The exchange energy of a uniform electron gas experiencing a new, flexible range separation

John A. Parkhill a,b, Jeng-Da Chai a,b,c, Anthony D. Dutoi d, Martin Head-Gordon a,b,*

a Department of Chemistry, University of California, Berkeley, CA 94720, United States
b Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States
c Department of Physics, National Taiwan University, Taipei 10617, Taiwan
d Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States
* Institute of Physical and Theoretical Chemistry, Ruprecht-Karls-University Heidelberg, Im Neuenheimer Feld, 229, 69120, Germany

A R T I C L E   I N F O

Article history:
Received 12 April 2009
In final form 16 July 2009
Available online 19 July 2009

1. Introduction

Despite rough beginnings [1], the local density approximation (LDA) has been developed through decades of work on the Kohn–Sham (KS) [2] construction into one of the most successful approximations in quantum chemistry and solid state physics. Within this framework, the LDA exchange correlation functional is combined by an adiabatic connection with a non-interacting wavefunction so that an approximate kinetic energy may be extracted and there is no need to develop accurate functionals for the kinetic energy [3] which have proven elusive. Along these lines, Becke [4] realized that the accuracy of Kohn–Sham energy functionals could be improved by the admixture of ‘exact’ exchange coming from the explicit exchange energy of the fictitious Kohn–Sham wavefunction. The resulting hybrid density functionals have been the most commonly applied model chemistry for many years [5] because they have been found to be remarkably accurate with computational costs virtually equivalent to those of the Hartree–Fock (HF) method.

One of the few remaining substantial defects of the Kohn–Sham construction which has attracted theoretical effort is the so-called self-interaction problem [6–12], and it is directly related to the treatment of the exchange energy [13]. In the HF energy expression the Coulomb repulsion of a one-electron function with itself is cancelled exactly by the corresponding exchange integral. In the KS construction with a pure, local functional the Coulomb energy is non-local, but the exchange energy is not. Considering the one-particle functions provided by the KS wavefunction we might say that the electron repels itself if the particle is spread over space because the antisymmetric complement of the Coulomb interaction, non-local exchange, is missing. At equilibrium geometries the effect on predicted ground state energies is not severe, but this defect means that dissociation problems may lead to fragments which only possess a fractional number of electrons [14,15], or response properties which reflect serious artifacts if charges are significantly redistributed [16]. If globally a fraction of the exchange energy of the KS determinant is mixed with the DFT exchange energy these artifacts are partially remediated. To a stranger unfamiliar with the history of hybrid DFT’s development the situation must seem confusing, because it is not obvious why any mixture of ‘exact’ exchange with (semi-)local exchange is advantageous. The answer is that the accuracy of most DFT functionals lies in a cancellation of errors between exchange and correlation functionals. Both are compensating for the single-reference nature of the fictitious KS determinant [17,18]. From another angle, one might say that these two non-local objects [19] are best considered together because what results is more local.

A way to preserve the local cancellation of errors yet recover correct exchange at long range has emerged in the form of range-separated hybrid functionals. The idea which goes back to the pioneering work of Gill and Savin [20–22], is to divide 1/r by multiplying it with a function which varies between 0 and 1, such that both this function and its complement are integrable. The greatest fraction of work in this very active area [23–28] has employed the standard error function to achieve this separation:
(1)

The LDA exchange functional corresponding to $\text{erf}(or)/r$, and integral kernel for the exact exchange over $\text{erf}(or)/r$ can then be derived so that locally exchange is provided by the LDA and at a distance exchange is provided by the KS wavefunction. The position where the transition is smoothly made between the two treatments is determined by the adjustable parameter $1/\omega$. The choice of the error function as a Coulomb attenuator is both practical (for most implementations one must be able to perform the integral of Gaussians over the function analytically [29]) and arbitrary because the error function is just one of many which possess this property. In some recent studies promising results have been attributed to more flexible range separation [30].

One can extend the idea of mixing ab initio and DFT strengths further by imagining range separation of the correlation part of the functional as well. In this scheme (which we will not pursue in this work beyond mention) the ab initio method is made responsible for static and long range dispersion effects while the LDA correlation functional is adjusted for the modified Coulomb interaction to avoid double counting. Savin and coworkers have experimented with the choice of another attenuated Coulomb interaction for these purposes [31], a linear combination of an error function and a Gaussian which offers a sharper separation:

$$v_{\text{erf gauss}} = \text{erf}(or)/r - (2\mu)/\sqrt{\pi} e^{-(1/3)\mu^2 r^2}. \quad (2)$$

More recently this erf–gau LDA functional was combined with a standard GGA in an attempt to surpass the accuracy of exchange-hybrid functionals based on erf [32] with a GGA correlation treatment. Another one-parameter attenuated Coulomb interaction, the Yukawa potential, has also been the subject of recent investigations [33,34]

$$v_{\text{erf yukawa}} = \exp(-r^2)/r. \quad (3)$$

Improved performance of the resulting functional was attributed to an increased fraction of short-range exchange [35].

