Van der Waals Density Functional Theory with Applications

D. C. Langreth,[1] M. Dion,[1] H. Rydberg,[2] E. Schröder,[2]
P. Hyldgaard,[2] and B. I. Lundqvist[2]

[1]Department of Physics & Astronomy and Center for Materials Theory,
Rutgers University, Piscataway, New Jersey 08854-8019

[2]Department of Applied Physics, Chalmers University of Technology
and Göteborg University, SE–412 96 Göteborg, Sweden

Abstract

The details of a density functional that includes van der Waals (vdW) interactions are presented. In particular we give some key steps of the transition from a form for fully planar systems [Phys. Rev. B **62**, 6997 (2000)] to a procedure for realistic layered compounds that have planar symmetry only on large distance scales, and which have strong covalent bonds within the layers. It is shown that the random-phase approximation of that original functional can be replaced by an approximation that is exact at large separation between vdW interacting fragments and seamless as the fragments merge. An approximation to the latter which renders the functional easily applicable and which preserves useful accuracy in both limits and in between is given. We report additional data from applications to forms of graphite, boron nitride, and molybdenum sulfide not reported in our previous communication [Phys. Rev. Lett. **91**, 126402 (2003)].

1 Introduction

The density-functional-theory (DFT) success story stretches over many decades. Today molecular and materials theories describe cohesion, bonds, structures, and other properties very well for densely-packed molecules and materials. However, sparsely-packed systems, including soft matter, van der Waals complexes, and biomolecules, are at least as abundant. The long-ranged forces between fragments across lower density regions are accounted

for within DFT, but not in local (LDA) no semilocal (GGA) approximations of the exchange-correlation (XC) functional. There are several approximate functionals proposed to account for effects in $E_{\rm xc}$, in particular for the asymptotic interaction between widely separated fragments [1, 2, 3, 4, 5, 6, 7]. Our current functionals (vdW-DF) approach the correct asymptotic dependence at large distances and are seamless at small distances. The first such form of the functional appropriate for layered systems has been applied to graphite, boron nitrate, and molybdenum sulfide [8, 9]. The characteristic result for the potential-energy surface of BN in Figure 1 illustrates what results can be obtained with a useful accuracy.

The plan of the paper is to review the considerations behind our vdW-DF, elaborate on several aspects that earlier have been only touched upon, present typical results from applications of vdW-DF, report on current work in progress, and look out over further developments. In Section 2 approximations for exchange are analyzed. Section 3 is devoted to correlation, in particular long-ranged effects, special features of the parallel configuration, the explicit assumptions of our nonlocal $E_{\rm xc}^{\rm nl}[n]$, and the saturation of this functional, providing a seamless connection as the fragments come close. A comment is also made on the local part of the correlation energy. In Section 4 exchange and correlation are put together, with an explicit account of our proposed density functional (vdW-DF) and some practical considerations for applying it. The applications to layered materials are reviewed in Section 5, and the paper ends with some conclusions and an outlook for the future.

2 Approximations for exchange

The van der Waals density functionals discussed later are approximations for the correlation energy alone. In order to make comparisons with experiment, we need to approximate the exchange energy as well. We choose to continue to use a standard GGA [10, 11, 12, 13] for exchange even though we replace GGA for long range correlation by our vdW functional. However, it is important to choose a GGA flavor whose exchange part does not produce a binding which is not present when the exchange is treated exactly.

The possibility for such an attraction can be discussed in terms of the behavior of the exchange energy $E_{\rm x}$ as a function of the reduced density gradient

$$s = \frac{|\nabla n|}{2k_{\rm F}n} \tag{1}$$

with $k_{\rm F}^3 = 3\pi^2 n$. The key point is that when two pieces of matter are brought close enough together to have a small overlap, the values of s are decreased in the overlap region, while the values of n are increased. The gradient dependent part of $E_{\rm x}$ typically becomes more negative as n increases, so the density overlap tends to produce binding. This can (and should) be negated by the s dependence, since the gradient part also typically becomes more negative as s increases, implying an antibinding effect from the smaller values of s in the overlap region. However many GGA's have exchange parts that for the very large values of s either fail to become more negative as s is further increased, and hence produce binding. The behavior of exchange functionals for such values of s can become important for some van der Waals systems.

Although it has sometimes been claimed that GGA's such as PW91 [10] can give van der Waals attraction, Wu et al. [14] have pointed out that most of this attraction comes from the exchange contribution. In particular, this GGA has been compared with exact Hartree Fock exchange in Figures 1a and 2a of that paper for the Ar and Kr dimers, respectively. Although this GGA for exchange alone gives binding very similar in magnitude and position to experiment for these two dimers, the exact exchange curve [14] shows no attraction whatever. Of course, van der Waals binding in reality is a correlation effect unrelated to exchange, and we must concur with the conclusion of Ref. [14] that this behavior is a coincidence. Such a conclusion is reinforced by the fact that for other dimers such as benzene [15] and a number of layered "dimers" such as graphene [8], boron nitride [9], and molybdenum sulfide [9], the exchange-driven GGA binding minimum occurs at a distance much larger than the physical van der Waals binding distance. Indeed any attempt go construct a GGA that scales properly to mimic the van der Waals interaction is undoubtedly doomed to failure, because there simply is not enough information in the small region of density overlap to provide a practical method of reconstructing the behavior of such a global property. In particular, atoms that have a strong influence on the vdW interaction are not all nearest neighbors with one on the other fragment, and hence may have a negligible effect on the overlap region.[1]

Because we wish to retain the success of PW91 [10] and similar GGA functionals [11, 12, 13] in describing strong covalent bonds, we seek a functional of the same type, but whose exchange part does not incorrectly give large binding in the van der Waals regime, and which gives the exchange

[1] An example would be an adatom a on a top site of a surface of atoms b in comparison with atom a the same distance from an otherwise isolated atom b.

part alone as accurately as possible. One of the above functionals [12] uses an exchange part with one of its parameter fitted to exact exchange calculations. This exchange functional of Zhang and Yang, $E_{\rm x}^{\rm ZY}$, like the others, saturates and fails to become more negative at sufficiently large s. However, in this case, the adverse consequences are much less than for the others, because here the saturation occurs at such a large value of s that the density is essentially zero, and the energy contribution small in magnitude; in short, the incorrect attraction here is negligible, and in any case its maximum is well separated from the van der Waals binding distance.

