using this relation the UW12 energy for a given geminal may be evaluated either using the virtual space or entirely within the occupied space. In the virtual space, the total UW12 energy may be evaluated for a fixed set of orbitals using equation (1). In the occupied space, the UW12 energy may be written solely in terms of the one-particle reduced density matrix allowing the orbitals to be optimized self-consistently within a generalized Kohn-Sham scheme.

The accuracy of the approximation depends on the choice of geminal functions w_{12}^{os} and w_{12}^{ss}. In previous work we have set $w_{12}^{\mathrm{ss}} = \kappa w_{12}^{\mathrm{os}}$, with $\kappa = 0.5$ in XCH-BLYP-UW12,[25] or $\kappa = 0$ in B-LYP-osUW12, where the same-spin contribution is neglected. In these functionals, the opposite-spin geminal has been a one-parameter exponential function:[36]

$$w_{12}^{\mathrm{os}} = -\frac{1}{2} r_{\mathrm{c}} \exp\left[-\frac{r_{12}}{r_{\mathrm{c}}} \right] \tag{19}$$

Integrals involving this kernel have been derived,[36] but we instead expand w_{12}^{os} as

$$w_{12}^{\mathrm{os}} = \sum_i c_{\mathrm{u}}^i \exp[-\gamma_i r_{12}^2] \tag{20}$$

for exponents $\{\gamma_i\}$ and coefficients $\{c_u^i\}$. For a fixed set of exponents the coefficients are fitted to the exponential function, and in this way the same code may be used to model other kernels.

Rather than constraining this expansion to fit a fixed function, the coefficients could be fitted directly to data, and this forms an integral part of the development described in this paper. In this way the opposite and same spin coefficients could be optimized separately, however, in this work we continue to use the opposite-spin only UW12 (osUW12) approach used in B-LYP-osUW12-D3(BJ).

It has been shown that electronic dispersion forces are crucial for a correct description of long-range electron correlation.[37–39] However, most standard XC functionals fail to account for these effects since they use local properties to calculate the XC energy. This lack of dispersion is a significant problem in DFT with a number of solutions suggested. Many modern density-functionals include a dispersion-correction term to account for these effects. One approach is to use empirical dispersion corrections, such as the D3(BJ) and D4 corrections.[40,41] While another is to create long-range (non-local) density functionals specifically designed to account for these effects, such as the VV10 functional.[32]

In this work we use the D3(BJ) correction

$$E_{\text{disp}}^{\text{D3(BJ)}} = -\sum_{n=6,8} \sum_{A<B} s_n \frac{C_{(n)}^{AB}}{R_{AB}^n} f_{\text{damp}}^{(n)}(R_{AB}), \tag{21}$$

where $C_{(n)}^{AB}$ are the dispersion coefficients between atoms A, B, with inter-atomic distance R_{AB}, scale factors s_n, and Becke-Johnson damping function $f_{\text{damp}}^{(n)}$ given by[42–44]

$$f_{\text{damp}}^{(n)}(R_{AB}) = \frac{R_{AB}^n}{R_{AB}^n + (a_1 R_0 + a_2)^n}. \tag{22}$$

Which includes functional-specific damping parameters a_0, a_1, and $R_0 = \sqrt{C_8^{AB}/C_6^{AB}}$. A future possibility would be to re-optimize the final functional to include VV10 rather than D3(BJ).

Table 1: Details of the 8 subsets of the Main-Group Chemistry Database (MGCDB84), including the multiplicative factors used in the weighting scheme. The naming scheme and factors f_d are taken from Ref. 29, where the labels 'easy" and "difficult" indicate whether the associated reactions are considered challenging for density-functionals.

Data-type	Details	Factor, f_d
BH	Barrier heights	10
IE	Isomerization (easy)	1000
ID	Isomerization (difficult)	10
NCED	Non-covalent (easy) dimers	100
NCEC	Non-covalent (easy) clusters	100
NCD	Non-covalent (difficult)	10
TCE	Thermochemistry (easy)	1
TCD	Thermochemistry (difficult)	1

The resulting functional formed by combining all these components is made up of the linear parameters c_x^{sr}, $\{c_x^i\}$, $\{c_{c,ss}^i\}$, $\{c_{c,os}^i\}$, $\{c_u^i\}$, s_6, and s_8. In addition to the non-linear parameters ω, γ_x, $\gamma_{c,ss}$, $\gamma_{c,os}$, γ_i, a_1, and a_2.

3   Datasets and Fitting

Following the method in Ref. 29, we use the Main-Group Chemistry Database (MGCDB84) to train and evaluate our new functional, for completeness we give a brief summary of the database and method here.

MGCDB84 is compiled from the work of numerous groups. It consists of 84 datasets with 4986 data-points, which includes a number of chemical properties. All but two datasets are arranged into 8 subsets based on the types of chemical reactions involved, these are summarised in table 1. The other datasets consist of absolute atomic energies (AE18) and rare-gas dimer potential energy curves (RG10).[45,46]

The full database is split into three sets for training, validation, and testing, with datasets from each subset distributed among these. We use the training set to fit the functional parameters; it contains 25 datasets and 870 data-points. We use the validation set to determine the whether the fitted functional is transferable; it consists of 35 datasets with 2964 data

points. We only use the test set to calculate the accuracy of the final functional; it contains 24 datasets with 1152 data-points. This same splitting was previously used to fit the ωB97M-V and ωB97M(2) functionals.[28,29]

The functional ωB97X-osUW12-D3(BJ) consists of multiple parameters; the linear parameters are selected and optimized using the procedure outlined below, while the non-linear parameters are fixed in advance of the optimization. The range-separation parameter $\omega = 0.3$ was chosen as it is the value used in the ωB97X and ωB97X-V functional.[27,31] The B97 parameters γ_x, $\gamma_{c,ss}$, $\gamma_{c,os}$ remain fixed using the original B97 values. The non-linear dispersion parameters were chosen as $a_1 = 0.00$ and $a_2 = 5.49$. These were taken from the ωB97X-D3(BJ) functional, a re-implementation of ωB97X-V with a D3(BJ) correction instead of the non-local VV10 component.[47] In addition these parameters are almost identical to those optimized for our dispersion-corrected B-LYP-osUW12-D3(BJ) functional, for which $a_1 = 0.0$ and $a_2 = 5.45$.[26]

Unlike the other non-linear parameters, the UW12 Gaussian exponents are not chosen to reproduce an observable effect per-se. Instead they are chosen to give a degree of flexibility in the final fit of the geminal function. To give sufficient flexibility, the exponents were chosen to be the set $\{\gamma^{-n}, \ldots, \gamma^n\}$ with $n = 4$ and $\gamma = 3.16228$. This covers a number of orders of magnitude for the scale of the geminal.

