A Harris-type van der Waals density functional scheme

Kristian Berland, Elisa Londero, Elsebeth Schröder and Per Hyldgaard
Microtechnology and Nanoscience, MC2, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

(Dated: March 14, 2013)

Large biomolecular systems, whose function may involve thousands of atoms, cannot easily be addressed with parameter-free density functional theory (DFT) calculations. Until recently a central problem was that such systems possess an inherent sparseness, that is, they are formed from components that are mutually separated by low-electron-density regions where dispersive forces contribute significantly to the cohesion and behavior. The introduction of, for example, the van der Waals density functional (vdW-DF) method [PRL 92, 246401 (2004)] has addressed part of this sparse-matter system challenge. However, while a vdW-DF study is often as computationally efficient as a study performed in the generalized gradient approximation, the scope of large-sparse-matter DFT is still limited by computer time and memory. It is costly to self-consistently determine the electron wavefunctions and hence the kinetic-energy repulsion. In this paper we propose and evaluate an adaption of the Harris scheme [PRB 31, 1770 (1985)]. This is done to speed up non-selfconsistent vdW-DF studies of molecular-system interaction energies. Also, the Harris-type analysis establishes a formal link between dispersion-interaction effects on the effective potential for electron dynamics and the impact of including selfconsistency in vdW-DF calculations [PRB 76, 125112 (2007)].

PACS numbers: 31.15.E-,31.15.ae, 71.15.Nc

I. INTRODUCTION

Density Functional Theory (DFT) is considered one of the best and most reliable condensed-matter tools for non-empirical studies of molecular, surface, bulk and compound properties. Standard implementations, using the Local Density Approximation (LDA) or the Generalized Gradient Approximation (GGA) for the exchange-correlation energy, provide an accurate description of the binding in regions characterized by high electron density. An even more widespread DFT usage will follow from an ability to address soft- and sparse-matter systems, structures that have internal voids or low-electron-density regions dominated by the van der Waals (vdW) forces, also called the London dispersion forces. While neither LDA nor GGA capture the truly nonlocal correlation effects that underpin those forces, the last decade has seen development of both vdW-extended DF and of regular nonlocal exchange-correlation functionals. The first class of methods are often atom centered and require use of a damping function or equivalent, while the second class of methods fits inside the regular DFT framework. Both types of sparse-matter DFT can describe, for example, vdW forces between molecules.

The van der Waals density functional (vdW-DF) method is a framework for approximating the exchange-correlation energy $E_{xc}[n]$. The method, summarized below, yields efficient general-purpose sparse-matter functionals that are non-empirical. The method employs the Coulomb gauge (with Green function $G = (r - r')^{-1}$) and a scalar dielectric function $\varepsilon$. In its most general form, the vdW-DF method is a reformulation of the adiabatic connection formula (ACF) and assumes that a plasmon-pole approximation for $\varepsilon$ can satisfy

$$
\int \frac{du}{2\pi} \text{Tr} \{\ln(\nabla G)\} \equiv E_{xc}[n] + E_{\text{self}},
$$

where $u$ denotes the complex frequency and where $E_{\text{self}}$ is the internal Coulomb self-energy of each electron. This equation summarizes an, in principle, exact description of the (longitudinal) electrodynamics in the inhomogeneous electron gas and reflects the use of a Dyson equation for handling screening. Like the GGAs, the vdW-DF method further uses physics-based constraints to approximate the plasmon-pole response and thus defines the functional form of a nonlocal correlation term, $E_{\text{nl}}$. In the recent explicit functional versions, called vdW-DF and vdW-DF, the nonlocal energy $E_{\text{nl}}[n]$ is expressed as a double integral over the density, weighted by a kernel. However, the plasmon basis still allows it to capture a collectivity that reflects the broader density variation. The vdW-DF method also involves picking a gradient-corrected exchange that reflects prespecified criteria, e.g., good all-round molecular-system performance or improved bulk-system properties. The vdW-DF shares the plasmon-pole emphasis with its LDA and GGA relatives, and the functionals “vdW-DF#” have both seamless integration in the homogeneous limit and a build-in conservation of the exchange-correlation hole. The non-empirical design suggests that the vdW-DFs can achieve a good transferability across systems, length scales, charging states, and binding morphologies.

The vdW-DF method has been and is being tested for many systems. It delivers a parameter-free atomic-scale characterization of the binding in complex sparse-matter systems. Selfconsistent (sc) vdW-DF calculations can be used to calculate stress within periodic unit cells and guide atomic optimization (relaxation). There
are performance tests for bulk, layered, absorption, molecule and atom adsorption, surface self-assembly, and molecular systems. The development of effective algorithms has allowed "sc vdW-DF calculations that have today little or no additional costs over GGA calculations. For very large systems, typical of biomolecular-interaction problems, the set of real-space vdW-DF calculation schemes effectively permits an order-\(N\) scaling in the evaluation of \(E^\text{nl}_c[n]\), as discussed for example in Ref. [31].

Nevertheless, the feasibility of sparse-matter DFT calculations for large molecular matter is still limited by the huge number of atoms that are usually involved in interacting molecular systems. The problem is confounded by the fact that (especially in biochemically relevant problems) we are often forced to address nonperiodic systems where size-convergence of the model unit-cell size becomes important. Molecular transport studies, and many molecular-crystal problems further exemplify the general need to address vdW binding in big systems. The challenge for typical sparse-matter DFT is increasingly becoming that of computing the steric hindrance effects that are described by the DFT kinetic-energy repulsion. When we reach thousands of atoms, all first-principle and vdW-extended DFT calculations are simply limited by memory requirements and computational costs of the wavefunction evaluation.

In this paper we propose and test an adaption of the Harris scheme to perform non-selfconsistent (nsc) vdW-DF calculations for sparse matter. The aim is to explore possibilities for reducing the computational cost of the wavefunction-evaluation bottleneck that could be impeding an even broader vdW-DF usage in, for example, biochemistry. The approach consists of using a superposition of the fragment (electron) densities, and is here termed sfd-vdW-DF. It can be seen as an alternative to the more costly nsc-vdW-DF or even sc-vdW-DF evaluations. It is just the regular vdW-DF Harris scheme if the fragment densities are based on vdW-DF calculations. However, the sfd framework also works without a vdW-DF implementation of the underlying DFT code, and we reserve the term sfd-vdW-DF to cases where the fragment densities are obtained from GGA calculations. The method uses our code for real-space vdW-DF evaluation based on the charge densities from the underlying DFT code.

We present formal analysis and a testing of the sfd-vdW-DF computational scheme for molecular systems. As an interesting aside, the analysis also identifies conditions for expecting significant vdW-induced changes in the bandstructure or electron dynamics for a given binding morphology. We separately test how well the sfd framework faithfully reproduces interaction effects that arise from the semi-local part of the vdW-DF functional. We test the performance of the sfd-vdW-DF scheme across the S22 benchmark set and for systems with a varying degree of static polarizations. We observe that while a vdW-DF Harris scheme (evaluated with vdW-DF fragment densities) is correct to second order in binding-induced density changes, \(\delta n = n_\text{sc} - n_\text{sfd}\), the proposed scheme is formally only correct to linear order in these density changes. However, we also find that the linear term \(\delta n\) is weighted by the changes in the effective Kohn-Sham (KS) potential that result from the shift from the GGA to the vdW-DF functional. The sfd-vdW-DF scheme may thus be broadly applicable in the absence of large static dipoles.

The Paper is organized as follows. In Section II we summarize the case for developing an accelerated (but approximate) vdW-DF description of biomolecular interaction. In Section III we present some details of the LDA, GGA, and vdW-DF family of constraint-based density functionals to facilitate a formal analysis and our proposal for the sfd-vdW-DF scheme in Section IV. Section V provides a brief summary of computational (and implementation) details. Section VI presents the results of the performance testing that we have carried out for the proposed sfd-vdW-DF scheme, while Sec. VII contains a discussion. Finally, Section VIII contains summary and outlook.

II. A CASE FOR FAST VDW-DF STUDIES OF BIOMOLECULAR INTERACTIONS

In a related work we report a vdW-DF mapping of the vdW attraction in a DNA dimer (two periodic double-helix strands). That pilot study illustrates that general conclusions can be made for the important biomolecular-interaction problem from computing merely how \(E^\text{nl}_c\) depends on the interaction geometry.