We illustrate the above points in Figure 2, which shows exchange calculations for the krypton dimer in the region of the binding distance. The figure shows the self consistent binding or interaction energy vs. separation distance as predicted by revPBE exchange and PW91 exchange in comparison with exact Hartree Fock exchange as calculated in Ref. [14]. One sees that the deep binding minimum of the PW91 exchange is absent for exact exchange and revPBE exchange. Since revPBE exchange does not show a tendency to bind any of the van der Waals systems we have tried it on, it provides a conservative starting point for adding vdW correlation functionals, thus guaranteeing that any binding we find will not be coming from exchange.

3 Approximation for correlation

Below, we give a brief account of the derivation of our van der Waals density functional theory. Like the earlier work directly calculating the nonlocal correlations between two jellium slabs [5], the theory [6] exploits assumed planar symmetry. We believe that the experience drawn from the planar case has a strong bearing on van der Waals density functionals for general geometries [15]. Moreover, the planar cases also serve as independent tests for this development work.

The theory [6] divides the correlation energy functional into two pieces

$$E_{\rm c}[n] = E_{\rm c}^{0}[n] + E_{\rm c}^{\rm nl}[n]. \tag{2}$$

$E_{\rm c}^{\rm nl}[n]$ is defined to include the longest ranged or most nonlocal terms that give the van der Waals interaction; it approaches zero in the limit of a slowly varying system. The first term $E_{\rm c}^{0}[n]$ is also nonlocal, but approaches the LDA in the limit of a slowly varying system. Different approximations are made for the two terms. In particular, we make the LDA for $E_{\rm c}^{0}[n]$, arguing that after separating off the principal longest range terms, the LDA will be

much more accurate. Ultimately we may derive gradient corrections to the LDA approximation to $E_c^0[n]$, but these are expected to be small, and in any case are not yet available. So in this work we use

$$E_c^0[n] \approx E_c^{\rm LDA}[n] \tag{3}$$

in (2). The principal argument for the simplification of the second term, $E_c^{\rm nl}[n]$, is that such long range terms are less sensitive to the details of the system's dielectric response than the short range terms, and that very simple approximations for the dielectric function in the long range terms may be made. An important feature of our approximation, however, is that it makes good predictions for the polarization properties of a single layer, as these are what determine the vdW interaction.

3.1 Correlation at long range

The exact correlation energy may be expressed by the adiabatic connection formula [16]

$$E_c = -\int_0^\infty \frac{du}{2\pi} \int_0^1 \frac{d\lambda}{\lambda} {\rm tr}\left[(1-\tilde{\chi}_\lambda V_\lambda)^{-1}\tilde{\chi}_\lambda V_\lambda - \tilde{\chi}_0 V_\lambda\right]. \tag{4}$$

Here $\tilde{\chi}$ is the density response to the full electric potential and V is the interelectronic Coulomb interaction, where we use an obvious matrix notation for the spatial variables, and u is the imaginary frequency replacing the time variable. The coupling constant λ multiplies e^2 wherever it appears explicitly or implicitly.

One obtains contributions to the long range vdW interaction by expanding the denominator in Eq. (4), as the Coulomb interaction can hop at long ranges between one chunk of matter and another. Another equally important contribution to the vdW interaction comes from what might be called *spectator* excitations on a distant fragment: even though $\tilde{\chi}$ may be short ranged with respect to its two spatial variables, its *value* can be affected by a concomitant excitation a long distance away. The role of the coupling constant integration with respect to long range vdW is to sort out the proper weights for these contributions. However, one gets the exact answer for the vdW asymptote if one (i) neglects the spectator excitations entirely *and* (ii) replaces $\tilde{\chi}_\lambda$ by $\tilde{\chi}_1$.[2] This allows the coupling constant integration

[2]This is because without spectator contributions, all the interfragment Coulomb potentials are shown explicitly in Eq. (4), which then can be expanded to second order in these couplings, obtaining the the same result as direct second order perturbation theory,

to be trivially performed.[3] We term this the *Full Potential Approximation* (FPA). It is an approximation, since it is only exact for the long range vdW asymptote. As discussed earlier, we divide the correlation energy into two parts. We only make this approximation in the second E_c^{nl}.

The FPA enables the expression of E_c^{nl} in terms of the electromagnetic response of the physical system. For systems with planar geometry [6], this takes on a particularly simple form.[4] Consider a system with planar symmetry in the xy plane that occupies the space $0 < z < d$, and place a sheet with an arbitrary surface charge distribution far to the right of it at $z = z_0$ ($z_0 \gg d$), whose Fourier transform with respect parallel wave vector k and frequency ω is $\sigma(\vec{k}, \omega)$. The (Fourier transformed) electric potential from this will take the form

$$\phi(z) = \begin{cases} a_L e^{kz} & z \ll 0 \\ a_R e^{kz} + b_R e^{-kz} & d \ll z < z_0, \end{cases} \tag{5}$$

where $\vec{k}$ and ω dependence of the quantities is implicit. One then uses the electrodynamic response functions of the system to solve the boundary value problem posed by Eq. (5), and the corresponding correlation energy is given by

$$E_c^{nl} = A \int_0^\infty \frac{du}{2\pi} \int \frac{d^2k}{(2\pi)^2} \ln \frac{a_R(\vec{k}, iu)}{a_L(\vec{k}, iu)}, \tag{6}$$

where A is the planar area. This gives the vdW asymptote exactly, but otherwise should be a reasonable approximation to E_c^{nl} via the FPA.

3.2 Asymptotic interaction between parallel two insulating layers

Since we wish to apply Eq. (6) to layered systems, we consider here the asymptotic form for the interaction of two layers centered at $z = 0$ and $z = d$

which is exact for finite fragments [17] or a fragment and an infinite surface [18]. The method is also exact for planar geometry of two parallel surfaces by extending the method used in Ref. [6]. Full details await a future publication [15].

[3] A corollary is that if one uses a simple coupling constant dependent model for $\tilde{\chi}_\lambda$ that does not include the spectator contributions, then it is simply wrong to make an effort to carry out the coupling constant integration, at least in order to obtain the long ranged interaction.

[4] Ref. [6] used the random phase approximation (RPA), but the considerations used there can equally well be applied to the FPA, which produces the exact vdW interaction at large distances.

respectively, for large d.[5] At large d it cannot matter for the interaction energy if we consider the layers to be for zero thickness, although of course this assumption will cause us to miss the correlation energy of each layer in isolation. In solving for the coefficients a_L and a_R we will only need to match discontinuity conditions, which will depend on the per area polarizabilities of the individual layers. Since large d will imply that only small values of k contribute, we can use the small k or large wavelength form for the polarizabilities and for the matching conditions. We assume here that the polarizability of each layer has a principal axis that is perpendicular to the layer. We also mention that for large d, the assumption of planar symmetry on the atomic scale is irrelevant and unnecessary.