For the linear parameters, we impose the uniform-electron gas limit $c_x^0 + c_x^{sr} = 1$, the long-range exchange limit ($c_x^{lr} = 1$), and the correct leading-order dispersion behavior ($s_6 = 1.0$). The remaining linear parameters may then be fitted using least-squares regression, with the power series expansions of the B97 functionals all calculated up to fourth order.[48]

Fitting is performed using energy components calculated on a fixed set of orbitals, therefore we must choose a functional with which to calculate the initial orbitals. Unlike the optimizations of ωB97X-V and ωB97M-V we chose an established functional to calculate the initial orbitals, namely ωB97X.[31]

The fitting procedure starts with the basic functional containing all zeroth order B97

components, the long- and short-range HF exchange, and the D3(BJ) components. This functional is optimized using a (weighted) linear regression on the training set. All possible combinations of additional parameters are added to this initial functional and re-optimized. The fitted parameters and weighted root-mean-square error (RMSE) for both the training set and validation set are recorded for each combination. The combinations which minimize the error across both these sets are then taken to be the optimal result.

Once the optimal combination of parameters has been chosen, the training set energies are re-evaluated using orbitals calculated with this new fit and the parameters re-fitted to these new energies.

The weighting scheme used is the same as the one used to optimise ωB97M(2).[29] Each data-point in a given dataset is assigned an initial weighting $w_i = (N_i \overline{\Delta E_i})^{-1}$ where N_i is the total number of data-points in the dataset, and $\overline{\Delta E_i}$ is the root-mean-square of the reaction energies in the dataset i. For each data-type, the weights are then normalised by dividing all values by the minimum initial weight $w_{\min}^d$, and exponentiated so that they lie in between 1 and 2. During this calculation the AE18 dataset is included in the TCE data-type. Each data-type is then given a multiplicative weight f_d, as shown in table 1. For the RG10 dataset, the set is subdivided the set into attractive and repulsive interactions, with factors $f_d = 10000$ and $f_d = 1$ respectively.[29]

4   Computational Details

All calculations were performed in Entos Qcore unless otherwise noted.[49] Calculations use the Def2-QZVPPD basis set.[50,51] The density-fitting approximation is used to calculate Coulomb, exchange, and UW12 contributions using the Def2-universal-JKFIT basis set.[52] The three-electron term in the UW12 approximation is calculated using an RI-approximation which utilizes the Def2-QZVP-RI basis set.[53 54] Exchange-correlation contributions were calculated numerically using Neese fixed pruned quadrature grids,[55 56] which use Laqua partitioning.[57]

Results for all previously established functionals with the exception of UW12 functionals are taken from Ref. 5.

5 Results

In order to test the fitting method, the procedure was used to fit an ωB97X-D3(BJ) type hybrid (without UW12). Fitting all possible combinations of parameters and analyzing the weighted error in the validation set results in the best possible functional with 8 linear parameters (see figure 2). When the procedure was repeated with the UW12 components included, the optimal number of parameters was found to be around 15. However, the fitted B97 coefficients for these functional combinations were found to be significantly larger than those present in the non-UW12 fitted functionals. While this effect was significantly reduced if our search was limited to combinations where higher order B97 terms are only included if all lower order terms are already present for each series expansion – where no B97 components were skipped. The coefficients were still significantly greater than in the non-UW12 functional.

To eliminate this problem, it was decided to use regularization on the regression. This was done using ridge regression where an additional parameter α which adds a penalty for the size of coefficients equal to the sum of squares. This method is equivalent to ordinary least-squares regression for $\alpha = 0$. To determine the optimal value of α, we looked at the weighted RMSE in the validation as a function of α for a single combination at a time. Figure 1 shows the weighted RMSE for the training and validation sets as a function of α for the fifteen parameter combination in which no B97 parameters are skipped and produced the lowest error in the validation set with linear regression. The plot shows a minimum in the validation error for α in the range 1–10, while the training set error remains almost constant. It was decided to use a value of α in this range, with the final value of $\alpha = 4$ chosen.

Figure 2 shows the weighted RMSE in the validation set for different combinations. All
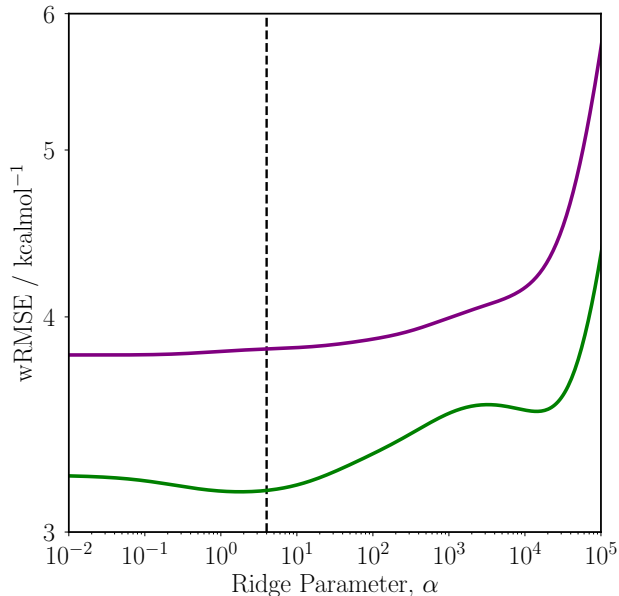
Figure 1: Weighted root-mean-square errors (wRMSEs) for the training (purple) and validation (green) sets as a function of the ridge regression parameter α. Errors are computed with the set of parameters used in the final functional, calculated using ωB97X orbitals. Dotted line shows at $\alpha = 4$ to demonstrate the final choice for α.

possible combinations were considered, though for ωB97X-osUW12-D3(BJ) only the lowest error combinations are plotted. Combinations which include no UW12 are also shown in the plot, these are also optimized using $\alpha = 4$, though the difference in these results is negligible compared to the standard linear fit.