Our group has recently published an analytical integral over a more general sort of Coulomb attenuator which allows for separate control of where and how rapidly the shift is made between parts of the Coulomb interaction [36]. It allows continuous variation of sharpness between the limits of erf and the Heavyside function. The function is a linear combination of two error functions (although note that the erf integral formulas are not sufficient to describe it) and so we have adopted the name ‘terf’:

$$\text{terf}_{\text{C0,or}}(r) = (1/2)\left(\text{erf}(or - or_0) + \text{erf}(or + or_0)\right). \quad (4)$$

Investigations into the performance of range-separated hybrids [37] have found that existing separations cannot simultaneously describe ground state and excited state properties with a single choice of $\omega$. Along these lines a common area of intersection has been located amongst many optimized attenuators at roughly 0.8 Bohr [38] in the $(r, \lambda_{\text{optimal}}(r))$ plane. The physical implication is that cancellation of GGA-exchange and GGA-correlation errors in this region is balanced with the error induced by semi-local exchange, and our attenuator should be shaped similarly in this region for thermochemical accuracy. Yet to repair self-interaction error the attenuator should reach its asymptote as rapidly as possible once we leave this region. The terf functional form can pass through this point while still reaching its asymptote more rapidly than erf, and so there is reason to hope that terf may be useful in this respect. Another nice feature of this choice of separation is that it reduces to the erf attenuator if the $r_0$ parameter is chosen to be zero. The attenuator controlling the fraction of exact exchange to include as a function of electronic separation is plotted for various choices of parameters in Fig. 1: it consists of terf itself as defined above, plus a fraction of short-range exact exchange (i.e. terfc) controlled by a third parameter, $c$:

$$L(r) = c_{\text{terf}}(or, or_0) + \text{terf}_{\text{C0,or}}(r). \quad (5)$$

The specific sets of parameters plotted in Fig. 1 are choices whose performance for chemical problems will be explored later in this paper. They all include roughly similar fractions of exact exchange in the mid-range region, as is believed to be important [38].

2. The terf-attenuated LDA

The exchange energy of a many-fermion system, charge balanced by a structureless positive background is our starting point. This matrix element is [39] where $\theta$ denotes the Heaviside function

$$E_x = \frac{-k^3}{12\pi^2} \int_0^\infty q^2 v_{\text{ee}}(q) \left(1 - \frac{3}{2}x^2 + \frac{1}{2}x^3\right)\theta(1 - x) dq,$$

where $k_i = (3\pi^2 n)^{1/3}, x = q/(2k_i)$. So we must obtain the Fourier transform of $\text{terf}_{\text{C0,or}}(r)/r$:

$$\mathcal{F}\left[\text{terf}_{\text{C0,or}}(r)/r\right] = v_{\text{ee}}(q) = \frac{4\pi\left(1 - e^{-\frac{q^2}{2\omega^2}}\cos\left(\frac{qor_0}{\omega}\right)\right)}{q^2} \quad (7)$$

The integration of this function is algebraically quite tedious, but can be done. Unfortunately the complex error function enters. Note that for $z \in \mathbb{C}$, $\text{erf}(z^*) = \text{erf}(z^*)$ and for $z \in \mathbb{R}$, $\text{erfi}(z) \in \mathbb{R}$. We report the exchange energy per particle of the Fermion gas experiencing this interaction, $E_x$, which may be readily implemented in any KS-DFT code.

$$E_x = \int n(R)e_x(n(R))dR. \quad (8)$$

The spinless $k_i$ can be easily replaced to obtain the spin-density functional

$$e^{\pm}(n) = \frac{\omega^3}{192\pi^3} \left\{8Ae^{\frac{-q^2}{2\omega^2}}\sqrt{\pi}\left[Aor_0\left(3 + A^2\left(4A^2\omega^2r_0^2 - 6\right)\right)\right] + \text{Im}\left[\text{erf}\left(\frac{1}{2A} - ior_0\right)\right] + \text{Re}\left[\text{erf}\left(\frac{1}{2A} - ior_0\right)\right] \right\} + \left\{-\left(3 + 2A^2 + 3A^4(\omega^2r_0^2 - 1)\right) + e^{\frac{-q^2}{2\omega^2}}16A^4\left(1 + 2A^2(\omega^2r_0^2 - 1)\right)\cos\left(\frac{qor_0}{\omega}\right) + Aor_0\sin\left(\frac{qor_0}{\omega}\right)\right\}. \quad (9)$$