The nonempirical nature and the regular-density-functional basis (no external parameters) makes vdW-DF well suited to pursue investigations of biomolecular interactions. The full DNA interaction problem touches on two general challenges for refining a computational description of life processes. First, it illustrates the workings of molecular recognition (the matching by weak forces of the genes or just the packing of our genome in its environment). Second, it reminds us of the challenge of characterizing these effects in a solution that contains counter ions. Since the counter ions, per se, can be expected to play a smaller role in the overall interfragment vdW attraction, it is natural to focus a sparse-matter DFT study on the charged biomolecular structures themselves. However, for this approach to become meaningful, we must be able to also characterize the vdW bonding at various charging states. Being a regular (parameter free) nonlocal density functional, vdW-DF has an inherent advantage for computational studies of biomolecular interactions.

The prospect of vdW-DF studies of biomolecular systems is as promising in terms of computational cost (and, currently, as restrained) as it is for other sparse-matter DFT approaches. The evaluation of the nonlocal correlation, \(E^\text{nl}_c\), causes no relevant slow down or bottlenecks.
thanks to the order-$N$ scaling in the real-space evaluation approach for large systems. In fact, our mapping of DNA attraction demonstrates that vdW-DF has, in practice, an efficiency similar to that of DFT-D for large biomolecular problems. There are for the DNA-dimer attraction problem no memory bottlenecks and essentially a linear scaling at least up to 1000+ cores for the $E_n^\text{nl}$ evaluation. It is already today possible to bring that functional evaluation to about ten minutes wall time. While this evaluation of $E_n^\text{nl}$ does cost more than adding the dispersion term in DFT-D, neither of these descriptions of the vdW attraction cause limitations: The computing expense is irrelevant when compared to the cost of converging the density in even one DNA (one periodically repeated coil of a DNA double helix) in DFT.

With such a promise, it is frustrating that a full DFT study of the DNA dimer is today impossible without an allocation at a petaflop facility. There is a need to accelerate the DFT determination of large-system wavefunctions and for sparse matter investigations in general. Here we explore an approach that focuses on accelerating the evaluation of the kinetic-energy repulsion, and which can, in principle, reduce the wavefunction-solution stage to just a single electronic iteration, as is possible with the regular Harris scheme.

Figure 1 illustrates the feasibility of using the sfd-vdW-DF for accelerated vdW-DF studies of biomolecular interactions. It reports a comparison of binding in an ammonia dimer and shows that the sfd-vdW-DF2 description is in excellent agreement with the binding predicted by sc-vdW-DF2 (indicated by a cross).

Figure 2 summarizes our overall assessment, further detailed in Sec. VI. It testifies to a high degree of robustness across the S22 benchmark set and for other molecular-interaction problems. The important observation for developing a biomolecular computational strategy in DFT is that the sfd-vdW-DF scheme can be as reliable as the often-used nsc-vdW-DF evaluation even if it bypasses the need for a sc determination of the wavefunctions. The sfd framework could thus be one ap-
approach to speed up vdW-DF studies of large biomolecular interaction problems at a limited cost in accuracy.

The (regular) Harris scheme works within a fixed functional choice and takes the following steps: First, self-consistent calculations of the densities \( n_{i=1,2,...} \) for each individual building block ‘\( i \)’; second, construction of a density \( n_{\text{sd}} = \sum_i n_i \) as a superposition of the building blocks, and of the effective single-particle potential \( V_{\text{eff,sd}}(r) = V_{\text{eff}}[n_{\text{sd}}](r) \) that corresponds to \( n_{\text{sd}} \); and third a nsc, that is, no-density update, calculation of the eigenvalues corresponding to the potential \( V_{\text{eff,sd}}(r) \). These eigenvalues help provide an estimate of the kinetic energy term in the Harris scheme.\(^{55}\) The Harris scheme is traditionally pursued in a LDA or GGA framework starting, e.g., from sc GGA input densities. It is today often used when pursuing bandstructure calculations in DFT (and therefore provides very accurate wavefunctions for a given \( V_{\text{eff,sd}}(r) \)). However, it is also possible to use it for its original purpose, namely for providing efficient but approximate DFT studies of interactions.

Pursuing a regular vdW-DF Harris scheme represents one alternative for accelerating vdW-DF studies. We point out that the original study by Harris\(^{55}\) (working with LDA) shows that the scheme works reasonably well also for describing the formation of covalent bonds between atoms in some molecules. It should be even better suited to describe the weaker vdW bonding. The overall criteria for the applicability of a Harris-type scheme is that the fully sc density solution \( n_{\text{sc}} \) should not differ significantly from the input superposition density \( n_{\text{sd}} \). It can work also when we consider systems, like biomolecules in solution, where the charging from the surrounding counter ions must be considered when providing the input densities (and hence \( n_{\text{sd}} \) and \( V_{\text{eff,sd}}(r) \)). A vdW-DF Harris scheme can be expected to work well for a study of supramolecular systems as long as the binding-induced charge relocations remain small.

A regular Harris vdW-DF scheme is, however, not the focus here. This is because we do not yet have the ability to both perform vdW-DF calculations and allow an externally-defined superposition of input density in the same code. We have chosen to build on the DACAP\(^{22}\) code (for input densities and for the nsc kinetic-energy evaluation\(^{17}\)) and on the vdW-DF postprocessing\(^{21,25,26}\) that we have previously used extensively for vdW-DF studies. A benefit of introducing sfd-vdW-DF as a supplement to a regular vdW-DF Harris scheme is that we thus provide a computational framework that can take input densities from an arbitrary code (through an adaptation of any of the real-space strategies for evaluation vdW-DFT\(^{10,13,17,27}\)).

### III. DENSITY FUNCTIONAL DESCRIPTIONS OF DENSE AND SPARSE MATERIALS

We begin with a description of differences between the regular form of the nonlocal functional vdW-DF and the traditional local (LDA) and semilocal (GGA) DFT descriptions. This facilitates our subsequent discussion of the approximation that can allow an accelerated evaluation in (nsc) vdW-DF studies.

#### A. Nonempirical approximations to the universal energy density functional, LDA, GGA, vdW-DF

Both GGA and vdW-DF are refinements of the earlier approach LDA that describes the universal exchange correlation functional

\[
E_{\text{x,c}}^{\text{LDA}}[n] = E_x^{\text{LDA}}[n] + E_c^{\text{LDA}}[n] \tag{2}
\]

in terms of exchange and correlation energy densities \( \epsilon_x^{\text{LDA}}(n(r)) \) that are just functions of the local density \( n(r) \). The behavior of the LDA was initially established by considering the self-energy shifts that result with a single-particle coupling to the collective plasmon excitations.\(^{22,23}\) The LDA description was later refined by considering Quantum Monte Carlo studies of the homogeneous electron gas and its response.\(^{24,25}\)

The GGA adds a functional dependence on the local gradient through dimensionless measures of the gradient, \( s(r) = |\nabla n|/(2k_F n) \) and \( t(r) \propto |n|/(nk_s) \), where \( k_F = (3\pi^2 n)^{1/3} \) is the Fermi wave number, \( k_s = (4k_F/\pi a_0)^{1/2} \) denotes the Thomas-Fermi screening wave number, and \( a_0 = \hbar^2/mc^2 \). The set of GGAs expresses the exchange-correlation energy

\[
E_{\text{x,c}}^{\text{GGA}}[n] = \int d^3 r n(r) [\epsilon_x^G(n(r),s(r)) + \epsilon_c^G(n(r),t(r))]; \tag{4}
\]

we keep a subscript ‘\( g \)’ on the GGA energy densities to stress that one must pick a particular design choice, although history has pulled towards a few major choices.\(^{11}\)

The inclusion of the dimensionless gradient permits a richer variation of functional forms. The set of constraint-based GGAs is among the most successful\(^{26,28}\) and these extend the plasmon picture of the LDA via a wavevector analysis,\(^{23,28,53}\) while also emphasizing conservation of the exchange-correlation hole. The development led to robust and very versatile GGA forms like the PBE.\(^{53}\) In the constraint-based GGAs, the form of the exchange energy density

\[
\epsilon_x^G(n(r),s(r)) = \epsilon_x^{\text{LDA}}(n(r)) F_x^G(s(r)), \tag{5}
\]

is given by an enhancement factor \( F_x^G(s(r)) \) that has been chosen to satisfy a number of scaling laws, formal constraints and guidelines.\(^{24,52,55}\) Formal analysis also guides the choice of the gradient-corrected correlation energy density \( \epsilon_c^G(n(r),t(r)) \). It is important to note that the richer variation that \( \epsilon_x^G \) supports is thus tempered by adherence to fundamental physics criteria and that the identification and use of these criteria has helped propel the GGA (and DFT) to a tremendous success.\(^{10}\)
The broad class of materials and systems that are characterized by sparseness does, however, require further refinements beyond GGA. For example, organics, biomater- 
and supramolecular systems are sparse in the sense that they have internal electron voids or other regions with a low electron distribution. Here the binding and function are dominated by the vdW forces that reflect an electrodynamical coupling and act across internal voids.

