IOP Publishing Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **26** (2014) 213202 (16pp) doi:10.1088/0953-8984/26/21/213202

Topical Review

Many-body van der Waals interactions in molecules and condensed matter

Robert A DiStasio Jr[1], Vivekanand V Gobre[2] and Alexandre Tkatchenko[2]

[1] Department of Chemistry, Princeton University, Princeton, NJ 08544, USA
[2] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4–6, 14195, Berlin, Germany

E-mail: tkatchenko@fhi-berlin.mpg.de

Received 8 January 2013, revised 23 January 2014
Accepted for publication 21 February 2014
Published 8 May 2014

Abstract

This work reviews the increasing evidence that many-body van der Waals (vdW) or dispersion interactions play a crucial role in the structure, stability and function of a wide variety of systems in biology, chemistry and physics. Starting with the exact expression for the electron correlation energy provided by the adiabatic connection fluctuation–dissipation theorem, we derive both pairwise and many-body interatomic methods for computing the long-range dispersion energy by considering a model system of coupled quantum harmonic oscillators within the random-phase approximation. By coupling this approach to density functional theory, the resulting many-body dispersion (MBD) method provides an accurate and efficient scheme for computing the frequency-dependent polarizability and many-body vdW energy in molecules and materials with a finite electronic gap. A select collection of applications are presented that ascertain the fundamental importance of these non-bonded interactions across the spectrum of intermolecular (the S22 and S66 benchmark databases), intramolecular (conformational energies of alanine tetrapeptide) and supramolecular (binding energy of the 'buckyball catcher') complexes, as well as molecular crystals (cohesive energies in oligoacenes). These applications demonstrate that electrodynamic response screening and beyond-pairwise many-body vdW interactions—both captured at the MBD level of theory—play a quantitative, and sometimes even qualitative, role in describing the properties considered herein. This work is then concluded with an in-depth discussion of the challenges that remain in the future development of reliable (accurate and efficient) methods for treating many-body vdW interactions in complex materials and provides a roadmap for navigating many of the research avenues that are yet to be explored.

Keywords: van der Waals dispersion interactions, intermolecular interactions, electron correlation, random phase approximation, density functional theory, adiabatic connection fluctuation-dissipation theorem

(Some figures may appear in colour only in the online journal)

1. Introduction

The importance of van der Waals (vdW) interactions in the structure, stability and function of molecules and materials can hardly be overemphasized [1–6]. Ubiquitous in nature, vdW interactions span a vast array of physically relevant distances, ranging from just a few Ångström to several nanometers [7–9], with recent experiments suggesting that vdW forces can even be significantly longer-ranged [10]. These non-bonded interactions are largely responsible for the formation of the gas-phase benzene dimer at low temperatures, the stabilization required for the formation of molecular crystals, and the binding of molecules to proteins and DNA inside living cells. In addition, vdW interactions play a central role in the fields of supramolecular chemistry and nano-materials, in which non-covalent binding is essential for structure and functionality.

In order to enable rational predictions and design of molecular and condensed-matter materials, including the interfaces between them, a reliable first-principles method is required that can describe vdW interactions both accurately and efficiently. However, forming an accurate description of vdW interactions is extremely challenging, since the vdW dispersion energy arises from the correlated motion of electrons and must be described by quantum mechanics. The rapid increase in computational power coupled with recent advances in the development of theoretical models for describing vdW interactions have allowed us to achieve so-called 'chemical accuracy' for binding between small organic molecules. However, the lack of accurate and efficient methods for treating large and complex systems hinders truly quantitative predictions of the properties and functions of technologically and biologically relevant materials.

In this topical review, we are mainly concerned with methods for describing long-range vdW interactions in the context of density functional theory (DFT). Therefore, we will not discuss the subtle effects that become important for macroscopic systems, such as relativistic retardation and effects of finite temperature. We refer the readers interested in this subject to standard textbooks [1, 2] and existing literature, e.g., references [11–13]. We remark that retardation and finite temperature effects will have to be eventually incorporated into microscopic approaches to vdW interactions to achieve a seamless link to macroscopic systems.

Many encouraging approaches have been proposed in recent years for approximate inclusion of long-range pairwise dispersion interactions in DFT [14–23]. Despite such significant progress in the field of modeling vdW interactions, many questions still remain and further development is required before a universally applicable method emerges. For instance, pairwise interatomic vdW methods are frequently employed to describe organic molecules adsorbed on inorganic surfaces [24–27], ignoring the relatively strong electrodynamic response screening present within bulk materials. On the flip side of the coin, the popular non-local vdW-DF functionals [28–30] utilize a homogeneous dielectric approximation for the polarizability, which is not expected to be accurate for molecules. Despite this fact, interaction energies between small organic molecules computed with such functionals turn out to be reasonably accurate. Understanding the physical reasons as to why these different approaches 'yield good results' outside of their expected domains of applicability is central to the development of more robust approximations.

Interatomic pairwise dispersion approaches based on the standard C_6/R^6 summation formula were popularized by the DFT-D method of Grimme [16] and are now among the most widely used methods [15, 18, 19] for including the dispersion energy in DFT. Despite their simplicity, these pairwise-additive models provide remarkable accuracy when applied to small molecular systems, especially when accurate dispersion coefficients (C_6) are employed for the atoms in molecules [31, 32]. Only recently have efforts been focused on going beyond the pairwise treatment of vdW contributions; for example, the importance of the non-additive three-body interatomic Axilrod–Teller–Muto term [33–35] was assessed, as well as the role of non-local screening in solids [36] and molecules adsorbed on surfaces [37]. Furthermore, an efficient and accurate interatomic many-body dispersion (MBD) approach has recently been proposed [38], which demonstrated that a many-body description of vdW interactions is essential for extended molecules and molecular solids, and that the influence of many-body vdW interactions can already become significant when considering the binding between relatively small organic molecules [38, 39].

In this work, we present a derivation of the pairwise and many-body interatomic dispersion energy for a model system comprised of isotropic polarizable dipoles from the adiabatic connection fluctuation–dissipation (ACFD) formula, which is an exact expression for the exchange–correlation energy. We distinguish and discuss two types of interatomic many-body contribution to the dispersion energy, which stem from beyond-pairwise non-additive interactions and self-consistent electrodynamic response screening. By using the ACFD formula we gain a deeper understanding of the approximations made in interatomic approaches, in particular the DFT+MBD method [38], providing a powerful formalism for the further development of accurate and efficient methods for computing the vdW dispersion energy.

Applications of the DFT+MBD method are presented for a wide variety of systems with finite electronic gaps, including benchmark databases of intermolecular interactions, the stability of extended and globular conformations of alanine tetrapeptide, the binding in the 'buckyball catcher' supramolecular host guest complex, and the cohesive energy of several oligoacene molecular crystals. For all of these cases, the role of the beyond-pairwise non-additive vdW interactions and electrodynamic response screening captured at the DFT+MBD level of theory is critically assessed and shown to contribute in a quantitative, and sometimes even qualitative, fashion. We conclude with a discussion of the challenges that remain in the future development of accurate and efficient methods for treating many-body vdW interactions in materials of increasing complexity.

As the modeling of vdW interactions is currently a very active field of research, it is impossible to cover all of the important developments in this article. For more information, we refer interested readers to the Ψ_k highlight [40] and review article [41] by Dobson and Gould, which discuss several different approaches for computing dispersion interactions; to the recent review by Klimes and Michaelides on dispersion methods within DFT [42]; and to the webpage for the recent vdW@CECAM workshop that brought together many of the key players in the development and application of vdW-inclusive first-principles methods [43].

2. Theory

The adiabatic connection fluctuation–dissipation (ACFD) theorem provides a general and exact expression for the exchange–correlation energy [44, 45], thereby allowing for the calculation of the dispersion energy in a seamless and accurate fashion which naturally incorporates higher-order many-body effects. In this section, we explore the use of the ACFD theoretical framework as a basis for the understanding and development of interatomic pairwise and many-body

dispersion methods. Beginning with a brief review of the ACFD correlation energy within the random-phase approximation (RPA), we then consider the ACFD-RPA correlation energy for a model system comprised of quantum harmonic oscillators (QHO) interacting via the dipole–dipole interaction potential. We derive the well-known C_6/R^6 interatomic pairwise summation formula from the second-order expansion of the ACFD-RPA correlation energy for this collection of N QHOs, each of which is characterized by an isotropic frequency-dependent point dipole polarizability. We then extend our treatment to account for spatially-distributed dipole polarizabilities and derive the corresponding Coulomb and dipole–dipole interaction potentials that naturally attenuate short-range interactions, thereby avoiding the near-field divergence associated with interacting point dipoles. By mapping the atoms in a molecular system of interest to this collection of QHOs, we then demonstrate how the Dyson-like self-consistent screening (SCS) equation of classical electrodynamics can be utilized to obtain the non-local response function as well as accurate molecular and atomic frequency-dependent dipole polarizabilities, all of which reflect the anisotropy arising from the underlying topology of the chemical environment as well as collective polarization and depolarization effects. This section is then concluded with a discussion of the DFT+MBD method, which is defined by the coupling of standard DFT functionals with the Hamiltonian corresponding to the aforementioned model system of QHOs utilizing a range-separated Coulomb (and dipole–dipole interaction) potential, allowing us to treat the full range of exchange and correlation effects in molecular systems of interest.

2.1. The ACFD-RPA correlation energy expression

For a system of nuclei and electrons, the ACFD theorem provides us with an exact expression for the exchange–correlation energy in terms of the density–density response function $\chi(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega)$ [44, 45], which measures the electronic response of the system at a point $\mathbf{r}$ due to a frequency-dependent electric field perturbation at a point $\mathbf{r}'$. Since the focus of this work is on dispersion interactions, which are quantum mechanical phenomena due to the instantaneous (dynamical) correlation between electrons, we begin with the ACFD formula for the correlation energy only (Hartree atomic units are assumed throughout):

$$E_\mathrm{c} = -\frac{1}{2\pi}\int_0^\infty \mathrm{d}\,\omega \times \int_0^1 \mathrm{d}\,\lambda \mathbf{Tr}[(\chi_\lambda(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega) - \chi_0(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega))v(\mathbf{r}, \mathbf{r}')]. \quad (1)$$

In this expression, $\chi_0(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega)$ is the bare or non-interacting response function, $\chi_\lambda(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega)$ is the interacting response function at Coulomb coupling strength λ, $v(\mathbf{r}, \mathbf{r}') = |\mathbf{r}-\mathbf{r}'|^{-1}$ is the Coulomb potential and **Tr** denotes the spatial trace operator (or six-dimensional integration) over the variables $\mathbf{r}$ and $\mathbf{r}'$.

Within the ACFD formalism, the adiabatic connection between a reference non-interacting system (defined at $\lambda = 0$) and the fully interacting system (with $\lambda = 1$), yields the correlation energy of the system of interest, which contains the many-body dispersion energy as well as other electron correlation effects. The interacting response function, χ_λ, is defined self-consistently via the Dyson-like screening equation, $\chi_\lambda = \chi_0 + \chi_0(\lambda v + f_\lambda^{\mathrm{xc}})\chi_\lambda$, which contains $f_\lambda^{\mathrm{xc}}(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega)$, the exchange-correlation kernel, an unknown quantity which must be approximated in practice. Neglecting the explicit dependence of f_λ^{xc} on the coupling constant allows for analytic integration over λ in the ACFD correlation energy expression in equation (1), and forms the basis for the most widely employed approximation for f_λ^{xc}, namely the random-phase approximation (RPA) [46, 47]. In what follows, we utilize the RPA, wherein $f_\lambda^{\mathrm{xc}} = 0$, which has been shown to yield reliable results for a wide variety of molecules and extended systems [48–65]. The ACFD-RPA correlation energy expression can be written as a power series expansion in $\chi_0 v$, following elimination of χ_λ using the Dyson equation and analytical integration over λ (see equation (1)) [66, 67]:

$$E_{\mathrm{c,RPA}} = -\frac{1}{2\pi}\int_0^\infty \mathrm{d}\,\omega \sum_{n=2}^\infty \frac{1}{n}\mathbf{Tr}[(\chi_0(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega)v(\mathbf{r}, \mathbf{r}'))^n]. \quad (2)$$

Using the Adler–Wiser formalism [68, 69], correlation energy calculations based on the ACFD and ACFD-RPA formulae in equations (1) and (2), respectively, are most often based on a reference bare response function computed using the set of single-particle occupied and virtual orbitals $\{\phi_i\}$ with corresponding energies $\{\epsilon_i\}$ and occupation numbers $\{f_i\}$ determined from semi-local DFT, Hartree-Fock or hybrid self-consistent field calculations, i.e.,

$$\chi_0(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega) = \sum_{ij}(f_i - f_j)\frac{\phi_i^*(\mathbf{r})\phi_i(\mathbf{r}')\phi_j^*(\mathbf{r}')\phi_j(\mathbf{r})}{\epsilon_i - \epsilon_j + \mathrm{i}\omega}. \quad (3)$$

Since the reference bare response function is not uniquely defined and response functions computed in this manner are computationally demanding (in terms of both construction and storage), it can be argued that such functions do not necessarily provide the ideal starting point for computing the ACFD correlation energy. In this work, we directly address these issues by positing an alternative framework for computing the bare or non-interacting response function based on a model system of QHOs—a scheme which is computationally efficient while also providing a vastly improved starting point for computing the long-range many-body dispersion energy for non-metallic systems.