This plot shows that the optimal number of parameters for ωB97X-osUW12-D3(BJ) is around 15. From this we investigated the lowest error functionals. We decided to use the lowest error 15 parameter functional for which no B97 components were skipped, which resulted in a final functional form containing the parameters: c_x^sr, $c_\mathrm{x}^0 = 1 - c_\mathrm{x}^\mathrm{sr}$, c_x^1, c_ss^0, c_ss^1, c_ss^2, c_ss^3, c_ss^4, c_os^0, c_os^1, c_os^2, s_8 as well as four osUW12 components c_u^{-4}, c_u^{-3}, c_u^{-2}, c_u^1 where coefficients c_u^n corresponds to the UW12 component Gaussian with exponent γ^n.

With this combination decided on, energies for all molecules in the training set were recalculated using the self-consistent orbitals obtained using this functional. Energies for all individual components were calculated using these orbitals and the coefficients re-fitted using this new data. This procedure was then repeated multiple times until the coefficients
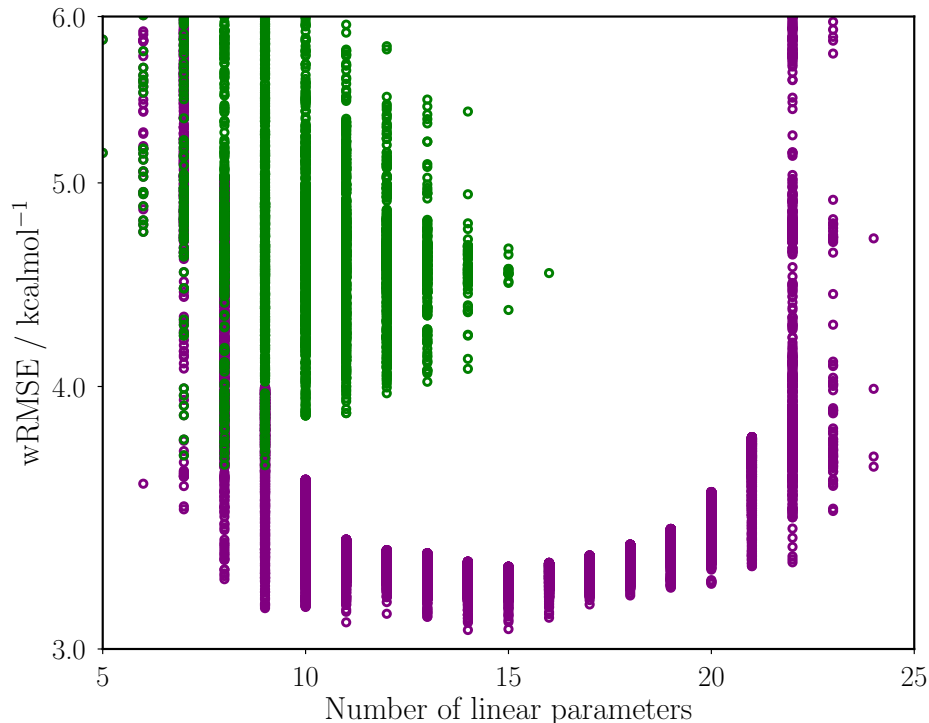
Figure 2: Weighted root-mean-square error (wRMSE) in the validation set as a function of the number of linear parameters. Each point represents a single combination of components. Only the best 1000 combinations are shown for each number of parameters. In each case coefficients were fitted to the training set data and the validation set errors recorded. All possible combinations of components were calculated for a functional with all ωB97X, osUW12 components and D3(BJ) (shown in purple). Optimized combinations containing no osUW12 contribution are shown in green.

converged. Table 2 shows the values of the coefficients at each cycle of the optimization. The zeroth cycle corresponds to the original fit using coefficients fitted using energies calculated on ωB97X orbitals. The final converged functional is shown in bold in the final column. We name this functional WM21-D3(BJ).

Throughout this optimization, a majority of the parameters remain mostly unchanged, with the difference between the final and initial values less than 0.05 in all but four cases. The c_u^{-4} value changes by 0.0554, while the first through fourth order same-spin B97 coefficients differ more significantly from the initial values than the other components. The greatest change is seen in the c_{ss}^1 coefficient (0.5958). The large change in these values is the reason for the numerous re-optimization cycles performed on the functional. Also of note is the

Table 2: Optimized coefficients for the ωB97X-osUW12-D3(BJ) functional at each stage of the optimization. The first set of coefficients are fitted using energies calculated using the ωB97X functional, subsequent fittings are performed on the training set data obtained using the functional defined in the previous column. The final WM21-D3(BJ) functional parameters are shown in bold.