The determination of the asymptotic behavior is thus reduced to combining the solution of Laplace's equation in interstitial regions

$$\phi(z) = \begin{cases} a_L e^{kz} & z < 0 \\ a_M e^{kz} + b_M e^{-kz} & 0 < z < d \\ a_R e^{kz} + b_R e^{-kz} & d < z, \end{cases} \tag{7}$$

with the appropriate matching conditions across the layers. These conditions are similar to those derived by Feibelman [19] or in a different form by Langreth [20]. The latter, when applied to the present problem simply give the discontinuity in the parallel and perpendicular components of the electric field $\vec{E}$ as one crosses from one side of the layer to the other. One has

$$\begin{aligned} \mathrm{disc} E_\parallel &= -4\pi \nabla_\parallel p_\perp \\ \mathrm{disc} E_\perp &= -4\pi \nabla \cdot p_\parallel \end{aligned} \tag{8}$$

Here $p_\perp$ and $\vec{p}_\parallel$ are the components of the polarization per unit area of the respective sheet, which are given in linear response by their respective polarizabilities α, which are functions of $\vec{k}$ and ω,

$$\begin{aligned} \vec{p}_\parallel &= \alpha_\parallel \cdot \vec{E}_\parallel \\ \vec{p}_\perp &= \alpha_\perp E_\perp, \end{aligned} \tag{9}$$

where E fields immediately outside the layer are to be used in the above definition; the small discontinuity in these fields would give a higher order correction to the asymptotic form when used in the right sides of Eq. (8). It should be emphasized that quite apart from any intrinsic anisotropies on

[5] But $d \ll ct$ where t characterizes the period for polarization fluctuations; such retardation effects normally occur at distances much greater than those of interest in electronic structure studies

the microscopic level, the polarizabilities in the parallel and perpendicular directions are different [7], because they are defined as the response to the external E-fields, not the local E-fields. The use of (9) in Eq. (8) along with the relation $\vec{E} = -\nabla\phi$ gives matching conditions for ϕ and ϕ', its derivative with respect to z

$$\begin{aligned} \mathrm{disc}\,\phi &= -4\pi\alpha_\perp\,\phi' \\ \mathrm{disc}\,\phi' &= 4\pi\vec{k}\cdot\alpha_\parallel\cdot\vec{k}\,\phi \end{aligned} \tag{10}$$

The use of the conditions (10) in Eq. (7) allows one to obtain

$$\frac{a_R}{a_L} = 1 - (2\pi k)^2 e^{-2kd} P_1(\vec{k},\omega) P_2(\vec{k},\omega), \tag{11}$$

where

$$P_i(\vec{k},\omega) = \left[\alpha^i_\perp(\vec{k},\omega) + \hat{k}\cdot\alpha^i_\parallel(\vec{k},\omega)\cdot\hat{k}\right] \tag{12}$$

where we use the index i ($i = 1, 2$) to label the layer and $\hat{k}$ is a unit vector in the direction of $\vec{k}$. Eq. (11) is now substituted into Eq. (6). At large d, the wave vector k is restricted to small values, and if the polarizabilities of both layers are well behaved for $k = 0$, one may simply expand the logarithm in Eq. (6) to lowest order e^{-2kd} obtaining

$$E_{\mathrm{vdW}} \to -\frac{C_4}{d^4}, \tag{13}$$

where

$$C_4 = \frac{3}{8}\int_0^\infty du\ \langle P_1(iu)P_2(iu)\rangle\,, \tag{14}$$

where the angle brackets indicate an angular average and the absence of $\vec{k}$ in the P's indicates that its magnitude is zero, and only directional information is retained. E_{vdW} is just the part of $E_{\mathrm{c}}^{\mathrm{nl}}$ representing the interaction between the sheets. The angular average takes a trivial form if the components of each of the polarizabilities in the in-plane principal directions are equal:

$$\langle P_1(iu)P_2(iu)\rangle = \langle P_1(iu)\rangle\,\langle P_2(iu)\rangle\,, \tag{15}$$

where

$$\langle P_i(iu)\rangle = \alpha^i_\perp(iu) + \alpha^i_\parallel(iu). \tag{16}$$

This covers the cases of interest here where the layers have square or hexagonal symmetry; other cases can easily be worked out from the more general Eq. (14).

As was pointed out by Barash [21] followed by a number of authors [22], the presence of metallic conductivity in the layers leads to the possibility of a singular response. In terms of our general formulas the effect would be that the α's in Eq. (11) would get large at small ω so that convergence of the u integral in Eq. (6) would require a finite k, and thus precluding the expansion made above, and producing a different form for the asymptote. Typically fractional power laws following off less rapidly than d^{-4} were obtained with various assumptions.[6]

3.3 Method for $E_{\rm c}^{\rm nl}$

An important part of our approximation to $E_{\rm c}^{\rm nl}$ is the use of a simple plasmon pole dielectric function [23]

$$\epsilon_k(z,\omega) = 1 + \frac{\omega_{\rm p}^2(z)}{\omega_q^2 - \omega^2 - \omega_{\rm p}^2} \tag{17}$$

where

$$\omega_q^2 = \omega_{\rm p}^2 + \frac{v_{\rm F}^2 q^2}{3} + \frac{q^4}{4m^2} \tag{18}$$

with $\omega_{\rm p}^2(z) = 4\pi n(z)e^2/m$, and $mv_{\rm F}(z) = (3\pi^2 n(z))^{1/3}$ where $n(z)$ is the local density. The quantity k is the Fourier transform variable in the parallel direction (x-y plane), and ω is the frequency. The quantity q is the three dimensional wave vector

$$q^2 = k^2 + q_\perp^2, \tag{19}$$

which gives $\epsilon_k(z,\omega)$ its k dependence. Since the dependence in the z direction is assumed local via the dependence on $n(z)$, we take the "perpendicular wave vector" $q_\perp$ as a fixed calculated quantity that guarantees the correct polarization properties. In particular, the polarizabilities that determine the longest range part of $E_{\rm c}^{\rm nl}$ are given by

$$4\pi\alpha_\perp(\omega) = \int dz\, \left[1 - \frac{1}{\epsilon(\omega)}\right] \tag{20}$$

and

$$4\pi\alpha_\parallel(\omega) = \int dz\, \left[\epsilon(\omega) - 1\right]. \tag{21}$$

[6] Among the materials discussed in this paper, only graphene exhibits metallic properties in some of its manifestations. Since it is in any case only marginally metallic, one could expect that the resulting singularity is weak, so the effect on $E_{\rm c}^{\rm nl}$ should be weak at the distances involved in our study, even though it would eventually dominate at very large distances.