Here we remind the reader that the ACFD-RPA expression is an approximation to the exact correlation energy, and the choice of χ_0, the reference bare response function, is not only the subject of this work but also the focus of current research in the field of DFT-based RPA theory. Moreover, we note in passing that this expression for the ACFD-RPA correlation energy is also not unique—for instance, the ACFD-RPA correlation energy can equivalently be written in terms of the fully interacting response function, i.e., through a power series expansion in $\chi_1 v$. For a more detailed review of the RPA approach for computing the correlation energy, see [40, 63, 65, 70] and references therein.

2.2. Derivation of the long-range interatomic pairwise dispersion energy

We now apply the ACFD-RPA approach to compute the correlation energy for a model system comprised of a collection of interacting QHOs. In doing so, we will first derive the standard C_6/R^6 interatomic pairwise summation formula for a system of two QHOs as the second-order expansion of the ACFD-RPA correlation energy within the dipole approximation. We will then demonstrate the validity of this formula for an arbitrary collection of N QHOs, providing a quantum mechanical derivation of the long-range interatomic pairwise summation formula that is customarily utilized for computing the dispersion energy.

To evaluate the ACFD-RPA correlation energy expression in equation (2), we first need the bare or non-interacting response function for this collection of QHOs, which can be assembled as a direct sum over the *individual* QHO response functions, $\chi_0(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega) = \chi_{0,p}(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega) \oplus \chi_{0,q}(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega) \oplus \cdots$. For a QHO located at position vector $\mathbf{R}_p = \{x_p, y_p, z_p\}$ and characterized by an isotropic frequency-dependent point dipole polarizability, $\alpha_p(\mathrm{i}\omega)$, the individual QHO response function takes on the following form [70]:

$$\chi_{0,p}(\mathbf{r}, \mathbf{r}', \mathrm{i}\omega) = -\alpha_p(\mathrm{i}\omega)\nabla_{\mathbf{r}}\delta^3(\mathbf{r}-\mathbf{R}_p)\otimes\nabla_{\mathbf{r}'}\delta^3(\mathbf{r}'-\mathbf{R}_p), \quad (4)$$

in which $\delta^3(\mathbf{r}-\mathbf{r}')$ is the three-dimensional Dirac delta function and $\otimes$ is the tensor (outer) product. An important point to note here is the stark contrast between this bare response function, which is constructed from completely localized (and thereby non-overlapping) individual QHO response functions, and the typical bare response function assembled using the aforementioned Adler–Wiser formalism (see equation (3)), which is constructed from the occupied and virtual *molecular* orbitals obtained from an underlying self-consistent field calculation and is therefore a relatively delocalized object by construction.

For the moment, we will assume that the QHOs are separated by a sufficiently large distance, allowing us to use the bare dipole–dipole interaction potential to describe the interoscillator couplings, a condition that will be relaxed when the general case is considered in the next section. This dipole–dipole interaction potential between oscillators p and q is straightforwardly obtained from the bare Coulomb potential, $v_{pq} = |\mathbf{R}_p - \mathbf{R}_q|^{-1}$, via

$$\mathfrak{T}_{pq} = \begin{cases} \nabla_{\boldsymbol{R}_p}\otimes\nabla_{\boldsymbol{R}_q} v_{pq} & \text{if } p \neq q \\ 0 & \text{if } p = q \end{cases} \quad (5)$$

and is therefore a 3×3 second-rank tensor with components given by

$$\mathfrak{T}_{pq}^{ab} = -\frac{3R_aR_b - R_{pq}^2\delta_{ab}}{R_{pq}^5}, \quad (6)$$

in which a and b represent the coordinates $\{x, y, z\}$ in the Cartesian reference frame, R_a and R_b are the respective components of the interoscillator distance R_{pq}, and δ_{ab} is the standard Kronecker delta.

With the individual QHO response functions and the dipole–dipole interaction tensor as defined above, we now consider the quantity $\chi_0 v$ in the ACFD-RPA correlation energy expression in equation (2), which can now be represented in matrix form as the product AT. Here, A is a diagonal $3N \times 3N$ matrix with $-\alpha_p(\mathrm{i}\omega)$ values on the 3×3 diagonal atomic subblocks, representing the bare or non-interacting response function for the collection of N QHOs. The dipole–dipole interaction matrix T is a $3N \times 3N$ matrix comprised of the 3×3 blocks of the $\mathfrak{T}_{pq}$ tensor as defined in equations (5) and (6).

For a system composed of two QHOs separated by a distance $R = |\mathbf{R}_p - \mathbf{R}_q|$ along the z-axis and characterized by isotropic point dipole polarizabilities $\alpha_p(\mathrm{i}\omega)$ and $\alpha_q(\mathrm{i}\omega)$, the AT matrix takes on the following form:

$$\chi_0 v \Leftrightarrow AT = \begin{pmatrix} 0 & 0 & 0 & -\frac{\alpha_p(\mathrm{i}\omega)}{R^3} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{\alpha_p(\mathrm{i}\omega)}{R^3} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{2\alpha_p(\mathrm{i}\omega)}{R^3} \\ -\frac{\alpha_q(\mathrm{i}\omega)}{R^3} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{\alpha_q(\mathrm{i}\omega)}{R^3} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2\alpha_q(\mathrm{i}\omega)}{R^3} & 0 & 0 & 0 \end{pmatrix}. \quad (7)$$

With the above matrix as input, the second-order ($n = 2$) term of the ACFD-RPA correlation energy expression in equation (2) yields the following energy expression for a collection of $N = 2$ QHOs:

$$\begin{aligned} E_{c,\mathrm{RPA\text{-}QHO}}^{(2)} &= -\frac{1}{2\pi}\int_0^\infty \mathrm{d}\omega\ \alpha_p(\mathrm{i}\omega)\alpha_q(\mathrm{i}\omega)\mathbf{Tr}[(\mathfrak{T}_{pq})^2] \\ &= -\frac{C_6^{pq}}{R^6} \quad (N=2) \end{aligned} \quad (8)$$

where we have used the fact that $\mathbf{Tr}[(\mathfrak{T}_{pq})^2] = 6/R_{pq}^6$ and the Casimir–Polder integral

$$C_6^{pq} = \frac{3}{\pi}\int_0^\infty \mathrm{d}\omega\ \alpha_p(\mathrm{i}\omega)\alpha_q(\mathrm{i}\omega) \quad (9)$$

to determine the C_6^{pq} dispersion coefficient from the corresponding pair of frequency-dependent point dipole polarizabilities. The above equation is nothing more than the familiar expression for the long-range dispersion interaction between two spherical atoms.

To demonstrate the validity of this formula for an arbitrary collection of N QHOs, one needs to consider the action of the spatial trace operator in equation (2) on the general $3N \times 3N$ AT matrix. As seen above, the second-order term in the ACFD-RPA correlation energy expansion requires the trace of the square of the AT matrix, for which the n-th diagonal element is simply the scalar product between the corresponding n-th column and n-th row of AT. As such, the overall trace corresponds to an accumulated sum of the diagonal elements contained in the smaller $(\mathfrak{T}_{pq})^2$ subblocks $\forall\ p, q$, each weighted by the product $\alpha_p(\mathrm{i}\omega)\alpha_q(\mathrm{i}\omega)$. Since

$\mathbf{Tr}[(\mathfrak{T}_{pq})^2] = 6/R_{pq}^6$ for any subblock $\mathfrak{T}_{pq}$, regardless of the geometry of the oscillator assembly, the second-order expansion of equation (2) reduces to the following energy expression for a collection of $N \geqslant 3$ QHOs:

$$E_{c,\text{RPA-QHO}}^{(2)} = -\frac{1}{2}\sum_{p\neq q}\frac{C_6^{pq}}{R_{pq}^6} \qquad (N \geq 3) \tag{10}$$

following the repeated use of the Casimir–Polder integral in equation (9) to determine the set of interoscillator dispersion coefficients. The reader will notice that this expression is the standard interatomic pairwise summation formula utilized by methods such as DFT-D to compute the dispersion energy corresponding to a collection of N atoms.

Although the second-order expansion of the ACFD-RPA correlation energy in equation (2) yields the familiar interatomic pairwise expression for the dispersion energy given by equation (10), the former equation is more general and provides a powerful formalism for the development of approximate methods for computing the dispersion energy in molecular systems of interest. For one, the ACFD-RPA correlation energy expression allows for the explicit utilization of the tensor form of the frequency-dependent dipole polarizability, enabling a fully anisotropic treatment of the dispersion energy. In this regard, anisotropy in the polarizability has been found to play a non-negligible role in the accurate description of intermolecular dispersion interactions [71, 72]. In the next two sections, we extend our model by considering QHOs characterized by spatially distributed (instead of point) dipole polarizabilities and describe a method for capturing the anisotropy in the frequency-dependent dipole polarizability based on the solution of the self-consistent Dyson-like screening equation of classical electrodynamics. Secondly, the use of the untruncated ACFD-RPA correlation energy expression allows for the explicit inclusion of the higher-order ($n > 2$) energetic contributions that arise naturally in the power series expansion of $\chi_0 v$. These terms include two distinct energetic contributions: the beyond-pairwise (non-additive) many-body interactions (to N^{th} order) and the higher-order electrodynamic response screening [73] (to *infinite* order). The first example of the beyond-pairwise many-body interactions is captured in the third-order expansion of the ACFD-RPA correlation energy (for a system with $N \geqslant 3$), which is the so-called Axilrod–Teller–Muto triple-dipole term [74]. The higher-order response screening is most easily illustrated by considering a system composed of two QHOs p and q and expanding equation (2):

$$E_{\text{c,RPA-QHO}} = -\frac{1}{2\pi}\int_0^\infty \mathrm{d}\omega\left(\frac{6\alpha_p(\mathrm{i}\omega)\alpha_q(\mathrm{i}\omega)}{R_{pq}^6} + \frac{9\alpha_p^2(\mathrm{i}\omega)\alpha_q^2(\mathrm{i}\omega)}{R_{pq}^{12}} + \ldots\right) (N=2) \tag{11}$$

in which the second-order term corresponds to the 'standard' C_6/R^6 pairwise dispersion interaction and the higher-order terms (which only survive with even powers of n) correspond to the electrodynamic screening of the polarizability of atom p by the presence of atom q and vice versa. In section 2.5, we extend our above treatment and describe a method that accurately and efficiently accounts for both beyond-pairwise non-additive many-body and higher-order electrodynamic response screening contributions to the dispersion energy for an arbitrary system of N QHOs.

2.3. Extension to spatially distributed dipole polarizabilities

In the previous section, we assumed that the QHOs were separated by a sufficiently large distance which allowed us to describe the interactions between them using the bare dipole–dipole interaction potential, a condition that will now be relaxed in order to consider the general case, in which QHOs can be separated by typical chemical bond distances. We remind the reader that in the limit of fluctuating *point* dipoles, the interaction between QHOs diverges when the interoscillator distances become relatively close (i.e., the so-called 'polarization catastrophe').