The vdW-DF method goes beyond LDA and GGA by introducing a truly nonlocal correlation contribution $E^{\text{nl}}[n]$ that makes the electrodynamical coupling explicit \cite{1,4,12,13} via Eq. (1). The vdW-DF exchange-correlation energy is thus written

$$E_{xc}^{\text{vdW-DF}}[n] = E_{xc}^{\text{c,0}}[n] + E_{xc}^{\text{nl}}[n]$$

(6)

where

$$E_{xc}^{\text{c,0}}[n] = E_{xc}^{\text{LDA}}[n] + E_{xc}^{\text{c}}[n]$$

(7)

denotes the semilocal part of the functional. We use a superscript ‘$c$’ to identify the vdW-DF version \cite{12,13} and stress that these in general have different exchange components $E_{xc}^{\text{c}}[n]$. We note that the different vdW-DF versions also have different forms of the nonlocal correlation term but we have chosen not to make that explicit in our discussion.

In the recent vdW-DF versions the nonlocal correlation term is represented by a second-order expansion in the plasmon-pole response \cite{13,24}

$$E_{xc}^{\text{nl}}[n] = \frac{1}{2} \int d^3r \int d^3r' n(r)\phi[n](r,r')n(r').$$

(8)

This nonlocal correlation term still captures the broader density variation through a collectivity that the plasmon reflects and \cite{5} is designed so that the vdW-DF method avoids double counting with the terms captured in the local correlation. The vdW-DF# have a seamless integration

$$E_{xc}^{\text{c}}[n] = E_{xc}^{\text{LDA}}[n] + E_{xc}^{\text{c}}[n],$$

(9)

and thus bypass the need for using a damping function. As stressed in the introduction, the vdW-DF# are derived as an approximation to the ACF. The recent more explicit functionals also use the ACF to link the plasmon pole to an inner functional \cite{11,12,13,24} that is also of the form (7). All functionals in the vdW-DF method build the nonlocal functional from the inside out, i.e., describe the electrodynamics of the dispersion forces by linking to response of the time-tested LDA/GGA plasmon description.

In this paper we use the vdW-DF method and work with both the vdW-DF1 version (in which the revPBE \cite{55} GGA is used for $E_{xc}^{\text{vdW-DF}}[n]$) and the vdW-DF2 version (in which the re-fitted PW91 \cite{55} GGA is used for $E_{xc}^{\text{vdW-DF}}[n]$). We note that in addition to the canonical Rutgers-Chalmers vdW-DF versions \cite{12,13,24}, there are also variants \cite{55,58,59}. These variants fit the outer exchange functional $E_{xc}^{\text{vdW-DF}}[n]$ to a form that is fitted to specific data sets, for example the S22.

**B. The exchange-correlation potentials**

For a discussion of the nature of the Harris and nsc-vdW-DF schemes (below) it is important to also characterize differences in the corresponding exchange-correlation potentials

$$\mu_{xc}^g[n](r) = \frac{\delta E_{xc}^g[n]}{\delta n(r)},$$

(10)

$$\mu_{xc}^v[n](r) = \frac{\delta E_{xc}^{\text{vdW-DF}}[n]}{\delta n(r)}.$$ 

(11)

We also use these potentials in a discussion of the error in the sfd-vdW-DF. Again we have used superscripts $g$ and $v$ to stress that for calculations we must pick specific versions of the GGA or of the vdW-DF.

Ref. \cite{1} provides a derivation of the effective-potential contribution $\Delta \mu_c^{v}[n](r)$ that arises from taking functional derivatives of the nonlocal correlation term $E_{xc}^{\text{nl}}$. From \cite{5} it follows that the vdW-DF exchange-correlation potential will also differ from a GGA exchange-correlation potential

$$\mu_{xc}^v[n](r) - \mu_{xc}^g[n](r) =$$

$$\Delta \mu_c^{v,0}[n](r) + \Delta \mu_c^{v}[n](r)$$

(12)

by a semilocal potential term $\Delta \mu_c^{v,0}[n]$. This semilocal potential change arises in part because vdW-DF subtracts off the gradient corrections to correlation. Also when discussing the difference from a given GGA ‘$g’ exchange-energy form $E_{xc}^{g}[n]$, the semilocal potential change must reflect the differences $E_{xc}^v[n] - E_{xc}^g[n]$.

**C. Selfconsistent DFT calculations of the total energy**

The KS energy can be written \cite{55}

$$E_{KS}[n] = T_0 + \int d^3r n(r) \left[ \frac{1}{2} \phi_n(r) + V_{ext}(r) \right] + E_{xc}[n] + E_N.$$ 

(13)

Here $V_{ext}(r)$ is the external potential and $\phi_n(r)$ the electrostatic potential at the given density $n(r)$,

$$\phi_n(r) = \int d^3r' n(r') \frac{n(r')}{|r-r'|}.$$ 

(14)

The first term $T_0$ and the last term $E_N$ of (13) express the kinetic energy of an effective single-particle wavefunction problem and the internuclear repulsion term, respectively.

Fully selfconsistent DFT calculations in the KS scheme proceed by finding the single-particle wavefunctional solution of an effective eigenvalue problem

$$\left\{ -\frac{1}{2} \nabla^2 + V_{eff}[n](r) - \epsilon \right\} \psi_\lambda(r) = 0$$ 

(15)
defined by the density-dependent (and density-functional specific) effective potential

\[ V_{\text{eff}}[n](\mathbf{r}) \equiv V_{\text{ext}}(\mathbf{r}) + \phi_n(\mathbf{r}) + \mu_{xc}[n](\mathbf{r}). \tag{16} \]

We use atomic units in all formal discussions of DFT calculations and the set of approximations. Selfconsistency is enforced by requiring that the resulting single-particle description of the electron density

\[ \tilde{n}(\mathbf{r}) = \sum_{\lambda} |\psi_{\lambda}(\mathbf{r})|^2 \tag{17} \]

actually coincides with the density that specified the effective single-particle potential \[ 15 \].

As an example of the total-internal energy DFT calculation that is thus made possible, we consider a two-fragment system with components separated by a distance \( d \). We here follow the presentation in Ref. 68 so as to simplify our subsequent discussion (Section IV). Summing up the set of occupied, single-particle energies \( \epsilon_{\lambda} \), leads to an incorrect counting of the total electron-electron interaction energy. However, this complication is easily adjusted for, giving \[ 19 \]

\[ E_{\text{KS}}(d) = \sum_{\lambda} \epsilon_{\lambda} - \int d^3r \ n(\mathbf{r}) \left\{ \frac{1}{2}\delta_{\text{nuc}}(\mathbf{r}) + \mu_{xc}[n](\mathbf{r}) \right\} + E_{xc}[n] + E_N(d). \tag{18} \]

All terms depend on the mutual fragment separation \( d \) (although we do not make that an explicit statement for all terms). We use \( E_{\text{vdW-DF}}^{\text{sc}}(d) \) or \( E_{\text{KS}}^{\text{vdW-DF}}(d) \) to specify whether the sc DFT total energy result was pursued in a vdW-DF or a GGA choice, respectively. Below we focus on the discussion of such fragment problems with mutual separation \( d \).