The difference between the two forms arises, because for constant fields $\vec{E}_\parallel$ is independent of z, while on the other hand it is $\epsilon E_\perp$ that is z independent. These considerations, of course, apply more generally than to the plasmon pole model Eq. (17) and the planar geometry, and similar considerations are important for concentric and parallel interactions between nanotubes [7].

To apply our scheme to a layered system, we start with the density as calculated in GGA, and average it in the directions perpendicular to the layers. This averaged density is used only for the calculation of the vdW part of the density functional, and not for any other component. This is a reasonable approximation for these long range parts that are less critically sensitive to the detailed placement of the atoms. This enables us to calculate the approximate dielectric function $\epsilon(\omega)$ according to Eq. (17). The quantity $q_\perp$ is then chosen so that $\alpha_\perp(0)$ according to Eq. (20) agrees precisely with a full GGA calculation of the perpendicular polarizability of a single layer.[7] The zero frequency polarizability in this perpendicular case is expected to set the scale for the frequency dependent polarizability, so we believe this step is a key to the quality of the results that we obtain.[8] A future improvement might be use TDLDA calculations of the finite frequency polarizability to improve a further parametrized ϵ. Such TDLDA calculations have been successfully used recently [24] to calculate the polarizabilities of several monomers, and to thereby obtain their vdW interactions to second order in the inter-monomer Coulomb interaction.

Once ϵ is determined, then it is a straightforward manner to integrate the Poisson equation

$$\frac{d}{dz}\left(\epsilon\frac{d}{dz}\phi\right) - k^2\epsilon\phi = 0 \tag{22}$$

to obtain ϕ and hence $E_{\rm c}^{\rm nl}$ by use of Eq. (6).[9] For bulk solids, it was found to be sufficient to make two applications of Eq. (22) and Eq. (6) using the bulk density truncated at 30 and 32 layers, respectively. The value of $E_{\rm c}^{\rm nl}$

[7] It should be noted that the value of $q_\perp$ in our approximation for ϵ is strongly dependent on the electron-electron coupling constant λ introduced in Section 3.1. The FPA determination of $q_\perp$, just described, is the place where our functional differs significantly from previous RPA-based treatements of the vdW interactions.

[8] Due to the probability of slow time scales for large distance movements of electrons, the static parallel polarizability $\alpha_\parallel(0)$, in addition to being more difficult to calculate accurately, cannot be trusted to yield the appropriate scale for typical frequencies contributing most to the vdW interaction at typical binding distances.

[9] In order to avoid exponentially increasing or increasing values as the equation is integrated, it is useful to accumulate the ratio of the potential ϕ with and without the sample, each calculated with the same algorithm. The asymptotic value of this ratio then determines a_R/a_L and thereby $E_{\rm c}^{\rm nl}$.

per layer for the infinite solid could then be accurately determined from the energy difference of the respective calculations.

3.4 Seamlessness

The vdW interaction must saturate and make a seamless connection as the interacting fragments are brought together. Such a seamless connection is a challenge for theory, and is typically not obtained in approaches that treat the interacting fragments as distinct entities interacting via interfragment Coulomb interaction. We discuss below how seamlessness occurs in our theory. The quantity E_c^{nl} can be written as

$$E_c^{nl} = -\int_0^\infty \frac{du}{2\pi} \int_0^1 \frac{d\lambda}{\lambda} \mathrm{tr}\left[(1-\tilde{\chi}_\lambda V_\lambda)^{-1}\tilde{\chi}_\lambda V_\lambda - (1-\epsilon_\lambda^{-1})\right]. \qquad (23)$$

where the use of the FPA is implied. The density response $\tilde{\chi}$ is simply expressed in terms of ϵ, since $4\pi e\delta n = \nabla\cdot(\vec{E}-\vec{D}) = \nabla\cdot(1-\epsilon)\cdot\vec{E} = \nabla\cdot(\epsilon-1)\cdot\nabla v/e$. Comparing with the definition $\delta n = \tilde{\chi}v$ gives

$$\tilde{\chi}_\lambda = (4\pi\lambda e^2)^{-1}\nabla\cdot(\epsilon_\lambda - 1)\cdot\nabla. \qquad (24)$$

Eq. (24) tells us how to apply the FPA to the above. One obtains

$$E_c^{nl} = -\int_0^\infty \frac{du}{2\pi}\,\mathrm{tr}\left[\ln(1-V\chi) - \ln\epsilon\right] \qquad (25)$$

where $\lambda = 1$ is implied in all quantities. It is shown in [6] that this can be evaluated exactly by use of Eq. (6). The key point to note is that the expression Eq. (25) vanishes for a uniform system. In this case, ϵ and $\tilde{\chi}$ are diagonal with respect to the spatial indices in the plane wave representation, so that from the above, $\tilde{\chi}V = 1-\epsilon$, causing the two terms in the integral in Eq. (23) to cancel. This is the key to saturation in our theory: as the fragments move closer and closer together, the density becomes more and more uniform, and the influence of this long range correlation piece gets less and less. This guarantees the correct behavior of saturation and a seamless connection. The saturation of E_c^{nl} is illustrated for graphite in Fig. 3. Shown also is the asymptotic curve. Notice that E_c^{nl} gives substantially stronger binding ($\sim 50\%$) in the region of the equilibrium c. This important effect is discussed in detail in Ref. [9] and means that one may not simply add the vdW asymptote to an existing functional, even if it were multiplied by a saturation function.

3.5 Method for E_c^0

The remaining piece of the correlation energy is what is left after comparing Eq. (4) and Eq. (23), that is

$$E_c^0 = -\int_0^\infty \frac{du}{2\pi} \int_0^1 \frac{d\lambda}{\lambda} \left[1 - \epsilon_\lambda^{-1} - \tilde{\chi}_0 V_\lambda\right], \tag{26}$$

an expression that is exact for a uniform system if the exact longitudinal dielectric function is used for ϵ. Of course it would be inappropriate to use the FPA in this term, because we would then, for example, incorrectly weight the second order exchange term which is necessary for getting the high density limit for uniform system correctly [25]. And it would be even more inappropriate to use the simple dielectric approximation of Eq. (17) or to assume locality in the perpendicular direction. Instead we replace ϵ and χ_0 by those appropriate for a uniform system of given density and average the density weighted result over space. In simpler language, we make the local density approximation for this term. We argue that the LDA should be a better approximation for E_c^0 than for the whole of E_c now that E_c^{nl} has been separated off. Furthermore is a consistent approximation in the sense that it becomes exact as the density becomes uniform, since E_c^{nl} vanishes in that limit. Our approximation might be improved once gradient or GGA expansions for E_c^0 alone are derived. But using the full GGA would be double counting, and for the moment we have to be satisfied with Eq. (3).