One may easily verify that this expression matches the known erf expressions for $e_x$ as $r_0 \to 0$ [20,40]. For large values of $A (>1)$, a
series expansion in powers of $1/A$ is employed up to 10th order in our implementation for purposes of numerical stability.

3. Application to range-separated hybrids

Without semi-local gradient information the thermochemistry of this functional would be undoubtedly poor and it would be difficult to determine if terf could improve functionals in use today. There are several recipes for combining this LDA exchange functional with a GGA enhancement factor ranging in degrees of technical difficulty and empiricism. Ideally the GGA factor will depend on the attenuating parameters [41], but recent results have shown that superior accuracy [13] can be obtained even if this is only done implicitly through optimized parameters of the GGA. At the end of the day the choice of $\omega$ parameter is quite empirical, as will be $r_0$, even if we introduce them for physical reasons. The final measure of a range-separated hybrid is optimization over a large training set, and evaluation over an independent test set, roughly a year of computer effort. We seek some justification for such an effort and so we combine the terfc-LDA exchange energy with the GGA exchange enhancement factor and correlation functionals of $\omega$B97X [13] and run some basic tests to establish whether the resulting functional shows promise. To be precise, the resulting functional is obtained directly from $\omega$B97X by replacing the $F(a_\nu)$ of Eq. (7) in that paper with the corresponding terfc-LDA $F(a_\nu)$ obtained from Eq. (8). To begin from a functional as close as possible to the parent (see the previous paper to clarify the notation), we also incorporate a variable fraction of short-range HF exchange in such a way that the UEG limit is respected (Eqs. (9) and (10))

$$E_{x}^{\text{SR—DFA}} = \sum_{g} \int \frac{\rho_{x}}{\rho_{g}} r_{g}^{\text{SR—DFA}}(s_{r}^{2}) dr.$$  \hspace{1cm} (10)

$$E_{0} = E_{x}^{\text{SR—HF}} + c_{x} E_{x}^{\text{SR—HF}} + (1 - c_{x}) E_{x}^{\text{SR—DFA}} + E_{x}^{\omega B97X}.$$  \hspace{1cm} (11)

Aside from the many parameters associated with the GGA we must choose reasonable guesses of $\{r_{0}, \omega, c_{x}\}$. The physically motivated guess is to reach the asymptote as rapidly as possible while still overlapping significantly with the established attenuators in the previously mentioned critical region [38]. An initial choice of parameters $\{r_{0}, \omega, c_{x}\} = [1.2, 1.016, c_{x}^{\omega B97X}]$ was made by this physically motivated criterion, and the usefulness of the resulting functional was assessed on noble-gas dimer cation dissociation (Figs. 2 and 3). Even with only very conservative changes made to the functional form of the attenuator, terf was able to significantly increase the accuracy (relative to its predecessor $\omega$B97X) of the dissociation asymptote associated with the self-interaction problem (SIE) (Fig. 3). In the case of He$_{2}^+$ the valence density lies so close to the neighboring atom that it seems challenging to reach a compromise between thermochemistry and exact exchange within a transferable range-separated exchange functional but the results for Ar$_{2}^+$ are encouraging. The solid thermochemical performance of the parent functional seemed more-or-less conserved, and so we investigated a little further. We make the approximation that the GGA parameters are unchanged between erf ($\omega$B97X) and terf attenuated Coulomb interactions. Undoubtedly this should be improved upon in future work and the literature already describes many ways this may be done.