### D. Non-selfconsistent vdW-DF calculations

We present an overview of the standard nsc-vdW-DF evaluation \[ 12-14 \] which has a total-energy variation \( E_{\text{vdW-DF}}^{\text{nsc}}(d) \). This energy variation is often \[ 12-14 \] found to be a close approximation to the energy variation \( E_{\text{vdW-DF}}^{\text{nsc}}(d) \equiv E_{\text{vdW-DF}}^{\text{sc}}(d) \) found by fully sc vdW-DF calculations, Sec. III.C.

The nsc-vdW-DF evaluations proceed for a given GGA ‘\( g \)’ by first completing self-consistent GGA calculations of both the electron density variation \( n_{\text{sc}}^g(d) \) and total GGA internal energy \( E_g(d) \). The GGA choice \( g \) is in practice often PBE and perhaps more seldom the revPBE version (that vdW-DF1 uses for its exchange component but that is of no concern in this formal discussion). We denote by \( E_{\text{vdW-DF}}^{\text{sc}}(d) \) the corresponding exchange and correlation energy that are evaluated for \( n_{\text{sc}}^g(d) \). The nsc-vdW-DF calculations proceed by simply adjusting for the nonlocal correlation and for the differences in semi-local correlation terms

\[ E_{\text{vdW-DF}}^{\text{nsc}}(d) = E_{\text{g}}(d) + E_{\text{c}}^{\text{nl}}[n_{\text{sc}}^g(d)] + \Delta E_{xc}^{\text{g}}[n_{\text{sc}}^g(d)]. \tag{19} \]

The semilocal functional component

\[ \Delta E_{xc}^{\text{g}}[n_{\text{sc}}^g] = E_{c}^{\text{LDA}}[n_{\text{sc}}^g] + E_{c}^{\text{g}}[n_{\text{sc}}^g] - E_{xc}^{\text{g}}(d) \tag{20} \]

not only extracts the gradient corrections to correlations but also implements a possible adjustment in the gradient-corrected exchange description (as necessary).

The nsc-vdW-DF calculations were for some time the only manner for completing a vdW-DF study: It permitted us to include van der Waals interactions in a computational efficient parameter-free single-density functional DFT. \[ 12-14 \] The approach can be motivated, in part by a Harris-type description (as substantiated further below) but the quantitative extent of the approximation could only be tested when efficient implementations of the sc-vdW-DF method were made available. \[ 11,21 \] The subsequent testing showed that nsc-vdW-DF often captures most of the binding of sc-vdW-DF. \[ 13 \]

### IV. HARRIS-TYPE EVALUATION SCHEMES

This paper formally proposes a computational strategy that combines nsc-vdW-DF calculations (above) with a further approximation inspired by the Harris scheme \[ 68,70 \] and other earlier suggestions of using frozen fragment densities.\[ 71,72 \] The approximation can limit the computational costs for molecular systems because it reduces the quality required for the input density in the nsc-vdW-DF evaluation. It comes with an accuracy cost, which as expected is found largest for systems with a static polarization, but it can provide a speed up.

#### A. Variational nature of the Harris scheme

The regular Harris scheme rests ultimately on the variational character of the KS formulation of the total energy for fully selfconsistent DFT evaluations. The KS energy form acquires a minimum at the correct ground-state density \( n_{\text{sc}} \).

\[ E_{\text{KS}}[n_{\text{sc}} + \delta n] = E_{\text{KS}}[n_{\text{sc}}] + C(\delta n)^2; \quad C > 0. \tag{21} \]

The Harris scheme rewrites the KS formulation of the total energy so as to avoid the need for updating the electron density in an estimate of the interacting energy. As mentioned in the introduction, the Harris scheme does not provide nor does it work with the sc density \( n_{\text{sc}} \), but rather utilizes a superposition density

\[ n_{\text{sfd}} = \sum_{i} n_{\text{sc},i} \tag{22} \]

defined from sc determinations of the electron densities \( n_{\text{sc},i} \) for each of the fragments of the weakly interacting system.
In a GGA study, for example, we can formally express the Harris interaction estimate\(^{63}\)

\[
E_{\text{Harris}}^g(d) \equiv \sum_{\lambda} \epsilon_{\lambda}^g - \int d^3 r \, n_{\text{eff}}^g(r) \left\{ \frac{1}{2} \phi_{n_{\text{std}}}^g(r) + n_{xc}^g [n_{\text{std}}^g](r) \right\} + E_{\text{xc}}^g [n_{\text{std}}^g] + E_N(d). \tag{23}
\]

Again, \(d\) is the distance between the fragments and \(\phi_{n_{\text{std}}}^g\) is the electrostatic potential defined by \(n_{\text{std}}^g\). In the Harris estimate \(^{23}\), the eigenvalues \(\epsilon_{\lambda}^g\) are the single-particle energy sum calculated within the Harris “one-shot” (no density update) GGA calculation for the frozen input superposition density \(n_{\text{std}}^g\) using the effective (Harris-GGA) potential

\[
V_{\text{eff, std}}^g(r) = V_{\text{eff}}^g[n_{\text{std}}^g(r)] = V_{\text{ext}}(r) + \phi_{n_{\text{std}}}^g(r) + n_{xc}^g [n_{\text{std}}^g](r). \tag{24}
\]

There does, of course, exist a corresponding expression for a Harris approximation to vdW-DF calculations, in which case the input density would be \(n_{\text{vdW-DF}} = \sum_{i} n_{\text{vdW-DF},i}\).

The central step in the Harris scheme is the assumption that the density change \(n_{\text{sc}} - n_{\text{std}}\) produces only a small change in the effective potential,

\[
\Delta V_{\text{eff}}(r) = \phi_{n_{\text{sc}}}^g(r) - \phi_{n_{\text{std}}}^g(r) + n_{xc}^g [n_{\text{sc}}^g](r) - n_{xc}^g [n_{\text{std}}^g](r), \tag{25}
\]

so that one can expand the difference in KS and Harris estimates for the single-particle energy sum

\[
\sum_{\lambda} \epsilon_{\lambda}^g - \sum_{\lambda'} \epsilon_{\lambda'}^g = \int d^3 r \, n_{\text{sc}}^g(r) \Delta V_{\text{eff}}(r) + \mathcal{O}(n_{\text{sc}} - n_{\text{std}})^2. \tag{26}
\]

The linear term cancels out corresponding linear terms in the expansion of \(E_{\text{KS}}^g[n]\) and in the calculation of the electrostatic potential\(^{63}\).

The resulting single-shot (no density update) DFT estimate is also variational

\[
E_{\text{Harris}}[n_{\text{std}}] = E_{\text{Harris}}[n_{\text{sc}}] + \mathcal{O}(n_{\text{sc}} - n_{\text{std}})^2, \tag{27}
\]

but it is not, in general, an extremum.\(^{63}\) This follows because there is no consistency between the Harris scheme input density \(n_{\text{std}}\), and the single-particle electron density \(^{17}\) that results with the Harris-scheme effective potential \(V_{\text{eff}}[n_{\text{std}}](r)\).

**B. Nature of and error in non-selfconsistent vdW-DF calculations**

The non-selfconsistent vdW-DF total energy \(E_{\text{vdW-DF}}^g(d)\) is an approximation to the fully selfconsistent vdW-DF result \(E_{\text{sc}}^g\) of \(E_{\text{KS}}^g\) (31).

However, unlike a regular vdW-DF Harris interaction estimate \(E_{\text{Harris}}^g(d)\) it is built from the sc GGA result for the entire system \(n_{\text{sc}}\), and not the superposition of sc-vdW-DF fragment densities, \(n_{\text{vdW-DF}} = \sum_{i} n_{\text{vdW-DF},i} \neq n_{\text{sc}}\). The nsc-vdW-DF approximation \(E_{\text{vdW-DF}}^g\) can, however, still formally be seen as a further extension of the ideas that underpin the Harris estimate Eq. (23).