The most straightforward way to avoid this near-field divergence is to account for interoscillator overlap effects by extending our treatment to spatially distributed frequency-dependent polarizabilities in the model system of QHOs. In fact, this extension to spatially distributed frequency-dependent dipole polarizabilities can actually be derived from first principles by utilizing fundamental quantum mechanics, i.e., the solutions of the Schrödinger equation for the QHO. To proceed, we first note that the ground-state (singly-occupied) QHO wavefunction, $\psi_0^{\text{QHO}}(\mathbf{r})$, is a spherical Gaussian function, and hence the corresponding ground-state QHO charge density is also a spherical Gaussian function (via the Gaussian product theorem), i.e.,

$$n_0^{\text{QHO}}(\mathbf{r}) = |\psi_0^{\text{QHO}}(\mathbf{r})|^2 = \frac{\exp[-\mathbf{r}^2/2\sigma^2]}{\pi^{3/2}\sigma^3}, \tag{12}$$

in which σ represents the width or spread of the spherical Gaussian function. The corresponding Coulomb interaction between two spherical Gaussian charge distributions associated with oscillators p and q can then be derived as [75]

$$\bar{v}_{pq} = \frac{\mathrm{erf}[R_{pq}/\sigma_{pq}]}{R_{pq}}, \tag{13}$$

in which $\sigma_{pq} = \sqrt{\sigma_p^2 + \sigma_q^2}$, is an effective width obtained from the Gaussian widths of oscillators p and q (see equation (12)), that essentially determines the correlation length of this interaction potential. Since the dipole polarizability relates the response of a dipole moment to an applied electric field, the σ parameters physically correspond to the spatial spread of the local dipole moment distribution centered on a given oscillator. In fact, these Gaussian widths are directly related to the polarizability in classical electrodynamics [76] and can be derived from the dipole self-energy (i.e., the zero-distance limit of the dipole–dipole interaction potential derived below in equation (14)) as $\sigma_p = (\sqrt{2/\pi}\,\alpha_p/3)^{1/3}$.

From equation (13), we now proceed to derive the corresponding dipole–dipole interaction tensor between two

QHOs p and q from this Coulomb potential, which takes on the following form after straightforward algebra (see equations (5) and (13)):

$$\overline{\mathfrak{T}}^{ab}_{pq} = \nabla_{\mathbf{R}_p} \otimes \nabla_{\mathbf{R}_q} \overline{v}_{pq} = -\frac{3R_aR_b - R_{pq}^2\delta_{ab}}{R_{pq}^5} \times \left(\mathrm{erf}[R_{pq}/\sigma_{pq}] - \frac{2}{\sqrt{\pi}} \frac{R_{pq}}{\sigma_{pq}} \exp[-(R_{pq}/\sigma_{pq})^2] \right) + \frac{4}{\sqrt{\pi}} \frac{R_aR_b}{\sigma_{pq}^3 R_{pq}^2} \exp[-(R_{pq}/\sigma_{pq})^2]. \quad (14)$$

We note that the above expression describes a dipole–dipole interaction potential that (i) attenuates the interaction between oscillators at short distances in comparison to the bare dipole–dipole interaction potential, converging to a finite value even in the zero-distance limit, and (ii) becomes equivalent to the bare dipole–dipole interaction potential for large interoscillator distances. Hence, the use of this dipole–dipole interaction potential for an arbitrary collection of QHOs (now described by spatially distributed frequency-dependent dipole polarizabilities) not only allows us to avoid the near-field divergence that plagues the short-range, but also provides us with the simultaneous ability to correctly describe the long-range electrodynamic response as described in the next section.

2.4. Electrodynamic response screening and polarizability anisotropy

Neglecting retardation effects due to the finite speed of light, the long-range dispersion energy between two atoms *in vacuo* originates from the electrodynamic interaction of 'atomic' dipolar fluctuations. However, when an atom is embedded in a condensed phase (or in a sufficiently large molecule), the corresponding dipolar fluctuations significantly differ from the free atom case, and in fact, this difference originates from both the local chemical environment surrounding the atom (decaying exponentially via the radial extent of the atomic density) and the long-range electrodynamic interactions with the more distant fluctuating dipoles (decaying via a $\sim 1/R^3$ power law). In other words, each atom located inside a molecule or material experiences a dynamic internal electric field created by both the local and non-local fluctuations associated with the surrounding atoms. Depending on the underlying topology of the chemical environment, this fluctuating internal electric field can give rise to either polarization or depolarization effects, and is largely responsible for the anisotropy in the molecular polarizability tensor [77, 78]. Therefore, it is essential to include the environmental screening effects arising from both the short- and long-range in accurate first-principles calculations of the dispersion energy.

To proceed, we now map the atoms in a given molecular system of interest to the model system of QHOs discussed in the previous sections. In what follows, each atom p in the molecular system will be represented by a single QHO characterized by a position vector $\mathbf{R}_p = \{x_p, y_p, z_p\}$ and a corresponding frequency-dependent dipole polarizability, $\alpha_p(\mathrm{i}\omega)$.

To account for the local chemical environment, we utilize the Tkatchenko–Scheffler (TS) scheme [31] for the frequency-dependent dipole polarizability, which retains the first term in the Padé series [79] for $\alpha_p(\mathrm{i}\omega)$, i.e.,

$$\alpha_p(\mathrm{i}\omega) = \frac{\alpha_p^0}{1 + (\omega/\omega_p)^2}, \quad (15)$$

in which the static dipole polarizability, $\alpha_p^0 = \alpha_p^0[n(\mathbf{r})]$, and the effective (or characteristic) excitation frequency, $\omega_p = \omega_p[n(\mathbf{r})]$, are defined as functionals of the ground-state electron density, obtained from an initial self-consistent quantum mechanical calculation using either semi-local or hybrid DFT [80]—methods that can accurately treat electrostatics, induction, exchange-repulsion and local hybridization effects, but lack the ability to describe long-range dispersion interactions [81].

Assuming that the system (whether it be an individual molecule, a collection of molecules or even condensed matter), has a finite electronic gap and can therefore be divided into effective atomic fragments, the Hirshfeld [82], or stockholder, partitioning of the total electron density, $n(\mathbf{r})$, into atomic electron densities, $n_p(\mathbf{r})$, is then utilized to account for the local chemical environment surrounding each atom. By exploiting the proportionality between volume and polarizability [83], the Hirshfeld volume of an 'atom-in-a-molecule,' $V_p[n(\mathbf{r})]$, can then be utilized to compute $\alpha_p^0[n(\mathbf{r})]$ by appropriately scaling highly accurate reference *free* atom data, as follows:

$$\alpha_p^0[n(\mathbf{r})] = \left(\frac{V_p[n(\mathbf{r})]}{V_p^{\mathrm{free}}} \right) \alpha_p^{0,\mathrm{free}} = \left(\frac{\int \mathrm{d}\mathbf{r}\; r^3 n_p(\mathbf{r})}{\int \mathrm{d}\mathbf{r}\; r^3 n_p^{\mathrm{free}}(\mathbf{r})} \right) \alpha_p^{0,\mathrm{free}} \quad (16)$$

in which the free-atom volume, V_p^{free}, was obtained from the free-atom electron density, n_p^{free}, and $\alpha_p^{0,\mathrm{free}}$ is the reference free-atom static dipole polarizability (which can be taken from either experimental data or high-level quantum chemical calculations).

Since the free-atom parameters are referenced to highly accurate reference data, short-range quantum mechanical exchange-correlation effects are accounted for in these quantities by construction. In fact, the frequency-dependent dipole polarizabilties defined in this manner (cf equations (15) and (16)) yield C_6 coefficients that are accurate to 5.5% when compared to reference experimental values for an extensive database of atomic and (small) molecular dimers [31]. Nevertheless, this parameterization of the frequency-dependent dipole polarizability clearly lacks the aforementioned long-range electrodynamic screening that extends beyond the range of the exponentially decaying atomic densities, and these effects must be accounted for self-consistently within this system of fluctuating QHOs.

To capture the long-range electrodynamic response screening and anisotropy effects, we self-consistently solve the Dyson-like screening equation utilizing the dipole–dipole interaction tensor derived above in equation (14) which accounts for spatially-distributed frequency-dependent polarizabilities, thereby improving upon our initial description of the bare response function corresponding to this collection of QHOs. To proceed forward, we recall that this initial bare response function, χ_0, was constructed as a direct sum over

the individual oscillator response functions given in equation (4), i.e., $\chi_0 = \chi_{0,p} \oplus \chi_{0,q} \oplus \ldots$, which now correspond to QHOs characterized by isotropic frequency-dependent dipole polarizabilities parameterized using the TS definitions for $\alpha_p^0[n(\mathbf{r})]$ (presented above) and $\omega_p[n(\mathbf{r})]$. Therefore, the real-space matrix representation of χ_0 is diagonal, and as a result, the corresponding self-consistent Dyson-like screening (SCS) equation is separable and can be recast as the following non-homogeneous system of linear equations for a given frequency ω:

$$\overline{\alpha}_p(\mathrm{i}\omega) = \alpha_p(\mathrm{i}\omega) - \alpha_p(\mathrm{i}\omega) \sum_{q \neq p}^{N} \overline{\overline{\mathfrak{T}}}_{pq} \overline{\alpha}_q(\mathrm{i}\omega) \quad p = 1, 2, \ldots, N. \tag{17}$$

In equation (17), the complexity associated with integrating over spatial variables $\mathbf{r}$ and $\mathbf{r}'$ has been absorbed into $\overline{\overline{\mathfrak{T}}}_{pq}$, the 3 × 3 block of the dipole–dipole interaction tensor in equation (14), which facilitates the use of overlapping spatially distributed frequency-dependent dipole polarizabilities by eliminating the issues associated with the near-field divergence. The set of $\overline{\alpha}_p(\mathrm{i}\omega)$ are the unknowns in the SCS equations and physically correspond to frequency-dependent dipole polarizabilities that account for both short-range (via the TS scheme) and long-range (via the solution of the SCS equations) electrodynamic response screening effects arising from the chemical environment. The solution of equation (17) with input polarizabilities computed at the TS level will be referred to as TS+SCS throughout the remainder of this work.

The SCS equations can be solved in matrix form at discretized frequencies, ω (chosen from the roots of the Gauss–Legendre numerical integration scheme), via inversion of the $3N \times 3N$ Hermitian $\overline{A}$ matrix, which contains the inverse of the atomic frequency-dependent dipole polarizability tensors, $\alpha_p^{-1}(\mathrm{i}\omega)$, along the diagonal 3 × 3 atomic subblocks, and the dipole–dipole interaction tensor, $\overline{\overline{\mathfrak{T}}}_{pq}$, in each of the corresponding 3 × 3 non-diagonal subblocks. Inversion of the $\overline{A}$ matrix yields the screened (dense) non-local polarizability matrix, $\overline{B}$, or equivalently speaking, the corresponding *interacting* response function, χ_1, for our model system of QHOs. From this screened polarizability matrix, one can obtain the screened *molecular* polarizability tensor by internally contracting over all atomic subblocks,

$$\overline{\alpha} = \sum_{pq}^{N} \overline{B}_{pq}, \tag{18}$$

and the screened set of *atomic* polarizability tensors by partial internal contraction,

$$\overline{\alpha}_p = \sum_{q}^{N} \overline{B}_{pq}. \tag{19}$$

Utilizing the SCS procedure, these quantities are computed in a fully self-consistent manner and respectively provide the induced dipole moments of the molecule and individual atoms resulting from the internal electric field (i.e., the electric field due to the presence of the atoms in the molecular system of interest). In addition to these screened molecular and atomic frequency-dependent dipole polarizabilities, the set of screened characteristic frequencies, $\overline{\omega}_p$, are also obtained during solution of the SCS equations as described in reference [38].