4 Exchange and correlation together

4.1 Our proposed density functional

Here we summarize the previous two sections on exchange and correlation to give our proposed functional. For exchange, we propose

$$E_x = E_x^{LDA} + E_x^{ZYgrad} \tag{27}$$

where E_x^{ZYgrad} is the piece of the Zhang-Yang [12] revPBE exchange functional depending on the gradient of the local density. Combining this with Eq. (2) for correlation gives

$$E_{xc} = E_{xc}^0 + E_c^{nl} \tag{28}$$

where

$$E_{xc}^0 = E_{xc}^{LDA} + E_x^{ZYgrad} + E_c^0 \tag{29}$$

where we have separated into $E_{\rm xc}^0$ all the terms that are fully local or semilocal. We make a similar separation in our full functional (including the single electron kinetic energy $T_{\rm s}$, Hartree, and potential terms)

$$E^{\rm vdW-DF}[n] = E^0[n] + E_{\rm c}^{\rm nl}[n], \tag{30}$$

and

$$E^0[n] = E^{\rm GGA}[n] - E_{\rm c}^{\rm GGA}[n] + E_{\rm c}^0[n], \tag{31}$$

where we have written the contents of Eq. (29) another way by adding and subtracting again the correlation part of the GGA functional that depends on the local density gradient.

4.2 Practical considerations for the application of the functional

We seek a description of the total energy $E_{\rm tot}$ that simultaneously accounts for both the strong covalent forces within the layers and the weaker dispersion forces between the layers. We thus seek to combine the high degree of traditional DFT in the GGA for regions of denser electron density with an account of the vdW attraction across voids of very low electron densities found in soft and sparse materials. We show here that the formulation of the contribution $E_{\rm c}^{\rm nl}$ provides us with a method to describe the total-energy variation and hence calculate the structure, the binding, and the elastic responses of the layered materials. This applies even though the approximations used in the formal derivation of $E_{\rm c}^{\rm nl}$ (especially the lateral density averaging used in the calculation of this part of the energy) contraindicates a direct application of this term to the in-plane binding.

The argument is based on a simple separation of energy contributions and proceeds as follows. For a simple layered materials characterized by the in-plane lattice constant a and the unit-cell height c we can formally express the cohesive energy

$$\begin{aligned} E_{\rm coh}(c,a) &\equiv E_{\rm tot}(c,a) - E_{\rm tot}(c\to\infty, a\to\infty) && (32)\\ &= [E_{\rm tot}(c\to\infty, a) - E_{\rm tot}(c\to\infty, a\to\infty)] \\ &\quad + [E_{\rm tot}(c,a) - E_{\rm tot}(c\to\infty, a)] && (33) \end{aligned}$$

The crucial part is now to make the following two observations about these terms: (i) The first square-bracketed term in (33) expresses the intra-layer binding and can simply be approximated by calculations of traditional DFT-GGA; (ii) The second term, in contrast, expresses the interlayer binding

and can be calculated by a consistent inclusion of the E_c^{nl} instead of the corresponding GGA term in a new vdW-DF giving $E^{vdW-DF} = E^0 + E_c^{nl}$.

For vdW-DF calculations of the interlayer binding (as shown below for MoS_2 and graphite) we compare the total energy variation at fixed in-plane lattice constant and can then focus on the second term. For calculations of the overall energy variation and for vdW-DF calculations of the materials structure and elastic response we need the full variation based on

$$E_{coh}(c,a) = E_{coh}^{GGA}(c\to\infty,a) + \Delta E^{vdW-DF}(c,a) \quad (34)$$

$$E_{coh}^{GGA}(c\to\infty,a) = E^{GGA}(c\to\infty,a) - E^{GGA}(c\to\infty,a\to\infty) \quad (35)$$

$$\Delta E^{vdW-DF}(c,a) = E^0(c,a) + E_c^{nl}(c,a) - E^0(c\to\infty,a) - E_c^{nl}(c\to\infty,a) \quad (36)$$

Our structure calculations presented in Table 1 are based on cohesive energy calculations using equation (34) with (35) and (36). Using the result $E^{vdW-DF} = E^0 + E_c^{nl}$ directly in equation (32) leads to worse results because our present approximation for E_c^{nl} with planar averaging of the electron density does not treat the intra-plane nonlocal interactions very well. However, the effect is small: for BN this would lead to a 0.6% and 0.1% increase in in-plane lattice constant a and unit-cell height c.

5 Applications to layered materials

We test the new vdW-DF by comparing the layer binding energy (adhesion), the structure, and the elastic response of several layered materials against experiments where available, Table 1. For graphite, hexagonal-BN, and for MoS_2 we calculate the total-energy variation as a function of the in-plane lattice constant a and the height c of the unit cell to determine the structure and elastic properties [26]. All of these materials are sparse and defined by an interlayer separation with voids of very low electron density where the material cohesion is provided by the interlayer vdW-binding. Table 1 shows that the inclusion of the nonlocal correlation effects in the new vdW-DF corrects the shortfalls of traditional DFT and provide us with a predictive materials-physics account without the need to invoke a full time-dependent DFT calculation.

The set of layered materials investigated here represent structurally similar systems with important differences in their electronic nature. Graphite is formed as a staggered A-B alternation of sheets of hexagonal carbon and has a strongly anisotropic dielectric function which even exhibits a semimetallic response for in-plane fields in the low frequencies. Hexagonal boron nitride is isoelectronic with graphite but is a very effective insulator. The

BN structure is similar to that of graphite but replaces every other carbon atom with B (and N) and also has a simpler stacking where the B atoms always sits on top of a N atoms in the adjacent layer. Finally, MoS_2 is similar to hexagonal BN but with the individual sheets formed by alternating Mo atoms and a pair of sulphur atoms slightly displaced above and below the Mo positions. The layered MoS_2 is a semiconductor which, however, supports edge states [27].

The differences in electronic nature implies differences in the detailed variation of the dielectric functions and hence, in principle, differences [22] in the interlayer van der Waals binding. The differences in the electronic structure control the form of the low-frequency response where, for example, the graphite dielectric function acquires a complex component. We argue, however, that since graphite is only semi-metallic, these effects remain insignificant for the strength of the van der Waals binding except at very large layer separations. We thus use a formally identical form of the dielectric function for all three materials (Eq. (17), but with different materials parametrizations). We provide vdW-DF calculations and predictions for total energy, structure, and elastic response that can be used to generally test the robustness of our vdW-DF description.