To establish a formal relation between sc and nsc vdW-DF calculations, we consider the differences in sc results that arise as we replace a GGA choice \(\hat{g}\) with a vdW-DF choice for the exchange-correlation functional. We introduce

\[
\delta n_{\text{sc}} = n_{\text{sc}}^\text{vdW-DF} - n_{\text{sc}}^g, \tag{28}
\]

\[
\Delta V_{\text{sc}} = \phi_{n_{\text{sc}}^\text{vdW-DF}} - \phi_{\bar{n}_{\text{sc}}^g} + \mu_{xc}^\text{vdW-DF} [n_{\text{sc}}^\text{vdW-DF}] - \mu_{xc}^g [n_{\text{sc}}^g], \tag{29}
\]

to identify the changes resulting in the density and in the effective potential, respectively. As in the original Harris analysis\(^{63}\) we can consider both \(\delta n_{\text{sc}}\) and \(\Delta V_{\text{sc}}\) small, and thus give rise only to linear changes

\[
\sum_{\lambda} \epsilon_{\lambda}^\text{vdW-DF} - \sum_{\lambda'} \epsilon_{\lambda'}^g \approx \int d^3 r \, n_{\text{sc}}^\text{vdW-DF} \Delta V_{\text{sc}}(r), \tag{30}
\]

\[
E_{\text{sc}}^\text{vdW-DF} \approx E_{\text{sc}}^g \approx \int d^3 r \, \delta n_{\text{sc}}(r) \mu_{xc}^\text{vdW-DF} [n_{\text{sc}}^g](r). \tag{31}
\]

Simply extending the analysis behind the Harris estimate therefore yields the formal relation

\[
E_{\text{sc}}^\text{vdW-DF} \approx E_{\text{KS}}^g + E_{\text{c}}^g [n_{\text{sc}}^g] + \Delta E_{\text{sc}}^\text{vdW-DF} \approx E_{\text{sc}}^g + \Delta E_{\text{sc}}^\text{vdW-DF} + \mathcal{O}(\Delta n_{\text{sc}})^2 \tag{32}
\]

\[
= E_{\text{vdW-DF}}^g(d) + \Delta E_{\mu,\text{sc}}^\text{vdW-DF} + \mathcal{O}(\Delta n_{\text{sc}})^2 \tag{33}
\]

with leading-order correction term

\[
\Delta E_{\mu,\text{sc}}^\text{vdW-DF} = \int d^3 r \, \delta n_{\text{sc}}(r) \left\{ \mu_{xc}^\text{vdW-DF} [n_{\text{sc}}^\text{vdW-DF}](r) - \mu_{xc}^g [n_{\text{sc}}^g](r) \right\}. \tag{34}
\]

**C. A sfd-vdW-DF scheme for accelerated calculations of molecular interactions**

We propose to pursue a Harris-type vdW-DF scheme that is based on the superposition \(n_{\text{std}}^g = n_{1}^g + n_{2}^g + \ldots\) of GGA fragment densities but which approximates the vdW-DF total energy by

\[
E_{\text{std}}^\text{vdW-DF}(d) \equiv E_{\text{Harris}}^g(d) + E_{\text{c}}^g [n_{\text{std}}^g] + \Delta E_{\text{sc}}^\text{vdW-DF} \approx E_{\text{Harris}}^g(d). \tag{35}
\]

Here, again, \(E_{\text{Harris}}^g(d)\) denotes the regular Harris estimate as described in the given GGA choice \(g\).
A formal analysis motivating the proposed vdW-DF approximation \cite{berlyand} is essentially already stated in Section III.B. We now consider slightly different density and effective-potential differences

\begin{align}
\delta n_{\text{sfd}} & \equiv n_{\text{sfd}}^{\text{vdW-DF}} - n_{\text{sfd}}^g, \\
\Delta V_{\text{sfd}} & \equiv \phi_{\text{sfd}}^{\text{vdW-DF}} - \phi_{\text{sfd}}^g \\
& + \mu_{\text{xc}}^{\text{vdW-DF}} [n_{\text{sfd}}^{\text{vdW-DF}}] - \mu_{\text{xc}}^g [n_{\text{sfd}}^g],
\end{align}

so that the \( g \to \text{vdW-DF} \) changes instead reflect the effects on the Harris-scheme single-particle eigenenergies \( \bar{\epsilon}_\lambda \).

The sfd-vdW-DF estimate can thus be expressed as an approximation to a regular vdW-DF Harris scheme

\begin{equation}
E_{\text{vdW-DF}}^{\text{Harris}}(d) = E_{\text{sfd}}^{\text{vdW-DF}}(d) + \Delta E_{\mu,\text{sfd}}^{\text{vdW-DF} \to g} + O(\delta n_{\text{sfd}})^2
\end{equation}

where we now have a slightly different leading-order correction term

\begin{equation}
\Delta E_{\mu,\text{sfd}}^{\text{vdW-DF} \to g} = \int d^3r \ \delta n_{\text{sfd}}(r) \ \{ \mu_{\text{xc}}^{\text{vdW-DF}} [n_{\text{sfd}}^{\text{vdW-DF}}](r) - \mu_{\text{xc}}^g [n_{\text{sfd}}^g](r) \}.
\end{equation}

V. COMPUTATIONAL DETAILS

This paper compares the interaction energy curves obtained with nsc Harris-type calculations with those of DFT calculations for selected non-covalently bound molecular dimers. Both sc and nsc calculations with the PBE version of GGA are performed using the Dacapo software. This planewave DFT code was chosen because it is straightforward in Dacapo to set the electronic densities equal to the sum of molecular (frozen input) densities \( n_{\text{sfd}} = n_1 + n_2 \) through an external manipulation in Asp\cite{asp} and thus to prepare the sfd calculations.

The non-local correlation energy is evaluated in a post-processing procedure both for regular DFT and Harris-type vdw-DF calculations. For these calculations, we use an efficient in-house real-space code, further described in Ref. \cite{asp} A radius cutoff of 6 Å is used for dense (full) sampling of the grid and a cutoff of 26 Å is used for sparse (double-spaced) sampling of the grid.

In the PBE calculations, relying on Vanderbilt ultrasoft pseudopotentials, we use plane-wave and density-sampling cutoffs of 500 eV. This cutoff choice has been used in many similar calculations\cite{vanderbilt,ultrasoft} and gives a relatively dense sampling of the density grid used to evaluate the non-local correlation. As long as the reference calculations have the same grid-sampling density, here secured by using the same size of the unit cell, the non-local correlation energy is typically converged to within about 1 meV.

FIG. 3: A comparison between the regular Harris scheme and self-consistent (sc) Kohn-Sham calculations for four different molecular pair configurations identified in the upper panel, all systems investigated in the PBE version of GGA. These are (left to right, top to bottom) the HF dimer in parallel configuration, the HF-benzene pair with the H of HF pointing towards the benzene center, the parallel benzene dimer, and the C60 dimer with the benzene rings facing each other. The middle panel shows the difference between the interaction energy in the sc and Harris PBE calculation. The lower panel shows (full curves) the interaction curves using the Harris functional and (dashed curves) the sc result. The abscissa label \( d \) denotes the separation between the closest atoms in separate molecules. The two curves involving the highly polar HF exhibit a non-negligible binding. The dotted vertical lines indicate the GGA-minimum. Among the investigated systems, the Harris estimate gives the largest overestimate (16\%) for HF-benzene interaction curve.
VI. RESULTS: ASSESSING THE SFD-VDW-DF EVALUATION

Four molecular pairs, depicted in the upper panel of Fig. 3, have been chosen for our comparison between the sfd scheme and regular DFT calculations. The first is a hydrogen fluoride (HF) dimer in parallel configuration. This configuration is not the optimal one but is here chosen as a representative for systems with large dipole-dipole interactions. The second is a molecular configuration where the hydrogen of HF points towards the center of a benzene molecule. Thus, one molecule has zero and the other a large dipole moment in vacuum. The third system is a benzene dimer in parallel sandwich configuration. The binding in this system is dominated by vdW (also called London dispersion) forces. The interaction in this system is representative of dilute sparse matter system, like a gas. The fourth system, a dimer of C60 with hexagonal rings facing each other, is also one where the binding is dominated by vdW forces. But because of the large size of C60, this attraction is much stronger than for the benzene dimer. This system is therefore more representative of compact molecular complexes that arise in bulk sparse matter.