2.5. The many-body dispersion energy: the DFT+MBD method

To compute the full many-body dispersion (MBD) energy, we directly solve the Schrödinger equation for this set of fluctuating and interacting QHOs (representing the N atoms in a given molecular system of interest) within the dipole approximation. The corresponding Hamiltonian for this model system is given by [84–90]:

$$H = -\frac{1}{2}\sum_{p=1}^{N} \nabla_{\boldsymbol{\mu}_p}^2 + \frac{1}{2}\sum_{p=1}^{N} \overline{\omega}_p^2 \mu_p^2 + \sum_{p>q}^{N} \overline{\omega}_p \overline{\omega}_q \sqrt{\overline{\alpha}_p^0 \overline{\alpha}_q^0}\, \boldsymbol{\mu}_p \overline{\overline{\mathfrak{T}}}_{pq} \boldsymbol{\mu}_q, \tag{20}$$

in which each QHO p is characterized by a screened static dipole polarizability, $\overline{\alpha}_p^0$, and screened excitation frequency, $\overline{\omega}_p$, computed at the TS+SCS level of theory (via solution of the SCS equations in equation (17)), an oscillator mass, $m_p = (\overline{\alpha}_p^0 \overline{\omega}_p^2)^{-1}$, and $\boldsymbol{\mu}_p = \sqrt{m_p}\boldsymbol{\xi}_p$, which is defined in terms of $\boldsymbol{\xi}_p$, the displacement of a given QHO from its equilibrium position. The first two terms in this Hamiltonian correspond to the single-particle kinetic and potential energy, respectively, while the last term describes the coupling between QHOs via the dipole–dipole interaction tensor ($\overline{\overline{\mathfrak{T}}}_{pq} = \nabla_{\mathbf{R}_p} \otimes \nabla_{\mathbf{R}_q} \overline{\overline{v}}_{pq}$), where $\overline{\overline{v}}_{pq}$ will be defined below).

For a model system of QHOs coupled within the dipole approximation, we have in fact proven the equivalence [91] between the full interaction energy obtained from the diagonalization of the Hamiltonian in equation (20) and the corresponding ACFD-RPA correlation energy expression in equation (2). Hence, the full ACFD-RPA correlation energy for this model system of QHOs can be obtained by taking the difference between the zero-point energies [92, 93] derived from the coupled and uncoupled QHO frequencies, i.e.,

$$E_{\mathrm{MBD}} = \frac{1}{2}\sum_{p=1}^{3N} \sqrt{\lambda_p} - \frac{3}{2}\sum_{p=1}^{N} \overline{\omega}_p = E_{\mathrm{c,RPA\text{-}QHO}} \tag{21}$$

in which λ_p are the Hamiltonian matrix eigenvalues, i.e., due to the bilinear nature of the interaction potential in equation (20), these eigenvalues simply correspond to the classical normal modes in this model system of QHOs. As such, the full ACFD-RPA correlation energy for this model system of QHOs can be efficiently computed by diagonalizing this $3N \times 3N$ Hamiltonian matrix, thereby enabling treatment of molecular systems containing thousands of atoms on a single processor—this is in stark contrast to the computationally expensive methods traditionally utilized for computing the ACFD-RPA correlation energy based on the Adler–Wiser bare response function (cf equation (3)).

Although the MBD energy is part of the long-range correlation energy, the full correlation energy in general also includes other contributions. In order to construct an electronic structure method that treats the full range of exchange and correlation effects, we need to couple the MBD energy for our model system of QHOs in equation (21) to an approximate semi-local

DFT functional. Instead of utilizing an *ad hoc* damping function, as typically employed in interatomic pairwise approaches, the coupling of MBD to an underlying functional (DFT+MBD) is achieved via the following range-separated Coulomb potential [94, 95], which suppresses the short-range interactions that are already captured at the DFT level,

$$\overline{\overline{V}}_{pq} = \left(1 - \exp(-(R_{pq}/\overline{R}_{pq}^{\mathrm{vdW}})^{\beta})\right)/R_{pq}, \qquad (22)$$

where β is a range-separation parameter that controls how quickly $\overline{\overline{V}}_{pq}$ reaches the long-range $1/R_{pq}$ asymptote, and $\overline{R}_{pq}^{\mathrm{vdW}} = \overline{R}_{p}^{\mathrm{vdW}} + \overline{R}_{q}^{\mathrm{vdW}}$ are the screened vdW radii as defined in [31, 38]. The value of this range-separation parameter, β, is obtained from global optimization of the total DFT+MBD energy on the S22 test set, a widely employed benchmark database of noncovalent intermolecular interactions [96, 97]. For the PBE [98] and PBE0 [99, 100] functionals, the optimized values of the β parameter were found as 2.56 and 2.53, respectively.

Finally, we remark that our choice of using the screened $\overline{\alpha}_p^0$ and $\overline{\omega}_p$ parameters as inputs in equation (20) is not unique, which is nothing more than a restatement of the fact that the choice of the reference bare response function in the ACFD formula for the correlation energy is not uniquely defined. Our experience with this choice for the reference bare response function is quite extensive (see sections 3 and 4 for an overview of some of the systems considered) and seems to indicate that the current choice is a good starting point for computing the long-range many-body dispersion energy in molecular systems of interest. However, there is one theoretical drawback associated with the present method that is worth mentioning: by using the screened $\overline{\alpha}_p^0$ and $\overline{\omega}_p$ parameters as inputs and diagonalizing the Hamiltonian in equation (20), the MBD method does not fully utilize the information contained in the non-local polarizability matrix obtained after the solution of equation (17). The use of this non-local information requires an interaction potential that couples the collective QHO modes and matches the underlying DFT functional. We are actively investigating alternatives which are certainly possible from the viewpoint of the ACFD formula that incorporate the improvements discussed in section 5, which we hope will allow us to treat the full spectrum of molecular systems of interest on an equal footing, ranging from small gas-phase molecules to complex materials of biological and nano-technological importance.

3. Applications: the role of self-consistent electrodynamic screening in molecular polarizabilities

The simplest possible model for the polarizability of molecules and solids consists of a sum over effective hybridized polarizable atoms, as given by equation (15). This model can be very effective in reproducing accurately known isotropic molecular polarizabilities and isotropic C_6 coefficients. For example, the TS method uses a localized atom-based model and yields an accuracy of ≈14% for the isotropic polarizabilities of more than 200 molecules [101] and 5.5% for the C_6 coefficients in 1225 cases [31]. However, one has to recognize that the

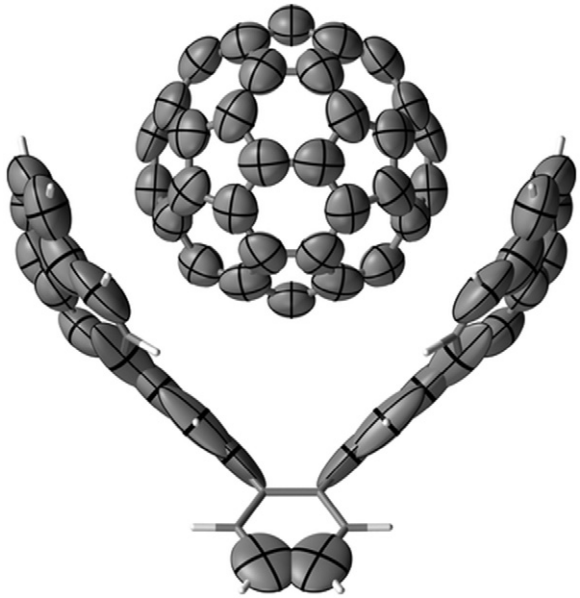

Figure 1. Illustration of the geometry and the anisotropy in the atomic TS+SCS polarizabilities of the $C_{60}@C_{60}H_{28}$ buckyball catcher complex. The atomic polarizability tensors are visualized as ellipsoids [102].

polarizability measures the response of a dipole moment to an applied electric field. Since both the dipole moment and the electric field are vector quantities, the dipole polarizability is evidently anisotropic and should be described by a second-rank tensor. Hence, the rather simplified additive model fails to correctly capture the anisotropy in the molecular polarizability [3]. Within the framework of electronic structure calculations, the static polarizability can be computed as the second derivative of the total energy with respect to an applied electric field. An alternative, but equivalent formulation for computing the polarizability is based on the fact that the single-particle orbitals in a molecule are electrodynamically coupled. The solution of the coupling equations leads to the many-electron frequency-dependent polarizability of the full system.

The TS+SCS method introduced above in equation (17) is based on such an electrodynamic interaction model. Upon obtaining effective isotropic parameters for atoms in a molecule or a solid from the ground-state electron density, the non-local polarizability tensor is determined from the solution of a system of dipole–dipole coupling equations. The dipole–dipole coupling between atoms naturally introduces anisotropy in the molecular polarizability, even if one starts with purely isotropic atomic polarizabilities (see figure 1). We now illustrate the importance of electrodynamic screening for three different cases: small and medium-sized molecules, a linear chain of H_2 molecules, and silicon clusters of increasing size.

3.1. Small and medium-sized molecules

Table 1 shows the three components of the molecular static polarizability, α_{xx}^0, α_{yy}^0, and α_{zz}^0, along with the isotropic static polarizability, α_{iso}^0, for a database of 18 molecules [78]. The calculation of the electron density was carried out using the PBE exchange–correlation functional [98], however using a

Table 1. The isotropic polarizability $\alpha^0_{\rm iso}$, along with its three components α^0_{xx}, α^0_{yy} and α^0_{zz} (in Bohr^3) for a database of molecules with experimental data taken from [78]. The z axis is taken along the principal molecular axis in the case of non-symmetric molecules. The mean absolute relative error (MARE) for the components corresponds to the error in the fractional anisotropy (see text). The results are reported for the TS method (with anisotropy computed from the Hirshfeld partitioning, where the $\boldsymbol{r}^3$ operator is partitioned as $(xx+yy+zz)$ **r**, and the TS+SCS method.

	Experiment				TS				TS+SCS			
Molecule	$\alpha^0_{\rm iso}$	α^0_{xx}	α^0_{yy}	α^0_{zz}	$\alpha^0_{\rm iso}$	α^0_{xx}	α^0_{yy}	α^0_{zz}	$\alpha^0_{\rm iso}$	α^0_{xx}	α^0_{yy}	α^0_{zz}
H_2	5.33	4.86	4.86	6.28	4.61	4.57	4.63	4.63	3.98	3.15	3.15	5.64
N_2	11.88	9.79	9.79	16.06	12.59	12.02	12.02	13.73	11.24	8.79	8.79	16.14
O_2	10.80	8.17	8.17	15.86	10.03	10.02	10.02	10.06	9.86	7.61	7.61	14.36
CO	13.16	11.00	11.00	17.55	14.62	13.80	13.80	16.27	13.21	10.76	10.76	18.13
ethane	30.23	26.86	26.86	37.05	33.72	33.18	33.18	34.79	31.86	28.78	28.79	38.02
propane[a]	43.05	38.74	38.74	51.69	49.04	47.68	48.88	50.55	46.66	39.75	42.79	57.43
cyclopentane	61.75	56.69	61.88	66.67	74.56	72.49	75.56	75.63	68.49	57.54	73.95	73.98
cyclohexane	74.23	63.30	79.70	79.70	90.59	88.36	91.70	91.70	83.27	67.91	90.96	90.96
dimethylether	35.36	29.63	33.34	43.05	39.24	38.53	39.47	39.70	37.82	32.11	32.70	48.66
P-dioxane[b]	58.04	47.24	63.43	63.43	70.50	69.68	70.17	71.64	65.76	53.12	67.20	76.97
methanol	22.40	17.88	21.80	27.60	24.44	23.99	24.61	24.72	23.11	19.96	21.44	27.92
ethanol	34.28	30.37	33.61	38.87	39.71	38.73	39.15	41.23	37.64	32.33	37.28	43.29
formaldehyde[b]	16.53	12.35	18.63	18.63	19.06	17.09	19.54	20.55	18.09	11.42	18.86	24.00
acetone[b]	43.12	29.83	49.74	49.74	49.07	45.80	50.22	51.18	48.05	35.41	49.90	58.83
acetonitrile	30.23	25.98	25.98	38.74	32.51	31.17	31.17	35.19	32.82	23.62	23.62	51.22
$(CH_3)_3CCN$	64.72	60.94	60.94	72.27	79.13	78.16	78.16	81.07	77.09	70.65	70.65	89.98
methane	17.68	17.68	17.68	17.68	18.90	18.90	18.90	18.90	17.39	17.39	17.39	17.39
benzene	69.70	45.10	82.00	82.00	75.29	71.82	77.02	77.03	71.95	33.02	91.41	91.42
MARE	—		—		13.2%		76.3%		9.1%		33.5%	

[a] The experimental values from [78] do not distinguish between α^0_{xx} and α^0_{yy} components, while first-principles calculations lead to visible differences in these components of the polarizability.