	0	1	2	**3**
c_x^sr	0.2645	0.2658	0.2642	**0.2638**
c_x^0	0.7355	0.7342	0.7358	**0.7362**
c_x^1	0.7686	0.7679	0.7690	**0.7660**
c_ss^0	-0.1882	-0.2017	-0.1901	**-0.1891**
c_ss^1	1.4625	1.3045	1.2285	**1.2545**
c_ss^2	-0.6501	-0.2124	-0.0306	**-0.0543**
c_ss^3	-2.8600	-3.2029	-3.3474	**-3.3595**
c_ss^4	2.3564	2.4251	2.4468	**2.4690**
c_os^0	0.9544	0.9670	0.9676	**0.9673**
c_os^1	-1.3819	-1.4163	-1.4170	**-1.4147**
c_os^2	1.1881	1.1774	1.1875	**1.1877**
c_u^{-4}	-0.0913	-0.1307	-0.1324	**-0.1467**
c_u^{-3}	0.0836	0.1092	0.1030	**0.1108**
c_u^{-2}	-0.1944	-0.2016	-0.1973	**-0.1995**
c_u^1	-0.0051	-0.0040	-0.0038	**-0.0026**
s_6	1.0000	1.0000	1.0000	**1.0000**
s_8	-0.1016	-0.1146	-0.1179	**-0.1123**

small value for the c_u^1 coefficient.

As the optimization proceeds, the weighted root-mean-square error (wRMSE) in the training set is reduced from $3.832\,\mathrm{kcal\,mol^{-1}}$ in the initial cycle to $3.692\,\mathrm{kcal\,mol^{-1}}$ in the final step. Results for most of the datasets improve with the re-optimization with the error in 17 of the 24 datasets reduced compared to the initial results. Of the remaining sets, the greatest increase occurs in the DBH24 diverse barrier height dataset where the RMSE increases by $0.08\,\mathrm{kcal\,mol^{-1}}$. Though the minimum RMSE observed for some sets is seen in the intermediate cycles. A summary of RMSEs for each training dataset at each stage of the optimization is presented in the supporting information.

Figure 3 shows the form of the UW12 geminal function present in the WM21-D3(BJ) functional as a function of inter-electronic distance r_{12}. We note a significant difference from the Slater-function used previously. Though while the decision to use a Slater-type geminal
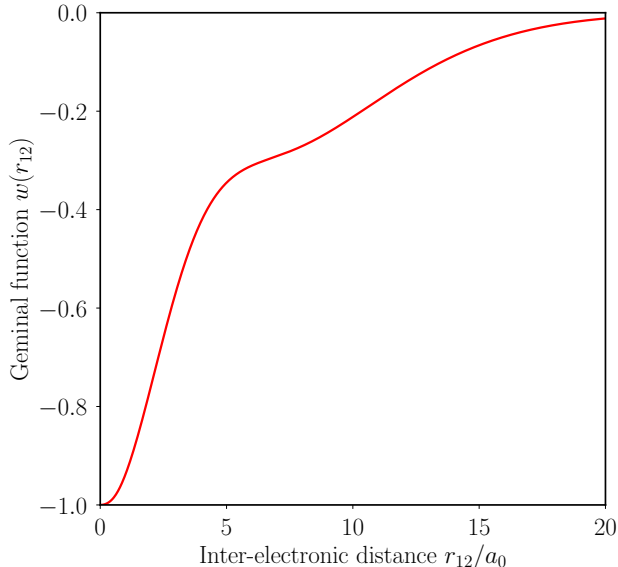
Figure 3: Plot showing the geminal function $w(r_{12})$ present in the final WM21-D3(BJ) functional. Scaled so that $w(0) = -1$.

function was made independent of the choice of the functional it was to be used in, the new geminal has been fitted in the presence of the other parts of the functional. In particular the long-range portion is optimized to work with the D3(BJ) dispersion correction, we have previously shown that this term can affect the range of the geminal function.[26]

Following the same procedure as for WM21-D3(BJ), we optimize the best standard hybrid functional of the type ωB97X-D3(BJ), looking at figure 2 it can be seen that the best possible hybrid contains 8 linear parameters, with the lowest error functional combination containing the coefficients c_x^{sr}, $c_x^0 = 1 - c_x^{sr}$, c_x^1, c_x^2, c_{ss}^0, c_{os}^0, c_{os}^1, c_{os}^2, and s_8. This combination has a weighted RMSE on the training and validation sets of $5.49\,\mathrm{kcal\,mol^{-1}}$ and $3.67\,\mathrm{kcal\,mol^{-1}}$ respectively. This combination of B97 components is slightly different than the one used in ωB97X-V, which does not include c_{os}^2, including c_{ss}^1 instead, this combination had weighted training and validation set RMSEs of $5.82\,\mathrm{kcal\,mol^{-1}}$ and $3.96\,\mathrm{kcal\,mol^{-1}}$ respectively. We refer to this functional as ωB97X21-D3(BJ).

As with WM21-D3(BJ) we re-optimize the coefficients after recalculating the energies in the training set self-consistently and repeating for multiple cycles until converged. Table 3 shows the coefficients for the hybrid functional at each stage of the procedure. Four of the

Table 3: Optimized coefficients for the ωB97X21-D3(BJ) functional at each stage of the optimization. The first set of coefficients are fitted using energies calculated using orbitals optimised with the ωB97X functional. Subsequent fittings are performed on the training set data obtained using the functional defined in the previous column. The final functional parameters are shown in bold.

	0	1	**2**
c_x^sr	0.1784	0.2183	**0.2164**
c_x^0	0.8216	0.7817	**0.7836**
c_x^1	0.7595	0.7911	**0.7916**
c_x^2	0.2515	0.0312	**0.0336**
c_ss^0	0.1775	0.0432	**0.0471**
c_os^0	1.2235	1.2744	**1.2728**
c_os^1	-2.3243	-2.7704	**-2.7611**
c_os^2	1.4293	2.0522	**2.0428**
s_6	1.0000	1.0000	**1.0000**
s_8	0.5272	0.4589	**0.4575**

coefficients undergo a significant change during the first optimization cycle, with c_os^1, c_os^2, c_ss^0, and c_x^2 all changing by more than 0.1, though these have converged after the second iteration. We note that the amount of same spin correlation contribution is greatly reduced by this re-optimization. Also we note that the non-linear parameters for both range-separation and dispersion are not re-optimized with the same ones used for both functionals.