Our vdW-DF total-energy calculations were performed as follows. We first calculated the electron density self-consistently within the PBE flavor [11][10] of GGA in the plane-wave DFT code DACAPO [29] using Vanderbilt ultrasoft pseudopotentials [30]. The pseudopotentials were chosen to have zero overlap of the pseudization spheres [30] and we chose an accordingly large accuracy.[11] From the PBE electron densities we subsequently extracted a (non-selfconsistent) determination for the revPBE total energy[12] E^{GGA} and the new local correlation energy E_{c}^{0}.

To complete the evaluation of the new nonlocal correlation energy $E_{\mathrm{c}}^{\mathrm{nl}}$ we also calculated in DACAPO the electrostatic response of individual layers and thus fixed the value of the static-polarizability parameter $q_{\perp}$, as discussed in section 3.3 above. Finally, these energy values are inserted in the calculations of $E^{\mathrm{vdW-DF}}$ and E_{coh}, Eqs. (34)–(36).

[10] The choice of GGA-PBE for these underlying calculations is made as PBE pseudopotentials are directly available in the Vanderbilt pseudopotential-generation code [28].

[11] These underlying DFT calculations are based on a 400 eV plane-wave energy cut off, a 12×12 k-point sampling, and a dense FFT grid sampling the electron-density variation.

[12] We tested the accuracy of this (non-selfconsistent) revPBE density and total energy E^{GGA} against self-consistent revPBE calculations in DACAPO using pseudopotentials both in the generic PBE flavor and in the corresponding revPBE flavor (obtained by adapting the Vanderbilt pseudopotential code).

Table 1 summarizes our vdW-DF calculation of the layered-materials structure and elastic coefficients. Using the procedure outlined above we calculated the total cohesive energy for a large number of different values for the in-plane lattice constants a and height c of the unit cell to establish the total-energy variation, for example, as plotted for hexagonal BN in Fig. 1. In our vdW-DF study of (bulk) MoS_2, we treated each MoS_2 as a composite layer and assumed perpendicular Mo-to-S atomic separations fixed at their optimal values ~ 1.58 Å [27]. Following and extending Ref. [26] we finally extracted the optimal structure parameters $(a_{\rm min}, c_{\rm min})$ and both the bulk modulus and elastic coefficient C_{33}.

Figure 4 shows our vdW-DF calculations for the layer binding (total unit-cell adhesion between the layers) for bulk MoS_2. The calculation is performed for optimal in-plane lattice constant, $a = a_{\rm min} = 3.23$ Å, and represents a cut through the data of our wider vdW-DF study of the cohesion, structure and elastic properties of bulk MoS_2 (outlined above). This calculation is thus effectively based on Eq. (36). The figure compares the variation with the unit-cell height c in the contribution from the new semilocal energy E^0 (dot-dashed curve), the new nonlocal-correlation contribution $E_{\rm c}^{\rm nl}$ (dashed curve) as well as the resulting vdW-DF variation (solid curve) $E^0 + E_{\rm c}^{\rm nl}$. The figure shows that the inclusion of the contribution from nonlocal correlation provides the interlayer binding and stabilizes the materials. We note that the optimal value of the unit-cell height $c_{\rm min} = 12.6$ Å calculated within vdW-DF is in fair agreement with experiments (12.24 Å), Table 1. We also note that the individual MoS_2 composite layer has a thickness of 3.16 Å and hence that corresponding MoS_2-to-MoS_2-layer separation $d = c/2 - 3.16$ Å $= 3.14$ Å which is typical of vdW bindings of in sparse, layered materials.

Figure 5 shows our vdW-DF calculation of the total interlayer binding in graphite. Again, the figure represents a cut through our overall vdW-DF investigations of the graphite total-energy variation at the calculated optimal value $a_{\rm min} = 2.47$ Å for the in-plane lattice constant. The predicted optimal height of the unit cell $c_{\rm min} = 7.52$ Å is in fair agreement but slightly larger than the experimental value, Table I. We note that none of the traditional GGA provide a satisfactory account of the graphite binding [8] but that the substitution of the contribution from nonlocal correlation $E_{\rm c}^{\rm nl}$ (dashed curve) remedies this situation. Taken together with the corresponding calculations for hexagonal-BN, Fig. 1, and for MoS_2, Fig. 4, this graphite result provides an illustration that we can consistently extend our materials-physics description by accounting for van der Waals description in sparse materials.

6 Conclusions and outlook

In summary, we have in a series of papers [6, 8, 9] developed a method for extending DFT to the description of vdW interactions which include a seamless saturation as the interacting pieces of matter merge into one. Our applications encompass a set of layered materials. We stress that this set of calculations provides a strong test of the approximation underlying the new vdW-DF. We emphasize that our materials-specific characterizations and predictions for layered materials range from the semi-metal graphite, across the semiconducting MoS_2, and to the insulating hexagonal BN. Such characterizations and predictions can be directly compared and tested in experiment to thus explore possible limitation on the validity of our underlying assumptions.

We believe the outlook for the future of this method is very positive. There are many layered materials where interplanar vdW interactions are important, giving a wealth of potential applications. We have also given several directions by which the accuracy of the functional can probably be improved. There is also the potential of generalizing this functional to other related geometries, with a significantly broader spectrum of applications possible [7]. Also exciting is the fact that the groundwork has been laid [31] for a corresponding general geometry functional, which has now been further developed and tested [15] on rare gas dimers and the benzene dimer. The indications of this preliminary work are that this method offers great promise as well.

Acknowledgments

We thank G. Scoles for providing data relevant to the exchange calculations of Ref. [14] and J. F. Dobson for a discussion of the asymptotic form for metallic layers. Financial support from the Swedish Foundation for Strategic Research via Materials Consortia #9 and ATOMICS and the Swedish Scientific Council are gratefully acknowledged. Work by M.D. and D.C.L. supported in part by NSF Grant DMR 00–93070.

References

[1] B. I. Lundqvist, Y. Andersson, H. Shao, S. Chan, and D. C. Langreth, Int. J. Quantum Chem. **56**, 247 (1995).

[2] Y. Andersson, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **76**, 102 (1996).

[3] J. F. Dobson and B. P. Dinte, Phys. Rev. Lett. **76**, 1780 (1996).

[4] W. Kohn, Y. Meir, and D. E. Makarov, Phys. Rev. Lett. **80**, 4153 (1998).

[5] J. F. Dobson and J. Wang, Phys. Rev. Lett. **82**, 2123 (1999).

[6] H. Rydberg, B. I. Lundqvist, D. C. Langreth, M. Dion, Phys. Rev. B **62**, 6997 (2000).