[b] The experimental values from [78] do not distinguish between α^0_{yy} and α^0_{zz} components, while first-principles calculations lead to visible differences in these components of the polarizability.

different functional would lead to a negligible change in the results [31]. The TS atomic partitioning of the polarizability integrated in different directions yields a mean absolute error of 13.2% for the isotropic molecular polarizability, and a much larger error of 76.3% for the fractional anisotropy (FA), defined as

$$\mathrm{FA} = \sqrt{\frac{1}{2}\frac{(\alpha^0_{xx}-\alpha^0_{yy})^2+(\alpha^0_{xx}-\alpha^0_{zz})^2+(\alpha^0_{yy}-\alpha^0_{zz})^2}{(\alpha^0_{xx})^2+(\alpha^0_{yy})^2+(\alpha^0_{zz})^2}}. \quad (23)$$

Upon including screening effects using the TS+SCS model (equation (17)), the isotropic polarizability is improved to 9.1%, and, more importantly, the accuracy of FA is improved by a factor of two to 33.5%. We believe that a substantial part of the remaining error stems from the isotropic input to the SCS model. Using the full electron density anisotropy at the TS level requires a substantial extension of the TS+SCS model, which is work that is currently in progress. We note that for the calculation of the vdW energy, what matters is the integration of the polarizability over imaginary frequencies, $\alpha(\mathrm{i}\omega)$, hence the error in the static polarizability is less significant when computing the vdW energy.

For a pair of atoms or molecules A and B, the C_6^{AB} coefficient determines their long-range vdW interaction energy. One of the main achievements of the TS method consists of an essentially first-principles model to determine the C_6^{AB} coefficients with an accuracy of 5.5% for a broad variety of small and medium-sized molecules (1225 C_6^{AB} coefficients). The performance of the TS method is shown in figure 2, where a remarkable correlation can be seen with reliable C_6^{AB} values computed from the experimental dipole–oscillator strength distributions (see [119] for a detailed analysis). The reason behind such good performance is that SCS effects beyond semilocal hybridization largely average out when computing C_6 coefficients for small molecules. In fact, the TS+SCS method yields an accuracy of 6.3% for the aforementioned 1225 C_6 coefficients and its performance is also shown in figure 2. We attribute the slight increase of the error with respect to TS as stemming from the approximation of the dipole moment distribution by a single isotropic QHO. The largest errors of TS+SCS are found for linear alkane chains, where the anisotropy along the chain is overestimated. Full tensor formulation of the input TS polarizabilities is under way and preliminary results indicate that the molecular anisotropy is improved.

3.2. A linear chain of H_2 molecules

We further illustrate the importance of SCS effects with the example of the linear $(H_2)_3$ chain, consisting of three H_2 dimers with alternating bond lengths (bond length of 2 Bohr inside the dimer and 3 Bohr between the dimers). An accurate calculation of the polarizability of such hydrogen dimer chains is considered to be a significant challenge for electronic structure theory [103]. We have calculated the reference frequency-dependent polarizability for $(H_2)_3$ using the linear-response coupled-cluster method (LR-CCSD) as implemented in the NWChem code [104, 105]. The LR-CCSD method is a state-of-the-art approach for computing static and frequency-dependent molecular polarizabilities, and it yields results that

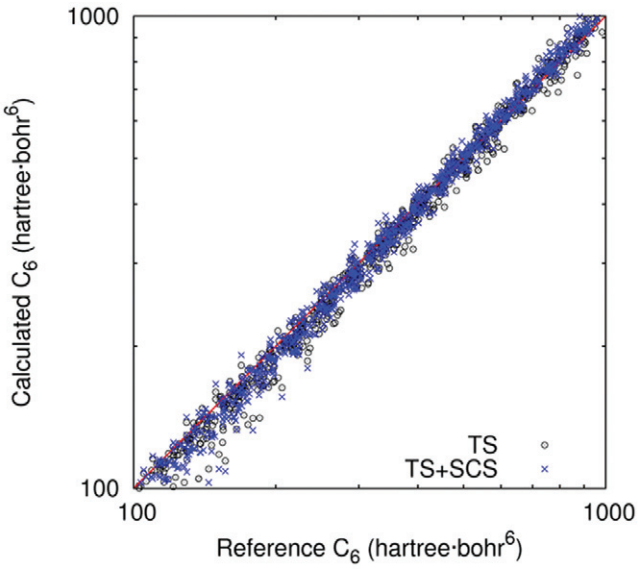

Figure 2. Isotropic C_6 coefficients for a database of 50 atoms and molecules (1225 data points) computed with the TS and TS+SCS methods, compared with reliable dipole oscillator strength distribution values obtained from experimental measurements [31].

agree to ≈3% when compared to reliable experimental values. The results for the isotropic and anisotropic C_6 coefficients for this chain at the TS, TS+SCS, and LR-CCSD levels of theory are shown in table 2. The TS method yields a vanishingly small anisotropy in the C_6 coefficients since it only accounts for the local environment. On the contrary, TS+SCS correctly captures the dipole alignment (polarization) along the $(H_2)_3$ chain, leading to a significant anisotropy that is in fair agreement with the reference LR-CCSD values. Also, the isotropic C_6 coefficient is noticeably improved when using the TS+SCS approach in comparison to TS.

3.3. Silicon clusters

We have shown that the TS+SCS method can rather effectively describe both anisotropy and polarization effects in molecules. We now illustrate that it can also treat depolarization in clusters and solids with the example of hydrogen-saturated silicon clusters of increasing size. The cluster–cluster C_6 coefficients are shown in figure 3. The reference values correspond to the TDLDA calculations of S. Botti *et al* [106]. We measured the accuracy of TDLDA using the experimentally derived C_6 coefficient for the SiH_4 molecule [107] and the $C_6^{\text{Si-Si}}$ coefficient in the silicon bulk determined from the Clausius–Mossotti equation with the experimental dielectric function. For the SiH_4 molecule, TDLDA yields a 13% overestimation and this error is further reduced to 3% for the silicon bulk. Therefore, we deem the TDLDA C_6 coefficients as good references for the larger silicon clusters. For smaller clusters, the TS values are in good agreement with experiment and TDLDA as expected. However, the error in the TS method increases progressively with the cluster size. For the largest $Si_{172}H_{120}$ cluster, the TS approach yields an overestimation of 27%. TS+SCS leads to an overall depolarization for the larger clusters, decreasing the error significantly in comparison to TDLDA. The depolarization effect is even larger for the Si bulk. The TS scheme yields an overestimation of 68% in the $C_6^{\text{Si-Si}}$ coefficient in comparison to the value derived from the experimental dielectric function, while the TS + SCS approach reduces the overestimation to just 8%.

Table 2. Anisotropic ($C_{6,\perp}$, $C_{6,\parallel}$) and isotropic ($C_{6,\text{iso}}$) C_6 coefficients for the linear $(H_2)_3$ chain using the TS and TS+SCS methods. The $C_{6,\perp}$ and $C_{6,\parallel}$ coefficients are defined with respect to the principal axis of the linear $(H_2)_3$ chain. Reference linear-response coupled-cluster (LR-CCSD) results are also shown. All values are in Hartree·Bohr^6.

	$C_{6,\perp}$	$C_{6,\parallel}$	$C_{6,\text{iso}}$
TS	166	161	165
TS+SCS	89	692	223
LR-CCSD	115	638	238

4. Applications: performance of the DFT+MBD method

Having established the accuracy of the TS+SCS method for computing the vdW coefficients for a wide variety of systems from molecules to solids, we now assess the performance of the DFT+MBD method based on the TS+SCS input (see section 2.5) for a broad variety of molecular systems. The cases studied herein include the binding energies of molecular dimers, conformational energetics of extended and globular alanine tetrapeptide, binding in the supramolecular host-guest buckyball catcher complex, as well as cohesion in molecular crystals composed of oligoacenes. The all-electron numeric atom-centered orbital code FHI-aims [108] was utilized for the DFT calculations discussed in this work.

4.1. Intermolecular interactions: the S22 and S66 databases

In order to assess the performance of the DFT+MBD method, we first study the S22 database of intermolecular interactions [96], a widely used benchmark database for which reliable binding energies have been calculated using high-level quantum chemical methods [96, 97]. In particular, we use the recent basis-set extrapolated CCSD(T) binding energies calculated by Takatani *et al* [97]. These binding energies are presumed to have an accuracy of ≈0.1 kcal mol^{-1} (1% relative error), and this level of accuracy is required for an unbiased assessment of approximate approaches for treating dispersion interactions.

Figure 4 shows the performance of the DFT+MBD method on the S22 database when used with the standard semilocal PBE [98] functional and the hybrid PBE0 [99, 100] functional which includes 25% Hartree–Fock exchange. The inclusion of the many-body vdW energy leads to a remarkable improvement in accuracy compared to the PBE+TS-vdW method [31]. The largest improvement when using the MBD energy over the pairwise TS-vdW energy is observed for the methane dimer and the parallel-displaced benzene dimer. We note that the methane dimer is bound by only 0.53 kcal mol^{-1} at the CCSD(T) level of theory, and the MBD energy reduces the

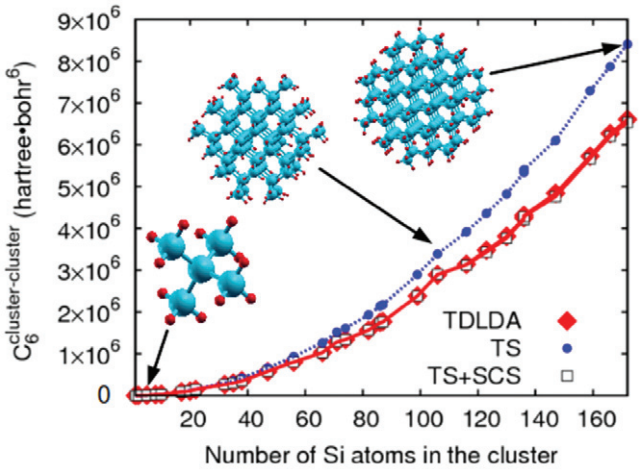

Figure 3. Cluster–cluster isotropic C_6 coefficients for hydrogen-terminated silicon clusters of increasing size. The TDLDA results are from [106].

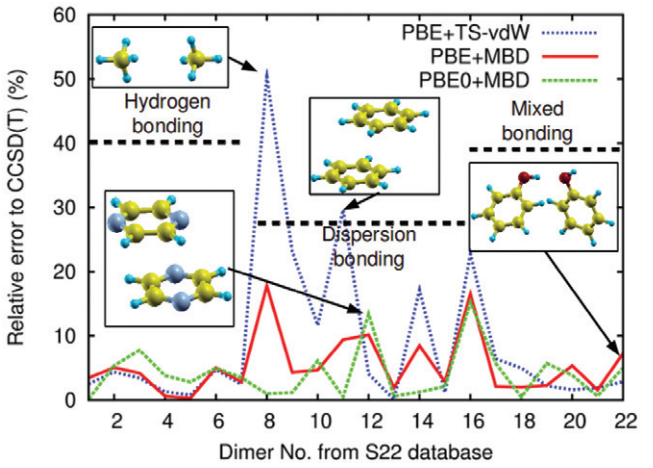

Figure 4. The performance of the PBE+TS-vdW method of Tkatchenko and Scheffler [31], PBE+MBD and PBE0+MBD on the S22 database of intermolecular interactions. The absolute relative error is reported to the basis-set converged CCSD(T) results of Takatani *et al* [97].

binding by 0.19 kcal mol^{-1} with respect to TS-vdW, explaining the large reduction in error for system 8 (methane dimer) observed in figure 4. This reduction does not come mainly from the many-body dispersion energy, rather it is due to a more physical definition of the short-range interactions in the MBD method arising from a range-separated Coulomb potential [38]. Taking the second-order expansion of the MBD energy, which yields a strictly pairwise energy, leads to a change of only 0.05 kcal mol^{-1} [91] compared to the full MBD energy. This simple test illustrates that the main difference between PBE+TS-vdW and PBE+MBD for the methane dimer stems from the different way of treating the short-range dispersion interactions. In addition, the inclusion of Hartree–Fock exchange in the PBE0 functional allows for a better description of permanent electrostatic moments and static polarizabilities for molecules, and leads to slightly improved binding energies (mean absolute relative error (MARE) of 4.1%) when compared to the semilocal PBE functional (MARE of 5.4%). We note that there are two systems in the S22 database for which the relative PBE0+MBD error exceeds 10% when compared to the CCSD(T) binding energies: pyrazine dimer (system 12) and ethene-ethyne (system 16). We attribute this finding to the remaining inaccuracy in the anisotropy for the molecular polarizabilities computed with the TS+SCS method. This issue will be analyzed in more detail for the case of the buckyball catcher complex below.