Unlike for the UW12 functional, upon re-optimization, the weighted RMSE of the training set increases compared to the initial orbital version going from $5.49\,\text{kcal}\,\text{mol}^{-1}$ to $6.45\,\text{kcal}\,\text{mol}^{-1}$ in the final functional. A complete summary of the RMSEs for each dataset in the training data at each stage in the optimization of ωB97X21-D3(BJ) is shown in the supporting information. Of the 24 datasets only 9 have reduced errors on re-optimization, though 6 of these errors are reduced by more than $0.1\,\text{kcal}\,\text{mol}^{-1}$. Large increases of more than $0.1\,\text{kcal}\,\text{mol}^{-1}$ in the RMSE are observed in 10 of the sets, with the greatest increases being for the AE18 atomization energies set, the CRBH20 barrier heights set and the G21EA electron affinity sets which all increase by more than $0.5\,\text{kcal}\,\text{mol}^{-1}$. The RMSE for the atomization energy set AE18 increases by $2.19\,\text{kcal}\,\text{mol}^{-1}$, by far the greatest increase. The cause of this increase is most likely due to the small amount of same-spin correlation in the resulting functional,

	B3LYP-D3(BJ)	CAM-B3LYP-D3(BJ)	M11-D3(BJ)	wB97X-D3	wB97X-V	wB97X21-D3(BJ)	B-LYP-osUW12-D3(BJ)	WM21-D3(BJ)
3B-69-DIM	0.22	0.31	0.36	0.20	0.14	0.30	0.24	0.13
3B-69-TRIM	0.59	0.81	0.76	0.49	0.31	1.19	0.75	0.34
AlkBind12	0.14	0.15	0.53	0.28	0.10	1.34	0.23	0.17
BSR36	3.90	4.26	1.31	5.14	2.55	0.89	0.46	0.28
C20C24	27.20	18.83	18.17	9.07	4.31	6.89	5.52	2.92
CE20	1.62	3.27	1.05	0.63	0.38	0.58	0.59	0.45
CO2Nitrogen16	0.06	0.17	0.29	0.45	0.09	0.52	0.32	0.18
H2O16Rel5	0.64	0.82	1.11	0.11	0.30	0.53	0.31	0.03
H2O20Bind10	8.84	22.12	6.03	6.36	1.17	0.96	7.93	1.06
H2O20Bind4	11.90	23.45	5.76	5.68	1.85	0.59	3.79	3.20
H2O20Rel10	0.23	0.37	0.70	0.27	0.07	0.13	0.17	0.52
H2O20Rel4	0.66	0.86	0.87	0.33	0.22	0.53	0.22	0.08
HB49	0.42	0.69	0.42	0.22	0.20	0.25	0.23	0.16
HNBrBDE18	4.15	4.79	1.29	3.86	2.35	1.21	2.00	1.69
Ionic43	0.64	0.94	1.09	0.71	0.61	1.04	1.67	0.81
Melatonin52	0.27	0.32	0.91	0.13	0.10	0.43	0.17	0.13
PlatonicHD6	2.98	1.22	10.77	6.47	4.78	1.45	1.70	2.12
PlatonicID6	7.74	9.65	6.94	7.47	3.94	3.93	2.07	1.70
PlatonicIG6	19.77	27.82	25.31	19.60	5.49	7.99	5.29	10.46
PlatonicTAE6	14.32	3.25	17.94	4.13	7.27	5.14	21.75	1.67
PX13	5.93	8.06	3.23	2.19	2.84	2.35	1.55	2.52
WCPT27	3.49	3.43	1.98	1.59	1.67	1.89	1.63	1.73
WCPT6	0.99	0.89	1.43	0.97	0.88	0.99	0.90	0.86
YMPJ519	0.37	0.34	0.67	0.35	0.24	0.35	0.27	0.19

with the amount of same-spin correlation greatly reduced in the orbitals compared to the initial ωB97X orbitals.

In order to assess the accuracy of our new functionals we compare the errors for the datasets in the testing set data, this was not utilized during the training and validation of the functionals. We compare to other (range-separated) hybrid functionals, in particular the ωB97X-V functional optimized using the same method.[14–16,27,34,40,58–60] Also shown is data from our previous functional B-LYP-osUW12-D3(BJ).

Table 4 shows the mean absolute errors for each dataset in the MGCDB84 testing set. Beginning with the ωB97X21-D3(BJ) functional, it can be easily observed that this functional is outperformed by ωB97X-V. However, it does give improved results for 8 of the 24 datasets, notably the $(H_2O)_{20}$ binding energy sets (H2O20Bind10 and H2O20Bind4), and the N−Br bond dissociation set (HNBrBDE18) where the functional results in the lowest overall errors. The largest errors for this functional are seen for the binding energy datasets – Alk-

Bind12, 3B-69-DIM, 3B-69-TRIM, and CO2Nitrogen16 – for which errors are significantly greater than ωB97X-V. Overall this functional demonstrates similar errors to the ωB97X-D3 functional, with lower errors for 10 of the datasets. Notable improvements over ωB97X-D3 are seen in the hydrocarbon bond-separation set (BSR36), as well as for three of the datasets containing reactions involving platonic hydrocarbon cages (PlatonicHD6, PlatonicID6, PlatonicIG6). This functional could be further improved by re-optimizing the D3(BJ) damping parameters or by replacing D3(BJ) with the non-local VV10 functional as in ωB97X-V.

The B-LYP-osUW12-D3(BJ) functional gives lower errors than ωB97X-V for 7 of the 20 datasets, for the platonic hydrocarbon cages dataset (PlatonicIG6) as well as proton exchange barrier height set (PX13), B-LYP-osUW12-D3(BJ) results in the smallest overall errors. This functional produces significant errors for the binding energies of ionic dimers (Ionic43) and total atomization energies (PlatonicTAE6) sets; the second of which is much larger than for any other functional considered.