[7] E. Schröder and P. Hyldgaard, Surf. Sci. **532**, 880 (2003); E. Schröder and P. Hyldgaard, Mater. Sci. Engin. C, in print.

[8] H. Rydberg et al., Surf. Sci. **532**, 606 (2003).

[9] H. Rydberg et al., Phys. Rev. Lett. **91**, 126402 (2003).

[10] J. P. Perdew, et al., Phys. Rev. B **46**, 6671 (1992); **48**, 4978(E) (1993).

[11] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[12] Y. Zhang and W. Yang, Phys. Rev. Lett. **80**, 890 (1998).

[13] B. Hammer, L. B. Hansen, and J. K. Nørskov, Phys. Rev. B **59**, 7413 (1999).

[14] X. Wu et al., J. Chem. Phys. **115**, 8748 (2001).

[15] M. Dion, et al. (to be published).

[16] D. C. Langreth and J. P. Perdew, Solid State Commun. **17**, 1425 (1975); O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976); D. C. Langreth and J. P. Perdew, Phys. Rev. B **15**, 2884 (1977).

[17] C. Mavroyannis and M. J. Stephen, Molecular Physics **5**, 629 (1962).

[18] E. Zaremba and W. Kohn, Phys. Rev. B **13**, 2270 (1976).

[19] P. J. Feibelman, Prog. Surf. Sci. **12**, 287 (1982).

[20] D. C. Langreth, Phys. Rev. B **39**, 10 020 (1989).

[21] Y. U. Barash, Fiz. Tverd. Tela (Leningrad) **30**, 1578 (1988) [Sov. Phys. Solid State **30**, 1580 (1988)].

[22] Y. U. Barash and O. I. Notysh, Zh. Eksp. Teor. Fiz. **98**, 542 (1990) [Sov. Phys. JETP **71**, 301 (1990)]; B. E. Sernelius and P. Björk, Phys. Rev. B **57**, 6592 (1998); M. Boström and B. E. Sernelius, Phys. Rev. B **61**, 2204 (2000); J. F. Dobson (private communication).

[23] B. I. Lundqvist, Phys. Kondens. Matter **9**, 236 (1969).

[24] A. J. Misauitta, B. Jeziorski, and K. Szalewicz Phys. Rev. Lett. **92**, 033201 (2003).

[25] M. Gell-Mann and K. A. Brueckner, Phys. Rev. **106**, 364 (1957).

[26] E. Ziambaras and E. Schröder, Phys. Rev. B **68**, 064112 (2003).

[27] M. V. Bollinger et al., Phys. Rev. Lett. **87**, 196803 (2001); M. V. Bollinger, K. W. Jacobsen, and J. K. Nørskov, Phys. Rev. B **67**, 085410 (2003).

[28] Vanderbilt ultrasoft pseudopotential generation code `http://www.physics.rutgers.edu/~dhv/uspp/` .

[29] DFT code DACAPO `http://www.fysik.dtu.dk/CAMPOS/` .

[30] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990); K. Laasonen, A. Pasquarello, R. Car, C, Lee, and D. Vanderbilt, Phys. Rev. B **47**, 10142 (1993).

[31] H. Rydberg, Ph D. Thesis (2001), Chalmers University of Technology, ISBN 91-7291-068-2.

[32] *Landolt-Börnstein search* (Springer, Berlin, 2003), `http://link.springer.de/` .

[33] Y. Baskin and L. Mayer, Phys. Rev. **100**, 544 (1955).

[34] L. X. Benedict *et al.*, Chem. Phys. Lett. **286**, 490 (1998).

[35] G. Kern, G. Kresse, and J. Hafner, Phys. Rev. B **59**, 8551 (1999).

[36] Th. Böker et al., Phys. Rev. B **64**, 235305 (2001).

TABLES

Table 1: Calculated properties of graphite, BN, and MoS_2, compared to experimental data, when available. The table shows the geometry (a, c), binding energy (E_0) with respect to isolated layers, bulk modulus (B_0), and elastic constant (C_{33}) obtained with our revised DFT, vdW-DF. The B_0 is calculated using the method of Ref. [26].

Substance	Method	a [Å]	c [Å]	E_0 [meV/atom]	B_0 [GPa]	C_{33} [GPa]
Graphite	vdW-DF	2.47	7.52	24	12	13
Graphite	Exp.	2.46[a]	6.70[b]	35 ± 10[c]	33[a]	$37 - 41$[a]
BN	vdW-DF	2.51	7.26	26	11	11
BN	Exp.	2.50[d]	6.66[d]			
MoS_2	vdW-DF	3.23	12.6	60	39	49
MoS_2	Exp.	3.16[e]	12.29[e]			

[a]Ref. [32], [b]Ref. [33], [c]Ref. [34], [d]Ref. [35], [e]Ref. [36].

FIGURES

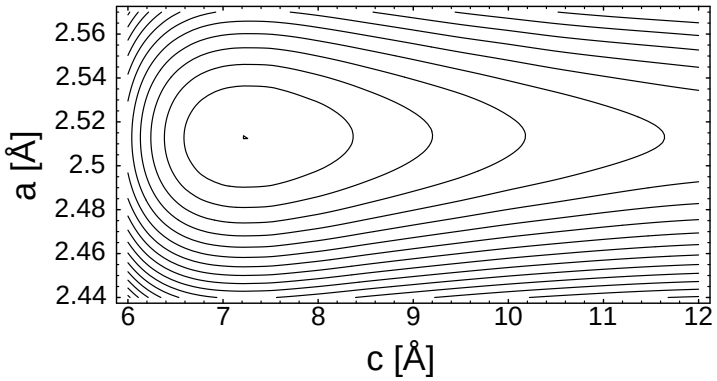

Figure 1: vdW-DF theory calculation of the cohesive-energy variation in hexagonal BN as a function of the in-plane lattice constant a and the unit-cell height c. Contour line intervals indicate a 20 meV/unitcell change in energy. The results are based on a large number of traditional DFT calculations for the variation of the electron density used in the new vdW-DF for layered materials, Eq. (30). The figure illustrates that the new vdW-DF can provide an accurate materials-physics description both of the intra-plane covalent binding and of the dispersion forces that bind the layers across the voids of ultra-low electron density. Using the shown variation in the vdW-DF binding energy we predict a structure $(a_{\rm min}, c_{\rm min})$ in fair agreement with the experimental values ($a_0 = 2.5$ Å, $c_0 = 6.7$ Å).