To put the performance of the DFT+MBD method in the context of other currently available approaches, we show the MARE on the S22 database for a variety of state-of-the-art methods in table 3 and in figure 5. The number of adjustable parameters employed for the dispersion energy in every method is also enumerated in table 3.

Recently, Hobza's group has significantly revised and extended the S22 database to include a broader variety of molecules and intermolecular interactions. The result of this effort is the so-called S66 database, composed of 66 molecular dimers [110]. The reference binding energies for the S66 database have been computed at the CCSD(T) level of theory employing medium-size basis sets, with an expected accuracy of ≈2–3% from the basis set limit. In order to cover non-equilibrium geometries, CCSD(T) binding energies have also been computed for 8 different intermolecular separations, ranging from a factor of 0.9 to 2.0 of the equilibrium distances. Therefore, the so-called S66x8 database contains binding energies for a total of 528 complexes computed at the CCSD(T) level of theory. The performance of the PBE0+MBD approach on the S66 database is comparable to the S22 results presented above. For equilibrium geometries in the S66 database, the mean absolute error (MAE) and MARE of the PBE0+MBD method are 0.38 kcal mol^{-1} and 6.1%, respectively. When all 528 equilibrium and non-equilibrium complexes are taken into account, the calculated MAE and MARE are 0.37 kcal mol^{-1} and 8.5%, respectively. The increase in the MARE stems from the S66(0.9x) and S66(0.95x) complexes with shorter-than-equilibrium interaction distances, where consistent underbinding can be observed. This is a well-known weakness of all dispersion-inclusive DFT methods, with errors increasing when considering shorter distances, since the dispersion energy contribution for such distances becomes very small. This observed underbinding can also be explained by the need to include higher multipoles in the vdW energy expression.

We conclude that the MBD energy beyond the TS pairwise approximation is not negligible even when studying the binding between rather small molecules. Empirical pairwise methods for the dispersion energy mimic some of the higher-order effects by adjusting sufficiently flexible damping functions, but this strategy is prone to fail for different molecular conformations and for more complex molecular geometries. We illustrate one such case in the next subsection.

4.2. Intramolecular interactions: conformational energies of alanine tetrapeptide

The study of biomolecules in the gas phase corresponds to ideal 'clean room' conditions, and recent progress in experimental gas-phase spectroscopy has yielded increasingly

Table 3. Performance of different methods on the S22 database of intermolecular interactions, measured in terms of the mean absolute relative error (MARE, in %). The errors are measured with respect to the basis-set extrapolated CCSD(T) calculations of Takatani *et al* [97]. The error is reported for hydrogen-bonded (H-B), dispersion-bonded (D-B) and mixed (M-B) systems. The number of free adjustable parameters used in every approach is shown in the 'N. param.' column.

Method	H-B	D-B	M-B	Overall	N. param.
MP2	1.8	37.4	14.8	18.9	0
EX+cRPA	11.2	21.6	14.8	16.1	0
vdW-DF1 [29]	15.2	13.0	10.8	13.0	0
PBE0-D3(Grimme) [34]	8.4	15.5	12.7	12.3	>3
EX+cRPA+SE [62]	5.9	11.6	5.4	7.8	0
vdW-DF2 [29]	5.3	6.8	10.8	7.6	1
PBE0+TS-vdW [31, 32]	3.4	12.0	6.0	7.3	2
rPW86+cPBE+VV10 [30, 109]	6.1	2.6	4.8	4.4	2
PBE0+MBD [38]	4.1	3.4	5.1	4.2	1

refined vibrational spectra for peptide secondary structures [111–113]. Joint experimental and *ab initio* theoretical studies can now successfully determine the geometries of small gas-phase peptides [114–116]. Polyalanine is a particularly good model system due to its high propensity to form helical structures [117], and its widespread use as a benchmark system for peptide stability in experiments and theory.

Here we assess the accuracy of the PBE0+MBD method for 27 conformations of alanine tetrapeptide (Ace-Ala_3-NMe, for brevity and consistency with previous biomolecular literature called Ala_4 here), for which benchmark CCSD(T) conformational energies were computed in reference [118], based on basis-set extrapolated values computed using Møller–Plesset second-order perturbation theory from references [119, 120] (the so-called MP2/CBS method). The Ala_4 conformations range from a β-sheet-like fully extended structure to a globular ('folded') conformer. The wide variety of interactions present in peptides ranging from hydrogen bonds to dispersion and electrostatics makes an accurate prediction of the conformational hierarchy of these systems quite a daunting task for affordable electronic structure calculations. We illustrate the performance of PBE0+TS-vdW and PBE0+MBD for Ala_4 conformers in figure 6. The PBE0+MBD method predicts a MAE of 0.29 kcal mol^{-1} with respect to the CCSD(T) reference, which is a significant reduction from 0.52 kcal mol^{-1} for PBE0+TS-vdW. We find that the main effect of the MBD energy over the pairwise TS-vdW approximation is to stabilize the extended conformations of Ala_4, bringing their energies in to much better agreement with the reference CCSD(T) values. Furthermore, PBE0+MBD corrects all hierarchy ordering errors present at the PBE0+TS-vdW level of theory.

4.3. Supramolecular systems: the buckyball catcher

Supramolecular host-guest systems play an important role for a wide range of applications in chemistry and biology. The prediction of the stability of host-guest complexes represents a great challenge for first-principles calculations due to the interplay of a wide variety of covalent and non-covalent

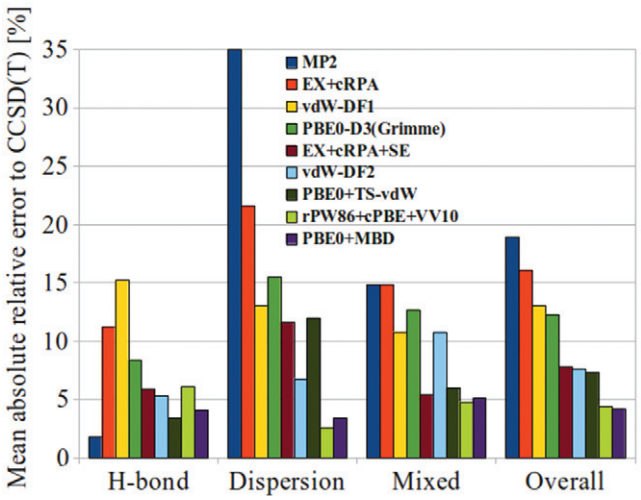

Figure 5. Performance of different methods on the S22 database of intermolecular interactions, measured in terms of the mean absolute relative error (MARE, in %). The errors are measured with respect to the basis-set extrapolated CCSD(T) calculations of Takatani *et al* [97]. Results are shown for MP2, EX+cRPA, EX+cRPA+SE [62], vdW-DF1 and vdW-DF2 [29], rPW86+cPBE+VV10 [30, 109], PBE0-D3 [34], PBE0+TS-vdW [31, 32] and PBE0+MBD [38].

interactions in these systems. Here we assess the performance of the DFT+MBD method on the binding of the so-called 'buckyball catcher' complex, $C_{60}@C_{60}H_{28}$, shown in figure 1. Since its synthesis [121], the buckyball catcher has become one of the most widely used benchmark systems for supramolecular chemistry. Recently a reliable binding energy of 26 ± 2 kcal mol^{-1} has been determined for the $C_{60}@C_{60}H_{28}$ complex from large-scale diffusion Monte Carlo (DMC) calculations [72]. This value is in excellent agreement with an extrapolated binding energy determined from the experimentally measured binding affinity [35]. It has to be mentioned that the extrapolation is done to 'remove' solvent and vibrational contributions from the measured binding affinities, therefore the extrapolation procedure contains uncontrollable approximations.

All pairwise-corrected dispersion-inclusive DFT calculations significantly overestimate the stability of the buckyball catcher complex, anywhere from 9 to 17 kcal mol^{-1} [72]. The PBE+MBD method yields a binding energy of 36 kcal mol^{-1}, improving the binding by 7 kcal mol^{-1} compared to the PBE+TS-vdW method. The inclusion of exact exchange using the PBE0+MBD method leads to a negligible change in the binding energy. Therefore, the PBE0+MBD method overestimates the binding by at least 8 kcal mol^{-1} compared to the DMC and extrapolated experimental reference binding energies.

In order to understand the most likely origin of why the binding energy of $C_{60}@C_{60}H_{28}$ complex is overestimated by PBE0+MBD, we show the projected polarizability tensors of the full complex resulting from the TS+SCS calculation in figure 1. One can clearly see that the polarizability distribution is highly anisotropic, with an increasing anisotropy close to the linker moiety that connects the two corannulene molecules of the catcher complex. While the approximation of isotropic C_6 coefficients used in DFT+MBD becomes sufficient as the distance

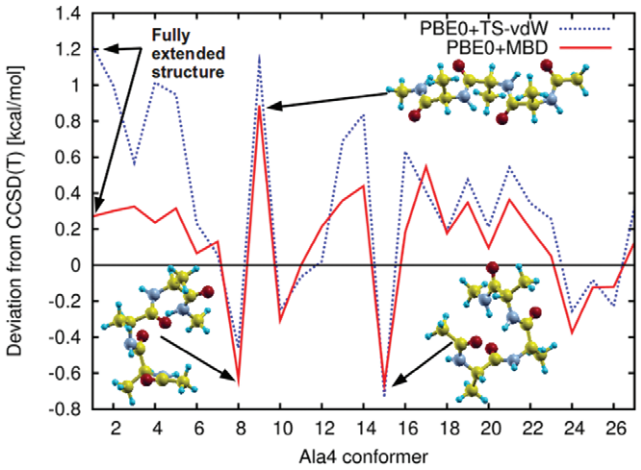

Figure 6. Performance of PBE0+TS-vdW and PBE0+MBD for the conformational energies of Ala_4. The reference CCSD(T) energies are taken from [118].

between the atoms is increased, at shorter interatomic distances the anisotropy plays a non-negligible role [71]. At present, there is no efficient method that can accurately account for the fully anisotropic dispersion energy at close interatomic distances. This statement applies to the widely employed interatomic dispersion energy methods, as well as the non-local density functionals (e.g., different variants of the vdW-DF method [17]). Work is currently in progress to seamlessly include anisotropy in dispersion energy expressions [71, 91]. In this regard, the anisotropy in the atomic polarizabilities will change the vdW energy contribution in different directions. In the case of the $C_{60}@C_{60}H_{28}$ complex, the polarizability of the $C_{60}H_{28}$ molecule is highly anisotropic as shown in figure 1. In the isotropic approximation, the dispersion energy between the C_{60} molecule and the corannulene moieties is therefore overestimated, because the polarization is artificially extended towards the C_{60} molecule. The fully anisotropic treatment of the dispersion energy is therefore likely to bring the binding energy closer to the DMC reference value.

4.4. Molecular crystals

The understanding and prediction of the structure and stability of molecular crystals is of paramount importance for a variety of applications, including pharmaceuticals, non-linear optics and hydrogen storage [122, 123]. The crystal structure prediction blind tests conducted by the Cambridge Crystallographic Data Centre have shown steady progress toward theoretical structure prediction for molecular crystals [124]. However, the insufficiency of DFT with pairwise dispersion corrections for the reliable predictions of molecular crystals is well documented, see e.g., references [33, 125–127].