The new WM21-D3(BJ) functional represents a significant improvement over B-LYP-osUW12-D3(BJ) for these sets, with lower errors for 19 of the 24 datasets. The two largest errors for B-LYP-osUW12-D3(BJ) are significantly reduced; for PlatonicTAE26, where B-LYP-osUW12-D3(BJ) resulted in the largest error, WM21-D3(BJ) gives the lowest error. The functional gives lower errors than ωB97X-V for 14 of the 24 sets, with large improvements seen for isomerization energies (C20C24), bond separation energies (BSR36), total atomization energies (PlatonicTAE6) and reactions involving hydrocarbon cages (PlatonicID6). This functional gives the lowest error for 10 of the 24 datasets. The only significantly large errors of note are seen in the PlatonicIG6 dataset of isogyric reaction energies, as well as in two of the $(H_2O)_{20}$ datasets, namely H2O20Bind4 and H2O20Rel10. The error observed in the H2O20Rel10 dataset is significantly greater than the other ωB97 functionals, though this remains within chemical accuracy. Also for the H2O20Bind10 set of binding energies which uses the same set of low energy $(H_2O)_{20}$ isomers, WM21-D3(BJ) results in one of the lowest errors. Counter-intuitively the opposite is true for the datasets with the higher energy

Table 5: Optimized coefficients for the ωB97X-MP2-D3(BJ) functional optimized using ωB97X orbitals.

c_x^{sr}	0.4681	c_{ss}^0	-0.1545	c_{os}^0	0.6354
c_x^0	0.5319	c_{ss}^1	2.0088	c_{os}^1	-0.0551
c_x^1	0.6052	c_{ss}^2	-2.0638	c_{os}^2	-0.6478
c_{os}^{MP2}	0.3108	s_6	0.3738		
c_{ss}^{MP2}	0.1719	s_8	0.5478		

$(H_2O)_{20}$ isomers; while WM21-D3(BJ) results in the lowest error for the isomerization energies (H2O20Rel4), a much greater error is observed for the binding energies (H2O20Bind4). Based on the comparison of ωB97X21-D3(BJ) and ωB97X-V, we note that WM21-D3(BJ) could be further improved by utilizing the VV10 non-local correlation functional instead of D3(BJ).

As an additional test of this functional, using the same procedure as ωB97X21-D3(BJ) and WM21-D3(BJ), we optimize a double hybrid density functional of the form ωB97X-MP2-D3(BJ). For this we allow the coefficients of opposite-spin and same-spin MP2 to be free parameters. In addition we let the s_6 coefficient in D3(BJ) be free to account for the long-range behavior of MP2. As with ωB97X21-D3(BJ), we use the same non-linear parameters as WM21-D3(BJ) and use ωB97X orbitals to calculate the energies. Following the procedure used previously and taking the combination which minimizes the weighted validation error with the minimum number of parameters without skipping any B97 terms. The resulting functional is a 12 parameter spin component scaled double hybrid. Unlike the other functionals we do not re-optimize this functional to work with self-consistent orbitals and use ωB97X orbitals for the final functional. Table 5 shows the coefficients in the ωB97X-MP2-D3(BJ) functional. The weighted RMSEs for this functional for the training and validation sets are $3.58\,\text{kcal}\,\text{mol}^{-1}$ and $2.48\,\text{kcal}\,\text{mol}^{-1}$ respectively. The validation set value is lower than the wRMSE for the initial WM21-D3(BJ) for which the training and validation errors are $3.83\,\text{kcal}\,\text{mol}^{-1}$ and $3.17\,\text{kcal}\,\text{mol}^{-1}$ respectively. However, this is before the parameters are re-optimized for self-consistent orbitals.
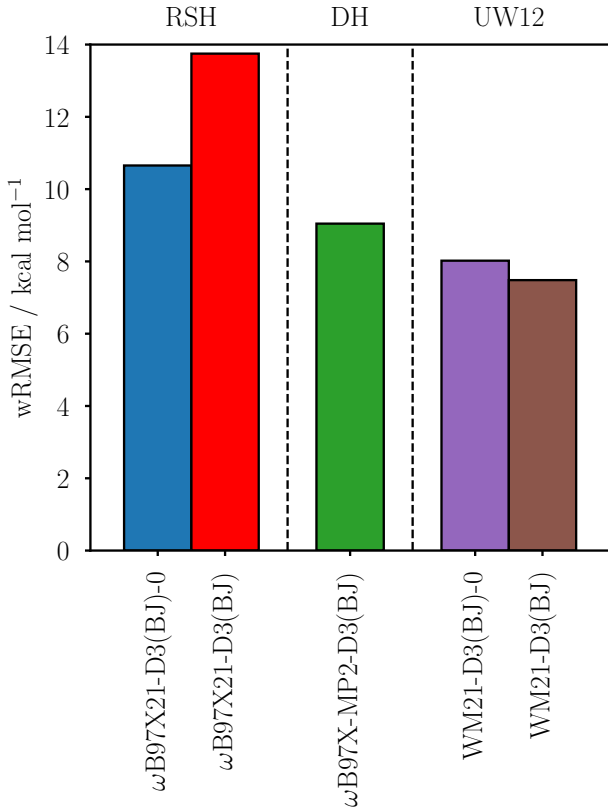
Figure 4: Weighted root-mean-square-error (wRMSE) in kcal mol^{-1} for the test set data for each of the functionals optimized in this paper. This includes the range-separated hybrid (RSH) ωB97X21-D3(BJ), the double hybrid (DH) ωB97X-MP2-D3(BJ), and the UW12 hybrid WM21-D3(BJ). Both ωB97X21-D3(BJ), and WM21-D3(BJ) refer to the final self-consistent versions of the functionals, while ωB97X-MP2-D3(BJ) is calculated using ωB97X orbitals. For comparison, the results for the initial fitted version of ωB97X21-D3(BJ) and WM21-D3(BJ) calculated using the ωB97X orbitals are also included. These are given the suffix '-0' to differentiate them from the final functionals.