A. Regular Harris scheme for GGA-PBE calculations

We first describe and illustrate the Harris scheme as it is used for GGA calculations. Obviously, the use of a GGA will not generally succeed in reproducing structural properties of typical sparse, weakly interacting molecular systems. Nevertheless, it is instructive to illustrate that the GGA Harris scheme is still generally able of faithfully reproducing the sc GGA calculations, including the sparse-matter GGA limitations.

Figure 3 compares the interaction curves for the four different molecular pairs as obtained with DFT and the Harris scheme using the PBE version of the GGA exchange-correlation functional. Only the HF dimer and the HF-benzene pair show an appreciable binding of respectively 173 and 180 meV using the Harris scheme for PBE and 158 and 156 meV using sc PBE calculations. For the parallel HF dimer system, which is dominated by dipole-dipole interactions, the Harris calculation overestimates the binding energy by 9% compared to regular GGA DFT calculations. For the HF-benzene system, where one of the molecules is highly polar and the other is not, the scheme overestimates the binding energy by 17%.

The discrepancies between the two methods can be understood from the significant dipole moment induced by the binding. At optimal separation (in the selected configurations), a dipole of 0.12 eÅ is induced for the HF dimer, while one of 0.15 eÅ is induced for the HF-benzene pair. These induced dipole moments are comparable to the dipole moment of the HF molecule itself (0.39 eÅ).

It is clear that molecular pairs involving one or more HF molecule(s) serve as tough tests for the feasibility of the Harris functional scheme.

For systems dominated by the vdW forces, the discrepancies are difficult to assess without including the effect of non-local correlation. We will therefore make this assessment in the next subsection.

FIG. 4: Comparison of the sfd-vdW-DF scheme for a benzene dimer and results obtained in nsc-vdW-DF evaluation. The top panel compares the differences between these two procedures as they arise for the total energy of vdW-DF1 (black curve with star dots) and vdW-DF2 (cyan/light curves with star dots); and the difference in the nonlocal correlation (two upper curves). The bottom panel compares the sfd-vdW-DF1 (black star-dotted curves) and sfd-vdW-DF2 (cyan/light star-dotted curves) against nsc-vdW-DF results (dashed curves). The somewhat larger shift (for most separations) in the non-local correlation energy when using the sfd scheme compared to the total vdW-DF energy, in particular for vdW-DF2, indicates a partial cancellation of the density sensitivity of the nonlocal and semilocal components.
B. Systems dominated by vdW attraction

The benzene dimer is a typical organic system bound by vdW forces. Since this system is weakly bonded, we can expect charge transfer to be small and thus the vdW-DF Harris scheme, and more generally the sfd framework, to be well suited to describe the system.

Figure 4 shows the comparison between the sfd and nsc vdW-DF calculations for the benzene dimer. The difference between the dashed and the full curves in the lower panel is barely distinguishable. At binding separation the sfd result is 2% below the nsc result for vdW-DF1 and merely 0.4% for vdW-DF2.

The upper panel of Fig. 4 shows that the non-local energy is somewhat affected by using the frozen density \( n_{\text{sfd}} = n_1 + n_2 \) in place of the one determined with a full GGA calculation, \( n_{\text{nsc}} \). It also reveals that there is some error cancellation between these shifts and the combined shifts in the other terms: the shifts obtained with the sfd scheme overestimate the non-local interaction energy, while the magnitude of the binding energy is underestimated. This trend is opposite to that exhibited for all four systems in the proper GGA Harris scheme, shown in Fig. 3. For vdW-DF2 this error cancellation is close to exact in a fairly wide region around the binding separation. For shorter separation between the molecules, corresponding to a larger density overlap, the absolute difference between schemes increases, as does the magnitude of the repulsive wall between molecules.

Figure 5 compares the two methods for a C60 dimer in the same fashion as for the benzene dimer. In this case the sfd-vdW-DF1 underestimates the binding energy by as little as 0.2%, while sfd-vdW-DF2 is spot on (within about 0.05 meV). This striking coincidence (arising from error cancellation) is likely fortuitous since the results are similar, but not this similar, in other regions of the interaction curve. For the C60 interaction curve, the sfd-vdW-DF calculations in some regions overestimate and in other regions underestimate the interaction energy.

The benzene and C60 dimer calculations indicate that the sfd scheme is an appropriate method to accelerate the evaluation of interaction energies in systems dominated by vdW interactions.

C. System with large induced charge: HF interacting with benzene

In the HF-benzene system the vdW forces contribute to the binding alongside electrostatic effects. The nsc-vdW-DF2 predicts a binding energy of 174 meV compared to that of 155 meV with the sc PBE calculations (in Fig. 3).

Figure 6 shows the results of the interaction curves obtained with sfd and nsc vdW-DF calculations. For this system, we also find that the vdW-DF2 calculation produces a larger binding energy than the vdW-DF1, which is opposite to the case for the benzene and for the C60 dimer. This switching of order is related to the fact that vdW-DF2 in general has a less repulsive exchange account and a less attractive non-local correlation account. Since the smaller size of this system decreases the magnitude of the non-local correlation, it shifts the balance between the repulsive and attractive terms.

The difference between the sfd and nsc vdW-DF calculations increases to as much as 30% for vdW-DF2, compared to 17% for the sfd and sc PBE calculations. The discrepancy is somewhat smaller for vdW-DF1. Note that this discrepancy arises mostly from the shift in the \( \Delta E_{xc}^{\text{vdW-DF1}} \) term and not from the non-local correlation, which contributes with 4 meV in the opposite direction of the total shift of \(-60\) meV.

Our results indicate that the increased inaccuracy of the sfd scheme for polar systems arises primarily from short-ranged effects. Thus, the inaccuracy may be reduced by starting from densities generated with revPBE exchange for sfd-vdW-DF1 and in the same vein PW86r for sfd-vdW-DF2.
FIG. 6: The sfd-vdW-DF description of binding between a HF molecule and benzene. Legends and details as in the lower panel of Fig. 1. The sizable discrepancy between the sfd and sc calculation in the PBE calculations, shown in Fig. 5, carries through to the case of (sfd and nsc) vdW-DF calculations. The non-local correlation has only a tiny effect on the discrepancy.

D. The nsc-vdW-DF approximation

We note that the leading-order difference \((34)\) between sc- and nsc-vdW-DF total-energy results is nominally linear in the density change, \(\delta n_{\text{sc}} = n^\text{sc,vdW-DF}_r - n^\text{sc}_r\). The nsc-vdW-DF calculations—while often very successful—need therefore not always be as robust as a regular Harris vdW-DF scheme would be.

On the other hand, the regular nsc-vdW-DF approach\(^2\) does have a mechanism for including some of the electron density rearrangement that arises from Pauli exclusion or from the formation of more traditional types of bonding (those that a GGA does capture). Fig. 1 shows that keeping such charge adjustments can be important for systems where at least one fragment has a large static polarization. Generally, we expect the nsc-vdW-DF approach to be more accurate except in cases with very weak intermolecular interactions. Of course, the only way to resolve the difference would be to perform a fully sc vdW-DF calculation\(^1\)(\(^1\)). The sc-vdW-DF is now becoming standard procedure for medium to large systems (system sizes approaching a thousand atoms). However, a performance testing comparing \(E^\text{sc,vdW-DF}_r\) and \(E^\text{sfd,vdW-DF}_r\) against \(E^\text{sc}_r = E^\text{KS}_r\) is for technical reasons beyond the scope of this paper (as discussed in Section II).

E. An organic-molecular testing of sfd-vdW-DF performance

Figs. 1 and 2 and Table 1 presented a summary of the further assessment we have performed of the accuracy of the sfd-vdW-DF scheme for the S22 benchmark suite.\(^1\) Here we provide some additional details.