To illustrate the role of MBD interactions in the stability of molecular crystals, we have studied a series of oligoacene crystals from naphthalene to pentacene. We have recently shown that reliable structures of oligoacene crystals (2% accuracy compared to low-temperature x-ray data) can be obtained with PBE+TS-vdW calculations, while MBD interactions play only a minor role in determining the geometry of these molecular

Table 4. Lattice energies of oligoacene crystals including zero-point energy corrections (PBE+TS-vdW and PBE+MBD calculations were carried out using optimized PBE+TS-vdW geometries). The range of experimental ('Exp.') 'lattice energies' were taken from [129] and extrapolated to 0 K. All values are in units of eV per molecule.

	PBE+TS-vdW	PBE+MBD	Exp.
naphthalene	−0.950	−0.862	−0.803 to −0.752
anthracene	−1.324	−1.206	−1.148 to −1.024
tetracene	−1.666	−1.587	−1.525 to −1.299
pentacene	−2.035	−2.018	−2.082 to −1.533

crystals [128]. However, the MBD energy plays a more significant role for the lattice energies of oligoacene crystals. Table 4 shows lattice energies at 0 K for naphthalene (2 benzene rings), anthracene (3 rings), tetracene (4 rings) and pentacene (5 rings) calculated using the PBE+TS-vdW and PBE+MBD methods, as well as a range of measured sublimation enthalpies extrapolated to 0 K. We have only taken those experimental values that are recommended as reliable after critical revision by the authors of reference [129], thus avoiding anomalously small or large sublimation enthalpies. Both naphthalene and anthracene crystals have been carefully studied, and their sublimation enthalpies are well known with a spread of 0.05 and 0.12 eV per molecule, respectively. There are fewer measurements available for tetracene and pentacene, and for the latter the three available experimental values deviate by 0.55 eV per molecule.

For naphthalene, anthracene and tetracene, the PBE+MBD method decreases and improves the binding by about 0.1 eV (2.3 kcal mol^{-1}) per molecule when compared to PBE+TS-vdW. This is a notable improvement, especially if viewed in the context of intermolecular interactions for the S22 and S66 databases. We remind the reader that the errors of PBE+TS-vdW and PBE+MBD for molecular dimers in the S22/S66 databases are well below 0.5 kcal mol^{-1}. The much larger difference between the pairwise PBE+TS-vdW approach and the many-body PBE+MBD method for molecular crystals can be explained by the presence of significant electrodynamic screening effects in extended systems, that are virtually absent in small molecules. We refer the reader to reference [128] for a detailed analysis of the importance of electrodynamic screening in molecular crystals.

The remaining slight overestimation of lattice energies in table 4 by PBE+MBD compared to the experimental range can be explained by the fact that the sublimation enthalpy is measured at finite temperatures, where the crystal unit cell undergoes thermal expansion. When using the experimental unit cell at 295 K for naphthalene, the PBE+MBD method yields a lattice energy that is increased by 50 meV per molecule, which places it essentially within the experimental range reported in table 4. Finally, we studied the influence of exact exchange for oligoacene crystals, finding that the PBE0+MBD method leads to an almost negligible difference when compared to PBE+MBD.

Our current work on a broad dataset of molecular crystals and their polymorphs [130, 131] shows that beyond-pairwise many-body vdW interactions can be even more

significant than found here for oligoacene crystals. In particular, DFT+MBD leads to further improvement in the computed crystal structures compared to DFT+TS-vdW and is able to account for the correct relative stabilities of polymorphic molecular crystals [131].

5. Remaining challenges

We have described a recently developed method for computing the many-body vdW dispersion energy based on a model response function corresponding to a collection of quantum harmonic oscillators (QHO). The resulting DFT+MBD approach does not contain any adjustable parameters for the determination of the frequency-dependent polarizability, and uses a single range-separation parameter for the coupling between the long-range many-body vdW energy and a given DFT functional. We view the DFT+MBD model as a crucial first step in the development of a reliable (accurate and efficient) method for describing many-body vdW interactions in complex materials.

Currently, the DFT+MBD method essentially amounts to solving the ACFD-RPA correlation energy equation for a system of localized screened QHOs in the dipole (long-range) approximation. There are several important extensions that can be accomplished within the ACFD framework that would allow us to go beyond the DFT+MBD method:

1. **Improving the anisotropic dipolar response.** The TS+SCS method defined in equation (17) yields the full non-local interacting response matrix as a function of atomic positions $\mathbf{r}$ and $\mathbf{r}'$. Currently, this information is not fully utilized in the DFT+MBD approach, since we use contracted isotropic TS+SCS atomic polarizabilities as input for the ACFD-RPA formula. In principle, the full response matrix can be used in the ACFD-RPA expression [132], however this requires a matching definition for the range-separated Coulomb potential. The interacting TS+SCS response matrix transforms the original atom-based representation to an eigenvector representation for the coupled modes of the system. The Coulomb interaction between the coupled modes needs to be extended from our current definition of range-separation that is based on atomic vdW radii.
2. **Going beyond the dipole approximation.** The QHO model possesses a response to infinite order in the multipole expansion. The current MBD method restricts the response to the dipole approximation, effectively allowing excitations only to the first excited state for every QHO due to the dipole selection rule. In principle, the full response function given by equation (3) can be computed for a system of QHOs up to an arbitrary energy cutoff for the excited states. This would allow us to treat multipole responses higher than dipole (quadrupole, octupole, etc.). The ACFD-RPA expression can still be utilized in this case, allowing us to compute dispersion interactions at shorter interatomic distances. It remains to be assessed whether or not this model will be useful, as a single QHO per atom might not be able to properly describe vdW interactions at shorter interatomic distances. However, in principle, our method can also be extended to represent every atom by several QHOs.
3. **Coupling between the long-range vdW energy and the DFT energy.** The DFT+MBD method couples the long-range vdW energy to the DFT energy by using a single range-separation parameter in the Coulomb potential. In order to improve this empirical component of the DFT+MBD method, the DFT functional has to be derived in the presence of the long-range vdW energy. To date, we have not used the fact that different functionals yield different results for the electron density tails; this information can be useful for developing a functional in which the long-range vdW energy is seamlessly integrated with the semilocal exchange-correlation functional.
4. **Simultaneous description of localized and metallic states.** Successful non-empirical DFT functionals are based on the local-density approximation (LDA) and converge to the LDA in the homogeneous electron gas (HEG) limit. LDA is an exact functional for the HEG, hence it includes vdW interactions inside the HEG. Therefore, a seamless vdW functional should yield a vanishing correction for the HEG. This can easily be accomplished by letting the polarizability vanish for slowly-varying regions of the electron density, as done in the vdW-DF [28] and VV10 [109] approaches. However, real materials (transition metals, nanostructures, etc.) are more complex than the rather simplified HEG model. In such systems, vdW interactions between ions are significant and are *screened* by the itinerant metallic electrons [133]. State-of-the-art vdW functionals do not correctly describe this complex situation. However, the DFT+MBD method can be extended to systems with localized and metallic states by introducing both localized and delocalized oscillators for every atom. The challenge consists of defining the oscillator parameters directly from the electron density and its gradient.
5. **Interatomic forces, geometry optimization and molecular dynamics.** Currently, the DFT+MBD method only yields the total energy for a specified geometry. In principle, geometry optimizations are possible by using the finite difference approximation for the interatomic forces. This is, however, computationally expensive especially in the case of molecular dynamics. Work is in progress to derive an analytic expression for the interatomic forces corresponding to the MBD energy [134]. Such development would allow for the routine application of the DFT+MBD method in large-scale molecular dynamics simulations.

6. Conclusions

There is mounting evidence that many-body vdW interactions, beyond the standard pairwise approximation, play a crucial role in the structure, stability and function of a wide variety of systems of importance in biology, chemistry and physics. We have illustrated the importance of including many-body vdW

interactions when describing small molecular dimers, conformational energies of peptides, binding in supramolecular systems, and cohesion in molecular crystals. We have presented a derivation of both the pairwise and many-body interatomic vdW dispersion energy for a model system of QHOs from the quantum-mechanical ACFD-RPA correlation energy expression. The ACFD formula provides us with a powerful framework for the understanding and future development of accurate and efficient electronic structure approaches.

The DFT+MBD method [38, 39] represents a first step towards the development of reliable methods for describing many-body vdW interactions in complex materials. In this work, we derived the MBD energy expression, discussed the approximations involved, and identified the remaining challenges that need to be addressed in future work. Over the next few years, we anticipate extensive development of new dispersion energy methods that will address the truly collective many-body nature of these ubiquitous quantum-mechanical forces.

Acknowledgments

The authors acknowledge the European Research Council (ERC starting grant VDW-CMAT) for support, and thank M Scheffler, R Car, X Ren, J F Dobson, O A von Lilienfeld, A Ambrosetti, A M Reilly and N Marom for enlightening discussions. RAD was supported by the DOE under grant no. DE-SC0005180 and by the NSF under grant no. CHE-0956500.

References

[1] Langbein D 1974 *Theory of van der Waals Attraction* (*Springer Tracts in Modern Physics* vol 72) (Berlin: Springer)
[2] Parsegian V A 2005 *van der Waals Forces: A Handbook for Biologists, Chemists, Engineers and Physicists* (Cambridge: Cambridge University Press)
[3] Stone A J 1996 *The Theory of Intermolecular Forces* (Oxford: Oxford University Press)
[4] Kaplan I G 2006 *Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials* (New York: Wiley)
[5] French R *et al* 2010 *Rev. Mod. Phys.* **82** 1887
[6] Jeziorski B, Moszynski R and Szalewicz K 1994 *Chem. Rev.* **94** 1887
[7] Beeby J L 1971 *J. Phys. C: Solid State Phys.* **4** L359
[8] Loskill P, Hähl H, Faidt T, Grandthyll S, Müller F and Jacobs K 2012 *Adv. Colloid. Interface Sci.* **179–82** 107
[9] Loskill P, Puthoff J, Wilkinson M, Mecke K, Jacobs K and Autumn K 2012 *J. R. Soc. Interface* **10** 20120587
[10] Dalvit D A R, Milonni P W, Roberts D C and Rosa F S S 2011 *Casimir Physics* (*Lecture Notes in Physics*) (Berlin: Springer)
[11] Buhl W 1976 *Z. Phys.* B **23** 221
[12] Linder B 1960 *J. Chem. Phys.* **33** 668
[13] Linder B 1964 *J. Chem. Phys.* **40** 2003
[14] Riley K E, Pitoňák M, Jurečka P and Hobza P 2010 *Chem. Rev.* **110** 5023
[15] Kannemann F O and Becke A D 2010 *J. Chem. Theory Comput.* **6** 1081
[16] Grimme S 2011 *Comput. Mol. Sci.* **1** 211
[17] Cooper V R, Kong L and Langreth D C 2010 *Phys. Proc.* **3** 1417
[18] Steinmann S and Corminboeuf C 2011 *J. Chem. Theory Comput.* **7** 3567
[19] Tkatchenko A, Romaner L, Hofmann O T, Zojer E, Ambrosch C and Scheffler M 2010 *Mater. Res. Soc. Bull.* **35** 435
[20] Misquitta A J, Podeszwa R, Jeziorski B and Szalewicz K 2005 *J. Chem. Phys.* **123** 214103
[21] von Lilienfeld O A, Tavernelli I, Rothlisberger U and Sebastiani D 2004 *Phys. Rev. Lett.* **93** 153004
[22] Johnson E R and Becke A D 2005 *J. Chem. Phys.* **123** 024101
[23] Silvestrelli P L 2008 *Phys. Rev. Lett.* **100** 053002
[24] Mercurio G *et al* 2010 *Phys. Rev. Lett.* **104** 036102
[25] Atodiresei N, Caciuc V, Lazic P and Blügel S 2009 *Phys. Rev. Lett.* **102** 136809
[26] Tonigold K and Gross A 2010 *J. Chem. Phys.* **132** 224701
[27] Stradi D *et al* 2011 *Phys. Rev. Lett.* **106** 186102
[28] Dion M, Rydberg H, Schroder E, Langreth D C and Lundqvist B I 2004 *Phys. Rev. Lett.* **92** 246401
[29] Lee K, Murray E D, Kong L, Lundqvist B I and Langreth D C 2010 *Phys. Rev.* B **92** 081101
[30] Vydrov O A and Van Voorhis T 2012 *J. Chem. Theory Comput.* **8** 1929
[31] Tkatchenko A and Scheffler M 2009 *Phys. Rev. Lett.* **102** 073005
[32] Marom N, Tkatchenko A, Rossi M, Gobre V V, Hod O, Scheffler M and Kronik L 2011 *J. Chem. Theory Comput.* **7** 3944
[33] von Lilienfeld O A and Tkatchenko A 2010 *J. Chem. Phys.* **132** 234109
[34] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *J. Chem. Phys.* **132** 154104
[35] Grimme S 2012 *Chem. Eur. J.* **18** 9955
[36] Zhang G-X, Tkatchenko A, Paier J, Appel H and Scheffler M 2011 *Phys. Rev. Lett.* **107** 245501
[37] Ruiz V G, Liu W, Zojer E, Scheffler M and Tkatchenko A 2012 *Phys. Rev. Lett.* **108** 146103
[38] Tkatchenko A, DiStasio R A Jr, Car R and Scheffler M 2012 *Phys. Rev. Lett.* **108** 236402
[39] DiStasio R A Jr, von Lilienfeld O A and Tkatchenko A 2012 *Proc. Natl Acad. Sci. USA* **109** 14791
[40] Dobson J F and Gould T 2011 Calculation of dispersion energies http://www.psi-k.org/newsletters/News_107/Highlight_107.pdf
[41] Dobson J F and Gould T 2012 *J. Phys.: Condens. Matter* **24** 073201
[42] Klimeš J and Michaelides A 2012 *J. Chem. Phys.* **137** 120901
[43] *Towards first-principles description of van der Waals interactions in complex materials (Lausanne, 2012)* http://www.cecam.org/workshop-791.html
[44] Gunnarsson O and Lundqvist B I 1976 *Phys. Rev.* B **13** 4274
[45] Langreth D C and Perdew J P 1977 *Phys. Rev.* B **15** 2884
[46] Bohm D and Pines D 1953 *Phys. Rev.* **92** 609
[47] Gell-Mann M and Brueckner K A 1957 *Phys. Rev.* **106** 364
[48] Fuchs M and Gonze X 2002 *Phys. Rev.* B **65** 235109
[49] Furche F and Van Voorhis T 2005 *J. Chem. Phys.* **122** 164106
[50] Furche F 2008 *J. Chem. Phys.* **129** 114105
[51] Scuseria G E, Henderson T M and Sorensen D C 2008 *J. Chem. Phys.* **129** 231101
[52] Janesko B G, Henderson T M and Scuseria G E 2009 *J. Chem. Phys.* **130** 081105
[53] Toulouse J, Gerber I C, Jansen G, Savin A and Ángyán J G 2009 *Phys. Rev. Lett.* **102** 096404
[54] Harl J and Kresse G 2008 *Phys. Rev.* B **77** 045136
[55] Harl J and Kresse G 2009 *Phys. Rev. Lett.* **103** 056401
[56] Lu D, Li Y, Rocca D and Galli G 2009 *Phys. Rev. Lett.* **102** 206411
[57] Dobson J F and Wang J 1999 *Phys. Rev. Lett.* **82** 2123
[58] Rohlfing M and Bredow T 2008 *Phys. Rev. Lett.* **101** 266106