We compare this functional to the others we have optimized in this paper for the test set data. Figure 4 shows the weighted root-mean-square-errors (wRMSEs) for the testing set data for each of the functionals fitted in this paper, with both the final self-consistent and initial ωB97X orbital versions included for ωB97X21-D3(BJ) and WM21-D3(BJ).

For ωB97X21-D3(BJ), the re-optimization results in an significantly increased error compared to the initial functional fit. This is consistent with what was observed for the training set data. The double hybrid functional ωB97X-MP2-D3(BJ) produces a lower overall error than the hybrid ωB97X21-D3(BJ), including the initial optimized version. However, both

versions of WM21-D3(BJ) result in a smaller wRMSE than the double hybrid functional for this set, with the fully-self-consistent WM21-D3(BJ) producing the lowest wRMSE value for these functionals.

6   Conclusions

We have presented WM21-D3(BJ) – a new range-separated hybrid functional with opposite-spin UW12 correlation and D3(BJ) dispersion which we have optimized using the survival of the fittest strategy on the main-group chemistry database (MGCDB84).[27–29] We have compared this functional to other range-separated hybrid functionals, including the ωB97X-V functional optimized using the same approach, and shown that for the test set data not used in the optimization procedure WM21-D3(BJ) results in lower errors for the majority of datasets considered without the VV10 non-local functional. The functional also offered significant improvement over the B-LYP-osUW12-D3(BJ) functional for a number of datasets.

We also optimized a standard range-separated hybrid ωB97X21-D3(BJ) using the same approach, this functional produced greater errors than ωB97X-V, despite containing similar components (with the exception of the dispersion component). The WM21-D3(BJ) functional results in much lower errors than ωB97X21-D3(BJ), showcasing the ability of UW12 correlation to improve the accuracy of a calculation.

In addition we compared results for WM21-D3(BJ) to a double hybrid functional ωB97X-MP2-D3(BJ) optimized using the same strategy, but without self-consistent optimization. We showed that the ωB97X-MP2-D3(BJ) double hybrid fitted to data calculated using ωB97X orbitals produces a greater overall error for the test set data than the initial version of WM21-D3(BJ) calculated in the same way. In addition, by using the self-consistent optimization procedure, the final WM21-D3(BJ) functional results in even lower errors for this set; demonstrating the advantage of the fully self-consistent UW12 approximation.

Overall, UW12 functionals allow increased accuracy compared to standard hybrid func-

tionals in a similar manner to double hybrids, but without using the virtual orbitals. In addition, the fully self-consistent nature of UW12 avoids a number of the problems associated with double hybrid functionals. The ability to capture correlation effects neglected by standard DFT functionals using only the occupied orbitals and a simple geminal function is both useful and theoretically interesting.

In future the functional could be re-optimized using the VV10 non-local correlation instead of D3(BJ) as this has been shown to offer great improvement for the standard hybrid ωB97X-V (compared to ωB97X21-D3(BJ)). In addition, following on from ωB97X-V the survival of the fittest strategy has also been used to optimize functionals with meta-GGA contributions such as ωB97M-V and the double hybrid ωB97M(2) – this could be added to a UW12 hybrid to further increase accuracy. In addition, while we have neglected same-spin UW12 contributions in this paper, a same-spin UW12 geminal could be optimized in a similar way to the opposite-spin UW12 with a separate geminal function for each spin pairing.

Supporting Information Available

The following files are included for additional information:

- `supplementary_information.pdf`: Information on the datasets in the main group chemistry database (MGCDB84)

- `supplementary_information.xlsx`: Tables detailing the individual energy contributions for each component used to optimize the UW12 functionals for all datasets and molecules included. In addition to Calculated energies for the final self-consistent functionals and a summary of functional coefficients.

Conflict of interest

FRM is Cofounder and CTO of Entos Inc.

Acknowledgement

References

(1) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.

(2) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(3) Goerigk, L.; Grimme, S. A Thorough Benchmark of Density Functional Methods for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670.

(4) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. 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.

(5) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(6) Chan, B.; Gill, P. M. W.; Kimura, M. Assessment of DFT Methods for Transition Metals with the TMC151 Compilation of Data Sets and Comparison with Accuracies for Main-Group Chemistry. *J. Chem. Theory Comput.* **2019**, 3610–3622.

(7) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(8) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(9) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(10) Perdew, J. P.; Kurth, S.; Zupan, A.; Blaha, P. Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation. *Phys. Rev. Lett.* **1999**, *82*, 2544–2547.

(11) Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Many-Electron Self-Interaction Error in Approximate Density Functionals. *J. Chem. Phys.* **2006**, *125*, 201102.

(12) Almbladh, C. O.; Von Barth, U. Exact Results for the Charge and Spin Densities, Exchange-Correlation Potentials, and Density-Functional Eigenvalues. *Phys. Rev. B* **1985**, *31*, 3231–3244.

(13) Levy, M.; Perdew, J. P.; Sahni, V. Exact Differential Equation for the Density and Ionization Energy of a Many-Particle System. *Phys. Rev. A* **1984**, *30*, 2745–2748.

(14) Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(15) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(16) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Fukuda, R. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(17) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **1996**, *105*, 9982–9985.

(18) Gill, P. M. W.; Adamson, R. D.; Pople, J. A. Coulomb-attenuated exchange energy density functionals. *Mol. Phys.* **1996**, *88*, 1005–1009.

(19) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

(20) Facco Bonetti, A.; Engel, E.; Schmid, R. N.; Dreizler, R. M. Investigation of the Correlation Potential from Kohn-Sham Perturbation Theory. *Phys. Rev. Lett.* **2001**, *86*, 2241–2244.

(21) Paier, J.; Janesko, B. G.; Henderson, T. M.; Scuseria, G. E.; Grüneis, A.; Kresse, G. Hybrid functionals including random phase approximation correlation and second-order screened exchange. *J. Chem. Phys.* **2010**, *132*, 094103.