Table 1 presents the calculated numbers of our comparison of sc-, nsc-, and sfd-vdW-DF1 and -vdW-DF2 results for interaction energies in the S22 set of molecular dimers. These interaction energies are all evaluated at the binding distance (identified in Ref. [19]) that minimizes respectively the sc-vdW-DF1 and sc-vdW-DF2 interaction-energy variation. The quantum chemistry computations are from Ref. [96]. For an actual S22-benchmarking of various sparse-matter DFT methods one should provide a full binding energy curve for each of the computational approaches. Here, our purpose is merely to complement our analysis based on binding curves of illustrative special cases with statistics for the S22 set of dimers that are seen as typical of organic-molecular interaction problems.

We note that system 1–7 can be labeled hydrogen-bonding dominated, while 8–15 can be labeled dispersion dominated, and the remainder mixed. By studying the table, it becomes clear that sfd tends to compare well with nsc results for dispersion-dominated systems, while the biggest discrepancies arise among the systems dominated by hydrogen bonds. This observation agrees well with our analysis based on Figs. 3–6.

Figure 2 conveys an overview and feeling for the quality of the sfd calculations compared to nsc and sc calculations. Together, the figure and table show that nsc and sc calculations are very similar. This is reassuring considering the fact that they are also based on different codes. As earlier discussed the sfd results compare well to the nsc results. Further, the inaccuracy introduced is overall smaller than the difference between sc-vdW-DF1 and QC results, while the inaccuracy is about equal to the difference between QC and vdW-DF2 results. vdW-DF2 has better performance for the S22 data set than vdW-DF1. The inaccuracy introduced by using sfd does not necessarily make results compare worse to QC.

VII. DISCUSSION

A. On vdW-bonding effects on electron dynamics

We begin with an interesting aside, noting the implications of our formal analysis on the expected error in nsc-vdW-DF and in sfd-vdW-DF, Eqs. (34) and (39).

It is known that the inclusion of nonlocal correlation and vdW forces often gives rise to indirect bandstructure effects\(^5\)\(^,\)\(^6\)\(^,\)\(^7\)\(^,\)\(^8\)\(^,\)\(^9\)\(^,\)\(^10\)\(^,\)\(^11\) because the vdW binding changes the morphology and hence the local environment for the electron dynamics. However, it is also interesting to identify conditions where one can also expect direct vdW bandstructure effects, that is, electron-dynamics changes that arise when—for given structure—the effective potential is changed from a GGA to the vdW-DF functional form.

The error estimate \((34)\) allows us to identify conditions for expecting the vdW-bonding to affect the bandstructure and, more generally, the electron dynamics.\(^8\)\(^,\)\(^10\) A
TABLE I: Interaction energies for pairs of small molecules from the S22 dataset. Quantum chemistry (QC) results from Ref. 96; the selfconsistant (sc) vdW-DF1 and vdW-DF2 results are from Ref. 19. All energies in meV/dimer.

| #  | Duplex                             | $E_{vdW-DF1}^{\text{sfd}}$ | $E_{vdW-DF1}^{\text{nsc}}$ | $E_{vdW-DF1}^{\text{sc}}$ | $E_{vdW-DF2}^{\text{sfd}}$ | $E_{vdW-DF2}^{\text{nsc}}$ | $E_{vdW-DF2}^{\text{sc}}$ | QC  |
|----|-----------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----|
| 1  | Ammonia dimer                      | 118                         | 111                         | 115                         | 136                         | 126                         | 134                         | 137 |
| 2  | Water dimer                        | 186                         | 171                         | 185                         | 220                         | 201                         | 218                         | 218 |
| 3  | Formic acid dimer                  | 736                         | 680                         | 690                         | 817                         | 745                         | 766                         | 815 |
| 4  | Formamide dimer                    | 610                         | 560                         | 587                         | 684                         | 619                         | 655                         | 699 |
| 5  | Uracil dimer                       | 807                         | 749                         | 767                         | 876                         | 805                         | 832                         | 897 |
| 6  | 2-pyridoxine – 2-aminopyridine     | 671                         | 623                         | 639                         | 728                         | 661                         | 687                         | 737 |
| 7  | Adenine – thymine                  | 652                         | 619                         | 609                         | 716                         | 661                         | 660                         | 726 |
| 8  | Methane dimer                      | 38                          | 39                          | 36                          | 29                           | 30                          | 30                           | 23  |
| 9  | Ethene dimer                       | 70                          | 67                          | 64                          | 65                           | 61                          | 65                           | 65  |
| 10 | Benzene – methane                  | 72                          | 70                          | 68                          | 64                           | 62                          | 63                           | 63  |
| 11 | Benzene dimer (slip-parallel)      | 136                         | 141                         | 136                         | 120                         | 124                         | 123                         | 114 |
| 12 | Pyrazine dimer                     | 189                         | 188                         | 185                         | 178                         | 177                         | 177                         | 182 |
| 13 | Uracil dimer (stacked)             | 414                         | 396                         | 403                         | 412                         | 391                         | 402                         | 422 |
| 14 | Indole – benzene (stacked)         | 231                         | 232                         | 206                         | 234                         | 199                         | 197                         | 199 |
| 15 | Adenine – thymine (stacked)        | 466                         | 456                         | 461                         | 457                         | 457                         | 466                         | 506 |
| 16 | Ethene – ethine                    | 74                          | 68                          | 69                          | 74                           | 67                          | 70                           | 65  |
| 17 | Benzene – water                    | 142                         | 125                         | 124                         | 148                         | 126                         | 129                         | 143 |
| 18 | Benzene – ammonia                  | 104                         | 97                          | 94                          | 101                         | 91                          | 92                           | 101 |
| 19 | Benzene – HCN                      | 194                         | 162                         | 166                         | 198                         | 159                         | 170                         | 197 |
| 20 | Benzene dimer (T-shape)            | 120                         | 115                         | 113                         | 108                         | 101                         | 105                         | 118 |
| 21 | Indole – benzene (T-shape)         | 240                         | 210                         | 214                         | 230                         | 192                         | 206                         | 243 |
| 22 | Phenol dimer                       | 267                         | 252                         | 254                         | 290                         | 270                         | 279                         | 307 |

A clear difference in sc and nsc vdW-DF total energies is required in order for such direct vdW-DF bandstructure effects to emerge.

Meanwhile, there now exists a significant experience with using vdW-DF for sparse matter systems and the calculations have shown that there is, in practice, often only limited differences in nsc-vdW-DF and sc-vdW-DF calculations. It follows that one must, in general, expect that direct bandstructure effects typically are small.

B. Overall assessment of the sfd-vdW-DF calculation scheme

The speed up gained when using the sfd calculations with DACapo are substantial yet somewhat modest. Computational costs are reduced by 40% and 55% for the benzene and C60 dimer respectively, when using standard cutoffs with a minimal number of bands. Considering that this software usually requires about 20 electronic iterations to converge at these system sizes (but more for large systems), this gain is less than one might anticipate.

However, we should consider that standard software like DACapo (that we here use) has been subjected to intense efforts to optimize its ability to simultaneously solve the problem of charge relaxation and determination of the KS eigenvalues. What formally constitutes Harris calculations in that code are today primarily used to obtain accurate values for the KS eigenvalues. We do not desire such enhanced accuracy for an actual sfd-vdW-DF study.

The fact that the here-proposed sfd-vdW-DF is still faster than nsc-vdW-DF (Sec. VI) is therefore promising. Furthermore, since the performance, documented here, is excellent for many molecular systems, there is room for more compromise on accuracy. We believe that the present results motivate the approach to be further evaluated in forthcoming studies.

C. Towards as fast biomolecular mapping of vdW interactions in biomolecular systems

We are ultimately interested in vdW-DF computational studies for large-scale interaction problems where there are relevant speed ups to be gained by not seeking the fully sc density (as described either in a full GGA or in a full vdW-DF study). The acceleration must come from minimizing the cost of DFT calculations of the steric-hindrance or kinetic-energy repulsion effects (as this is the large-system bottleneck).

We note that a parallel study, Ref. 76, pursues a closely related computational strategy and begins a first-
principle DFT mapping of the morphology variation on the vdW attraction in a DNA dimer by first focusing on an efficient evaluation of $E_{nl}^{|sfd|}$. The DNA dimer system is then taken as an example of a typical large-scale biomolecular interaction problem.