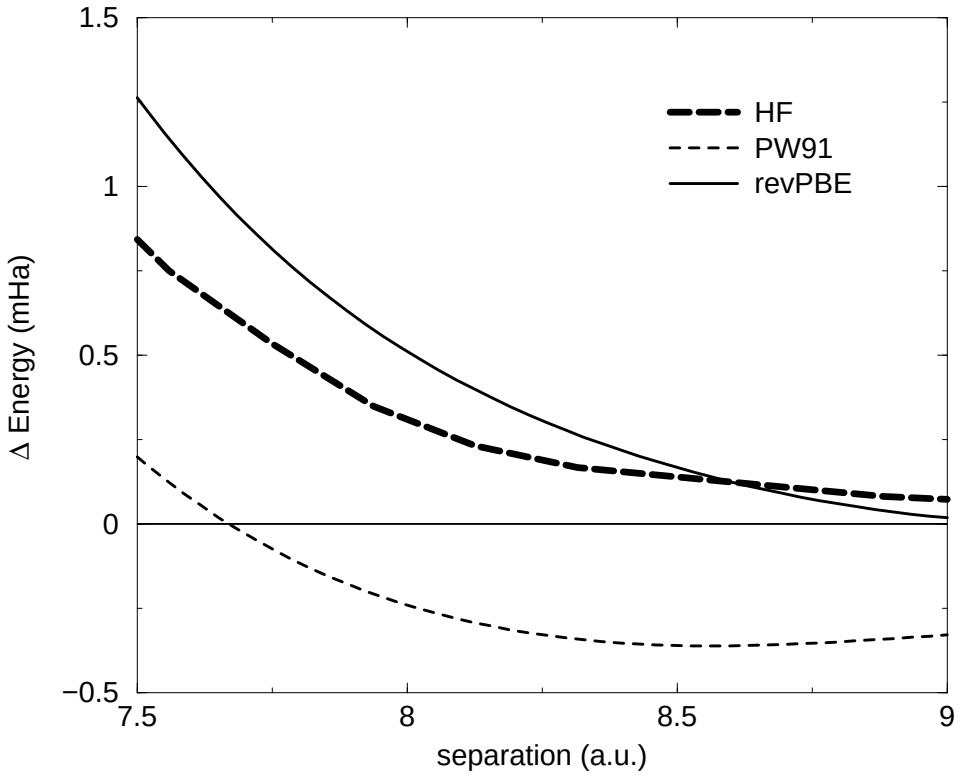

Figure 2: Exchange functionals vs. exact Hartree Fock exchange for the interaction energy for the Kr dimer (HF data from Ref. [14]).

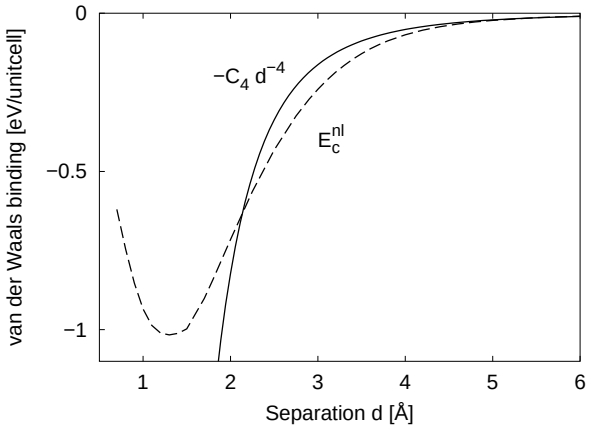

Figure 3: Enhancement and eventual saturation of the nonlocal-correlation energy contribution $E_{\rm c}^{\rm nl}$ evaluated for a pair of graphite sheets as a function of the interlayer separation d.

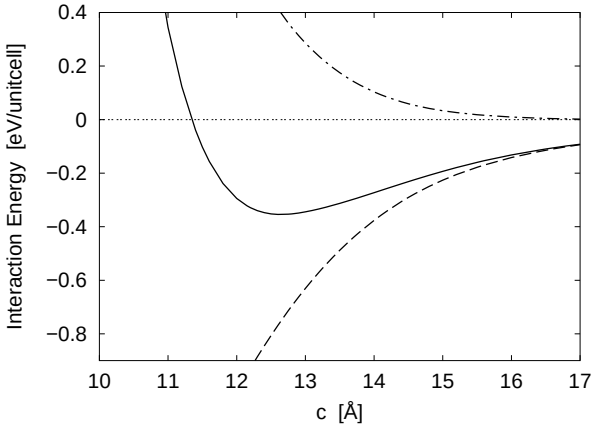

Figure 4: New vdW-DF theory description of the interlayer binding (solid curve) in MoS_2 as a function of the height c of the MoS_2 unit cell. The figure shows that this vdW-DF variation arises as a competition between the repulsion in the semi-local energy contribution E^0 (dash-dotted curve) and the vdW attraction described in the new nonlocal correlation expression E_c^{nl} (dashed curve). Our description is based on a large number of underlying calculations of the electron density in the revPBE flavor of GGA at general values of c and of the in-plane lattice constant a from which we establish the structure. The figure shows results evaluated at the optimal value $a_{min} =$ 3.23 Å that we thus calculate in vdW-DF theory. Note that the individual MoS_2 compound-layer has a thickness of 3.16 Å and the optimal value of c thus corresponds to a interlayer separation $d = c/2 - 3.16$ Å= 3.14 Å very close to those that characterizing the corresponding vdW interlayer binding in both hexagonal BN and in graphite.

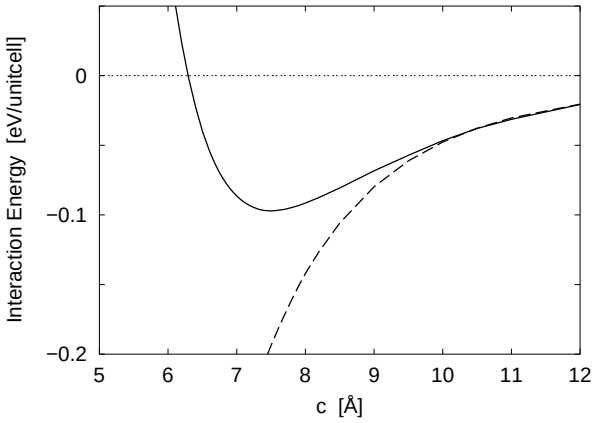

Figure 5: vdW-DF calculations of the total interlayer binding in graphite. The figure compares the results of the vdW-DF calculation (solid curve) and the contribution E_c^{nl} (dashed curve) as functions of the unit-cell height c. The curve shows a cut through the overall total-energy variation at the optimal vdW-DF in-plane lattice constant $a_{min} = 2.47$ Å.