(22) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(23) Goerigk, L.; Grimme, S. Double-hybrid density functionals. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 576–600.

(24) Martin, J. M. L.; Santra, G. Empirical Double-Hybrid Density Functional Theory: A 'Third Way' in Between WFT and DFT. *Isr. J. Chem.* **2020**, *60*, 787–804.

(25) Wiles, T. C.; Manby, F. R. Wavefunction-like Correlation Model for Use in Hybrid Density Functionals. *J. Chem. Theory Comput.* **2018**, *14*, 4590–4599.

(26) Williams, Z. M.; Wiles, T. C.; Manby, F. R. Accurate Hybrid Density Functionals with UW12 Correlation. *J. Chem. Theory Comput.* **2020**, *16*, 6176–6194.

(27) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904.

(28) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(29) Mardirossian, N.; Head-Gordon, M. Survival of the most transferable at the top of Jacob's ladder: Defining and testing the ω B97M(2) double hybrid density functional. *J. Chem. Phys.* **2018**, *148*, 241736.

(30) Becke, A. D. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560.

(31) Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

(32) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.

(33) Mardirossian, N.; Head-Gordon, M. Survival of the Most Transferable at the Top of Jacob's Ladder: Defining and Testing the ω B97M(2) Double Hybrid Density Functional. *J. Chem. Phys.* **2018**, *148*, 241736–241736.

(34) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(35) Becke, A. D. Density Functional Calculations of Molecular Bond Energies. *J. Chem. Phys.* **1986**, *84*, 4524–4529.

(36) Ten-no, S. Initiation of explicitly correlated Slater-type geminal theory. *Chem. Phys. Lett.* **2004**, *398*, 56–61.

(37) Grimme, S. Density Functional Theory with London Dispersion Corrections. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211–228.

(38) Klimeš, J.; Michaelides, A. Perspective: Advances and Challenges in Treating van Der Waals Dispersion Forces in Density Functional Theory. *J. Chem. Phys.* **2012**, *137*, 120901.

(39) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.

(40) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(41) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, 034112.

(42) Becke, A. D.; Johnson, E. R. A density-functional model of the dispersion interaction. *J. Chem. Phys.* **2005**, *123*, 154101.

(43) Johnson, E. R.; Becke, A. D. A post-Hartree–Fock model of intermolecular interactions. *J. Chem. Phys.* **2005**, *123*, 024101.

(44) Johnson, E. R.; Becke, A. D. A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. *J. Chem. Phys.* **2006**, *124*, 174104.

(45) Chakravorty, S. J.; Gwaltney, S. R.; Davidson, E. R.; Parpia, F. A.; Fischer, C. F. Ground-State Correlation Energies for Atomic Ions with 3 to 18 Electrons. *Phys. Rev. A* **1993**, *47*, 3649–3670.

(46) Tang, K. T.; Toennies, J. P. The van Der Waals Potentials between All the Rare Gas Atoms from He to Rn. *The Journal of Chemical Physics* **2003**, *118*, 4976–4983.

(47) Najibi, A.; Goerigk, L. The Nonlocal Kernel in van Der Waals Density Functionals as an Additive Correction: An Extensive Analysis with Special Emphasis on the B97M-V and ωB97M-V Approaches. *J. Chem. Theory Comput.* **2018**, *14*, 5725–5738.

(48) Entos-Qcore uses the LIBXC exchange-correlation functional library,[61,62] which calculates these expansions up to fourth order.

(49) Manby, F. R.; Miller III, T. F.; Bygrave, P.; Ding, F.; Dresselhaus, T.; Batista-Romero, F.; Buccheri, A.; Bungey, C.; Lee, S.; Meli, R.; Miyamoto, K.; Steinmann, C.; Tsuchiya, T.; Welborn, M.; Wiles, T. C.; Williams, Z. M. 2019, DOI: 10.26434/chemrxiv.7762646.v2.

(50) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys.* **2010**, *133*, 134105.

(51) Weigend, F.; Furche, F.; Ahlrichs, R. Gaussian basis sets of quadruple zeta valence quality for atoms H-Kr. *J. Chem. Phys.* **2003**, *119*, 12753–12762.

(52) Weigend, F. Hartree-Fock exchange fitting basis sets for H to Rn. *J. Comput. Chem.* **2008**, *29*, 167–175.

(53) Hättig, C. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core-valence and quintuple-ζ basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.

(54) The three-electron term uses an ABS+ type RI approximation, in which the identity operator is formed from the union of the atomic orbital basis set and an auxiliary RI basis sets.[63] All UW12 calculations utilise the Def2-QZVP-RI basis set except for the molecules in the H2O20Bind10/H2O20Rel10 set where Def2-SVP-RI was used as the auxiliary basis.

(55) These grids were used in Entos Qcore with index 4.

(56) Krack, M.; Köster, A. M. An adaptive numerical integrator for molecular integrals. *J. Chem. Phys.* **1998**, *108*, 3226–3234.

(57) Laqua, H.; Kussmann, J.; Ochsenfeld, C. An improved molecular partitioning scheme for numerical quadratures in density functional theory. *J. Chem. Phys.* **2018**, *149*, 204111.

(58) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation. *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817.

(59) Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. Revised Damping Parameters for the D3 Dispersion Correction to Density Functional Theory. *J. Phys. Chem. Lett.* **2016**, *7*, 2197–2203.

(60) Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. Long-Range Corrected Hybrid Density Functionals with Improved Dispersion Corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263–272.

(61) Marques, M. A.; Oliveira, M. J.; Burnus, T. Libxc: A library of exchange and correlation functionals for density functional theory. *Comput. Phys. Commun.* **2012**, *183*, 2272–2281.

(62) Lehtola, S.; Steigemann, C.; Oliveira, M. J.; Marques, M. A. Recent developments in libxc — A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1–5.

(63) Valeev, E. F. Improving on the resolution of the identity in linear R12 ab initio theories. *Chem. Phys. Lett.* **2004**, *395*, 190–195.