A few more sets of attenuator parameters were obtained by maximizing the least-squared overlap of a terf attenuator with those of $\omega$B97 and $\omega$B97X varying $r_{0}, c_{x}$ for a given $\omega$. Noble-gas dissociation curves and atomization energies were calculated for the standard G2 [44] thermochemical test set in the

<table>
<thead>
<tr>
<th>$\omega$ (a.u.)</th>
<th>$r_{0}$ (a.u.)</th>
<th>$c_{x}$</th>
<th>MAE</th>
<th>He$_{2}^+$ error</th>
<th>Ne$_{2}^+$ error</th>
<th>Ar$_{2}^+$ error</th>
<th>Kr$_{2}^+$ error</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLYP$^a$</td>
<td>-</td>
<td>-</td>
<td>2.09</td>
<td>-87.55</td>
<td>-80.37</td>
<td>-48.51</td>
<td>-49.0</td>
</tr>
<tr>
<td>0.3$^b$</td>
<td>0</td>
<td>0.1577</td>
<td>2.53</td>
<td>-35.1</td>
<td>-31.3</td>
<td>-9.7</td>
<td>-7.7</td>
</tr>
<tr>
<td>0.4$^b$</td>
<td>0</td>
<td>0</td>
<td>5.71</td>
<td>-31.7</td>
<td>-29.3</td>
<td>-6.7</td>
<td>-4.2</td>
</tr>
<tr>
<td>1.016</td>
<td>1.2</td>
<td>0.2395</td>
<td>4.02</td>
<td>-24.9</td>
<td>-24.7</td>
<td>-3.0</td>
<td>-1.5</td>
</tr>
<tr>
<td>1.4</td>
<td>1.345</td>
<td>0.248</td>
<td>5.77</td>
<td>-26.6</td>
<td>-26.2</td>
<td>-2.9</td>
<td>-1.2</td>
</tr>
<tr>
<td>2</td>
<td>1.329</td>
<td>0.316</td>
<td>9.16</td>
<td>-23.2</td>
<td>-21.9</td>
<td>-0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>0.98</td>
<td>0.138</td>
<td>7.38</td>
<td>-13.0</td>
<td>-16.2</td>
<td>-0.03</td>
<td>0.0</td>
</tr>
</tbody>
</table>

$^a$ $\omega$B97X.

$^b$ $\omega$B97.

$^c$ Pure Becke 88 exchange [42] and LYP [43] correlation, errors in this row are upper bounds.
6-311+G(3df,3pd) basis with a saturated quadrature grid. The purpose was not to search for parameters which would surpass \( \omega B97X \), because gradient corrections are vital to thermochemical accuracy, but rather to document the balance between thermochemistry and correction of the self-interaction error (Table 1). As expected the results were quite sensitive to the steepness of the attenuator and the amount of middle-range exchange, but note that the further this attenuator departs from erf the more severe becomes the GGA approximation. An accurate asymptote cannot be obtained by simply increasing the amount of exact exchange in the attenuator (because eventually the correlation part of the problem is disturbed), but by maximizing overlap with the established ones in the critical region we do obtain reduction of SIE with increasing ‘exact’ exchange. At the moment where the singly-occupied MO’s bond is breaking if this MO’s density around atom 2 is far enough from the bulk of this MO’s density around atom 1 so that it experiences Hartree exchange the density will localize on an atom (as it should physically). Smaller atoms cause greater difficulty in general depending specifically on shell structure. The results suggest that if the GGA enhancement factor of the functional were re-optimized for the modified attenuator it would possess thermochemistry much like \( \omega B97X \) with a significantly larger amount of exact exchange, and thus reduced self-interaction. Even in its current incarnation, the modified functional is accurate enough to be used in lieu of others for problems where ‘exact’ exchange might be important.

4. Discussion and conclusions

Owing largely to the popularity of hybrid functionals, range separation of exchange has become an intense area of research, and more flexible range separation may prove desirable [18]. Indeed this has already been done with the erf–gau type attenuator [32], although in this case this was done at the expense of abandoning a physically motivated choice of a parameter. An analytical formula [36] is available for the exact exchange energy with the terf attenuation, and this expression has already been efficiently implemented in the publicly available release of the Q-CHEM package [45]. This paper provides the other building block, the short-range LDA exchange energy and a proof-of-concept GGA functional. Preliminary results with unoptimized parameters indicate that the new functional may be a useful improvement and development should continue in the area of more general exchange attenuators. Further improvement over the functional developed here might be realized through complete reoptimization [13]. Alternatively one could derive the corresponding PBE type GGA-functional [46,47,41]. In either case, the path is clear and only limited by one’s curiosity. It will be especially interesting to see if the flexibility of the new attenuator can simultaneously describe ground state electronic structure and excited states. Special attention should be paid to the size of the chromophore relative to the scale of the attenuator, and the distance over which electron density is redistributed. This direction is currently being pursued in our group.

Acknowledgements

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the US Department of Energy under Contract DE-AC0376SF00098, and by a Grant from the SDSC Program.

References