The vdW-DF exploration that is proposed in Ref. 76 bypasses the need for performing the expensive computations of the kinetic-energy repulsion term (which, in any case, is not relevant outside the binding regime that extends to a nearest-atoms separation of about 4.0 Å, Ref. 30). The first-principle vdW-DF survey of DNA attains thus achieves a dramatic speed up but can still be applicable for large systems bound by vdW forces, the next step would be to investigate if the scheme can be further accelerated, in particular for large supramolecular systems. To this end, implementations for a full sfd-vdW-DF scheme for other DFT codes (beyond Dacapo) are being tested.

VIII. SUMMARY AND OUTLOOK

To accelerate large-scale vdW-DF characterization of biomolecular systems it seems useful to adapt the ideas of the Harris scheme 63,67,72 as is indicated and explored here. A Harris-type approximation which works reasonably well for describing the kinetic-energy effects of forming covalent bonds of atoms in some molecules should have a good chance of describing the simpler kinetic-energy repulsion (steric hindrance) of molecules in supramolecular systems.

Here, we have put this expectation to the test. Our results indicate that this scheme is promising for describing supramolecular systems bonded primarily by vdW forces. However, if one or more fragments are highly polar, this comes at the cost of accuracy.

This paper is also supplemented by a related publication, which presents a vdW-DF study that maps out the nonlocal correlation of large biopolymers within the presented sfd-vdW-DF scheme.

The pair of papers suggest a possible computational strategy for the study of binding in large supramolecular systems. The suggestion is to begin the structure and interaction-morphology search by essentially cost-free evaluations of the $E_{nl}^{|sfd|}$ variation. That $E_{nl}^{|sfd|}$ step is available simply from relatively cheap calculations of fragment electron densities. One can in turn search for relevant binding motifs, given the linking to the here-proposed and tested sfd-vdW-DF scheme. This strategy eventually leads to a complete vdW-DF estimate of the variation in total interaction energy. The strategy exists as an alternative to implementing a real-space vdW-DF version in a code that realizes genuine order-N scaling for large systems.

Having established the promise of the sfd scheme for systems bound by vdW forces, the next step would be to investigate if the scheme can be further accelerated, in particular for large supramolecular systems. To this end, implementations for a full sfd-vdW-DF scheme for other DFT codes (beyond Dacapo) are being tested.

Acknowledgments

The authors thank Pieremanuele Canepa and Timo Thonhauser for discussions and Kuyho Lee for supplying atomic coordinates for a full range of 14 various separations for each of the dimers in the S22 benchmark set. Partial support from the Swedish Research Council (VR) (two grants) and the Chalmers Area of Advance ‘Materials’ is gratefully acknowledged. The computations were performed on high-performance computing resources provided by the Swedish National Infrastructure for Computing (SNIC) at the C3SE and HPC2N metanodes.

* Electronic address: schroder@chalmers.se; Corresponding author

1. K. Burke, J. Chem. Phys. 136, 150901 (2012).
2. D.C. Langreth, B.I. Lundqvist, S.D. Chakarova-Käck, V.R. Cooper, M. Dion, P. Hyldgaard, A. Kelkkannen, J. Kleis, L. Kong, S. Li, P.G. Moses, E. Murray, A. Puzder, H. Rydberg, E. Schröder, and T. Thonhauser, J. Phys.: Cond. Matter 21, 084203 (2009).
3. G.D. Mahan, J. Chem. Phys. 43, 1569 (1965).
4. K. Rapcewicz and N.W. Ashcroft, Phys. Rev. B 44, 4032 (1991).
5. Y. Andersen, D.C. Langreth, and B.I. Lundqvist, Phys. Rev. Lett. 76, 102 (1996).
6. H. Rydberg, B.I. Lundqvist, D.C. Langreth, and M. Dion, Phys. Rev. B 62, 6997 (2000).
groups, [http://fy.chalmers.se/~scheroder/vdWDF](http://fy.chalmers.se/~scheroder/vdWDF)

S.V. Aradhya, M. Frei, M.S. Hybertsen, and L. Venkataraman, Nature Materials 11, 872-876, (2012).

N. Marom, A. Tkatchenko, S. Kapishnikov, L. Kronik, and L. Leiserowitz, Cryst. Growth. Des. 11, 3332 (2011).

J. Harris, Phys. Rev. B 31, 1770 (1985).

G.D. Bellchambers and F.R. Manby, J. Chem. Phys. 135, 084105 (2011).

W.M.C. Foulkes and R. Haydock, Phys. Rev. B 39, 12520 (1989).

V.K. Nikulin, Zh. Tekhn. Fiz. XLI, 41 (1971) [Sov. Phys. - Techn. Phys. 16, 28 (1971)].

R.G. Gordon and Y.S. Kim, J. Chem. Phys. 56, 3122 (1972).

The approach is sometimes also called the Harris-Foulkes scheme.

While beyond the present scope, we note that the Harris scheme with refinements can, for instance, be used for an approximate system characterization along a reaction pathway in organic systems.

P. Jurečka, J. Šponer, J. Černý, and P. Hobza, Phys. Chem. Chem. Phys. 8, 1985 (2006).

E. Londero, P. Hyldgaard, and E. Schröder, A van der Waals density functional mapping of vdW attraction in DNA dimers, preprint (2013).

A single electronic iteration is here taken to mean an evaluation stage during which the external potential is kept fixed. In most DFT codes, there are several wavefunction evaluation steps in each such electronic iteration. When pursuing a standard sc DFT determination, the requested accuracy on the wavefunction solution remains moderate. For bandstructure calculations, the DFT code is formally performing a Harris-type evaluation, but a code like Dacapo then forces many wavefunction-evaluation steps to deliver a higher accuracy on the wave functions. However, that is not the spirit of the original Harris scheme, nor do we intend this here with the proposed sfd-vdW-DF scheme.

Open-source, plane-wave-based DFT code DACAPO, [http://www.fysik.dtu.dk/CMAPOS/](http://www.fysik.dtu.dk/CMAPOS/)

D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. 45, 566 (1980).

S.H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. 58, 1200 (1980).

J.P. Perdew and Y. Wang, Phys. Rev. B 45, 13244 (1992).

J.P. Perdew, Phys. Rev. B 55, 1665 (1985).

D.C. Langreth and M.J. Mehl, Phys. Rev. B 28, 1809 (1983); D.C. Langreth and J.P. Perdew, Phys. Rev. B 21, 5469 (1980).

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

P. Elliott and K. Burke, Can. J. Chem. 87, 1485 (2009).

Y. Zhang and W. Yang, Phys. Rev. Lett. 80, 890 (1998).

J.P. Perdew and Y. Wang, Phys. Rev. B, 33, 8800 (1986).

J. Klimeš, D.R. Bowler, and A. Michaelides, J. Phys.: Condens. Matter 22, 022201 (2010).

BEEF, an approach to select the exchange choice for vdW-DF calculations to a specific target application, based on Bayesean logic and training, J. Wellendorff, K.T. Lundgaard, A. Mogelhøj, V. Petzold, D.D. Landis, J.K. Nørskov, T. Bligård, and K.W. Jacobsen, Phys. Rev. B 85, 235149 (2012).

One can of course evaluate the KS and the Harris schemes as well as the sfd-vdW-DF approximation for a general specification of the fragment orientation, motif-alignment, and arbitrary choice of displacement vector.

E. Zaremba, J. Phys.: Cond. Matter 2, 2479 (1990).

B. Farid, V. Heine, G.E. Engel, and I.J. Robertson, Phys. Rev. B 48, 11602 (1993).

Python-based atomic simulation environment, see [http://wiki.fysik.dtu.dk/ase/](http://wiki.fysik.dtu.dk/ase/)

W. Klopper, M. Quack, and M.A. Suhm, J. Chem. Phys. 108, 10096 (1998).

R.A. Oriani, C.P. Smyth, J. Am. Chem. Soc. 70, 125 (1948).

T. Takatani, E.G. Hohenstein, M. Malagoli, M.S. Marshall, and C.D. Sherrill, J. Chem. Phys. 132, 144104 (2010).

N. Marom, J. Bernstein, J. Garel, A. Tkatchenko, E. Joelvich, L. Kronik, and O. Hod, Phys. Rev. Lett. 105, 046801 (2010).