[59] Ren X, Rinke P and Scheffler M 2009 *Phys. Rev.* B **80** 045402
[60] Schimka L, Harl J, Stroppa A, Grüneis A, Marsman M, Mittendorfer F and Kresse G 2010 *Nature Mater.* **9** 741
[61] Li Y, Lu D, Nguyen H-V and Galli G 2010 *J. Phys. Chem.* A **114** 1944
[62] Ren X, Tkatchenko A, Rinke P and Scheffler M 2011 *Phys. Rev. Lett.* **106** 153003
[63] Ren X, Rinke P, Joas C and Scheffler M 2012 *J Mater. Sci.* **47** 7447
[64] Dobson J F 2006 *Time-Dependent Density Functional Theory* (*Lecture Notes in Physics* vol 706) M A L Marques *et al* (Berlin: Springer) p 443
[65] Eshuis H, Bates J E and Furche F 2012 *Theor. Chem. Accounts* **131** 1084
[66] Lu D, Nguyen H-V and Galli G 2010 *J. Chem. Phys.* **133** 154110
[67] Ren X, Rinke P, Blum V, Wieferink J, Tkatchenko A, Sanfilippo A, Reuter K and Scheffler M 2012 *New J. Phys* **14** 053020
[68] Adler S L 1962 *Phys. Rev.* **126** 413
[69] Wiser N 1963 *Phys. Rev.* **129** 62
[70] Dobson J F 1994 *Topics in Condensed Matter Physicsed* M P Das (New York: Nova) p 121
[71] Krishtal A, Vannomeslaeghe K, Geldof D, Van Alsenoy C and Geerlings P 2011 *Phys. Rev.* A **83** 024501
[72] Tkatchenko A, Alfè D and Kim K S 2012 *J. Chem. Theory Comput.* **8** 4317
[73] MacRury T B and Linder B 1973 *J. Chem. Phys.* **58** 5388
[74] Axilrod B M and Teller E 1943 *J. Chem. Phys.* **11** 299
[75] Helgaker T, Jørgensen P and Olsen J 2000 *Molecular Electronic Structure Theory* (Chichester: Wiley)
[76] Mayer A 2007 *Phys. Rev.* B **75** 045407
[77] Oxtoby D W and Gelbart W M 1975 *Mol. Phys.* **29** 1569
[78] Thole B T 1981 *Chem. Phys.* **59** 341
[79] Tang K T and Karplus M 1968 *Phys. Rev.* **171** 70
[80] Parr R G and Yang W 1989 *Density Functional Theory of Atoms and Molecules* (Oxford: Oxford Science)
[81] Koch W and Holthausen M C 2002 *A Chemist's Guide to Density Functional Theory* (New York: Wiley)
[82] Hirshfeld F L 1977 *Theor. Chim. Acta* **44** 129
[83] Brink T, Murray J S and Politzer P 1985 *J. Chem. Phys.* **98** 4305
[84] Bade W L 1957 *J. Chem. Phys.* **27** 1280
[85] Cao J and Berne B J 1992 *J. Chem. Phys.* **97** 8628
[86] Donchev A G 2006 *J. Chem. Phys.* **125** 074713
[87] Cole M W, Velegol D, Kim H-Y and Lucas A A 2009 *Mol. Simul.* **35** 849
[88] Berthoumieux H and Maggs A C 2010 *Europhys. Lett.* **91** 56006
[89] Purcell E M and Pennypacker C R 1973 *Astrophys. J.* **186** 705
[90] Lakhtakia A 1990 *Opt. Commun.* **79** 1
[91] Tkatchenko A, Ambrosetti A and DiStasio R A Jr 2013 *J. Chem. Phys.* **138** 074106
[92] Higgs P W 1953 *J. Chem. Phys.* **21** 1300
[93] Andricioaei I and Karplus M 2001 *J. Chem. Phys.* **115** 6289
[94] Whitfield T W and Martyna G J 2006 *Chem. Phys. Lett.* **424** 409
[95] Jones A, Thompson A, Crain J, Müser M H and Martyna G J 2009 *Phys. Rev.* B **79** 144119
[96] Jurečka P, Šponer J, Černý J and Hobza P 2006 *Phys. Chem. Chem. Phys.* **8** 1985
[97] Takatani T, Hohenstein E G, Malagoli M, Marshall M S and Sherrill C D 2010 *J. Chem. Phys.* **132** 144104
[98] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[99] Perdew J P, Ernzerhof M and Burke K 1996 *J. Chem. Phys.* **105** 9982
[100] Adamo C and Barone V 1999 *J. Chem. Phys.* **110** 6158
[101] Gobre V V and Tkatchenko A 2013 *Nat. Commun.* **4** 2341
[102] Trueblood K N, Bürgi H-B, Burzlaff H, Dunitz J D, Gramaccioli C M, Schulz H H, Shmueli U and Abrahams S C 1996 *Acta Cryst.* A **52** 770
[103] Liu R-F, Ángyán J G and Dobson J F 2011 *J. Chem. Phys.* **134** 114106
[104] Hammond J R, Valiev M, de Jong W A and Kowalski K 2007 *J. Phys. Chem. A* **111** 5492
[105] Hammond J R, Kowalski K and de Jong W A 2007 *J. Chem. Phys.* **127** 144105
[106] Botti S, Castro A, Andrade X, Rubio A and Marques M A L 2008 *Phys. Rev. B* **78** 035333
[107] Kumar A, Kumar M and Meath W J 2003 *Chem. Phys.* **286** 227
[108] Blum V, Gehrke R, Hanke F, Havu P, Havu V, Ren X, Reuter K and Scheffler M 2009 *Comput. Phys. Commun.* **180** 2175
[109] Vydrov O A and Van Voorhis T 2010 *J. Chem. Phys.* **133** 244103
[110] Řezáč J, Riley K E and Hobza P 2011 *J. Chem. Theory Comput.* **7** 2427
[111] Chin W, Piuzzi F, Dimicoli I and Mons M 2006 *Phys. Chem. Chem. Phys.* **8** 1033
[112] Simons J P 2009 *Mol. Phys.* **107** 2435
[113] Bierau F, Kupser P, Meijer G and von Helden G 2010 *Phys. Rev. Lett.* **105** 133402
[114] Gaigeot M P 2010 *Phys. Chem. Chem. Phys.* **12** 3336
[115] Stearns J A, Seaiby C, Boyarkin O V and Rizzo T R 2009 *Phys. Chem. Chem. Phys.* **11** 125
[116] Rossi M, Blum V, Kupser P, von Helden G, Bierau F, Pagel K, Meijer G and Scheffler M 2010 *J. Phys. Chem. Lett.* **1** 3465
[117] Baldwin R L 2007 *J. Mol. Biol.* **371** 283
[118] Tkatchenko A, Rossi M, Blum V, Ireta J and Scheffler M 2011 *Phys. Rev. Lett.* **106** 118102
[119] DiStasio R A Jr , Jung Y and Head-Gordon M 2005 *J. Chem. Theory Comput.* **1** 862
[120] DiStasio R A Jr , Steele R P, Rhee Y M, Shao Y and Head-Gordon M 2007 *J. Comput. Chem.* **28** 839
[121] Sygula A, Fronczek F R, Sygula R, Rabideau P W and Olmstead M M 2007 *J. Am. Chem. Soc.* **129** 3842
[122] Bernstein J 2002 *Polymorphism in Molecular Crystals* (New York: Oxford University Press)
[123] Price S L 2008 *Phys. Chem. Chem. Phys.* **10** 1996
[124] Bardwell D A *et al* 2011 *Acta. Cryst.* B **67** 535
[125] Beran G J O and Nanda K 2010 *J. Phys. Chem. Lett.* **1** 3480
[126] Wen S, Nanda K, Huang Y and Beran G 2012 *Phys. Chem. Chem. Phys.* **14** 7578
[127] Hongo K, Watson M A, Sanchez-Carrera R S, Iitaka T and Aspuru-Guzik A 2010 *J. Phys. Chem. Lett.* **1** 1789
[128] Schatschneider B, Liang J-J, Reilly A M, Marom N, Zhang G-X and Tkatchenko A 2013 *Phys. Rev.* B **87** 060104
[129] Roux M V, Temprado M, Chickos J S and Nagano Y 2008 *J. Phys. Chem. Ref. Data* **37** 1855
[130] Reilly A M and Tkatchenko A 2013 *J. Phys. Chem. Lett.* **4** 1028
[131] Marom N, DiStasio R A Jr , Atalla V, Levchenko S, Chelikowsky J R, Leiserowitz L and Tkatchenko A 2013 *Angew. Chem. Int. Edn* **52** 6629
[132] Davies B 1972 *Chem. Phys. Lett.* **16** 388
[133] Rehr J J, Zaremba E and Kohn W 1975 *Phys. Rev.* B **12** 2062
[134] Ambrosetti A, Reilly A M, DiStasio R A Jr and Tkatchenko A 2014 *J. Chem. Phys.* **140** 18A508