Non-Covalent Interactions in Molecular Crystals with Local Orbitals Using the Exchange-Hole Dipole Moment Model

Luc M. LeBlanc,[*,†] Joseph A. Weatherby,[†] Alberto Otero-de-la-Roza,[*,‡] and Erin R. Johnson[*,†]

†*Department of Chemistry, Dalhousie University, 6274 Coburg Rd, Halifax, Nova Scotia, B3H 4R2, Canada*

‡*Faculty of Chemistry, Universidad de Oviedo, 8 Avenida Julián Claveria, 33006, Oviedo, Asturias, Spain*

E-mail: luc.leblanc@dal.ca; aoterodelaroza@fluor.quimica.uniovi.es; erin.johnson@dal.ca

1   Abstract

We present the first implementation of the exchange-hole dipole moment (XDM) model in combination with a numerical finite-support local orbital method (the SIESTA method) for the modeling of non-covalent interactions in periodic solids. The XDM model is parameterized for both the B86bPBE and PBE functionals, using double-ζ- and triple-ζ-quality basis sets (DZP and TZP). The use of finite-support local orbitals is shown to have minimal impact on the computed dispersion coefficients for van-der-Waals-bound molecular dimers and small molecular solids. However, the quality of the basis set affects quite significantly the accuracy of calculated dimer binding energies and molecular crystal lattice energies, and the size of the counterpoise correction indicates that this is caused by basis-set incompleteness error. In

the case of the DZP basis set, its performance for weakly bound gas-phase dimers is similar to a double-ζ Gaussian basis set without diffuse functions. The new XDM implementation was tested on graphite and the X23 benchmark set of molecular crystal lattice energies. Our results indicate that lattice energies similar to plane-wave calculations can be obtained only if the counterpoise correction is applied. Alternatively, the calculated equilibrium geometries are reasonably close to the plane-wave equivalents, and composite approaches in which a single-point plane-wave calculation is used at the XDM/DZP equilibrium geometry yield good accuracy at a significantly lower computational cost.

2   Introduction

The treatment of London dispersion forces[1] is crucial for the description of intermolecular interactions in materials and solids. Weak dispersion forces play an important role in determining the structural and physical properties of systems such as 2D materials[2-4] and molecular crystals.[5-8] There have been longstanding efforts to develop methods capable of modeling dispersion interactions accurately, especially in the field of density-functional theory (DFT).[9-12] While satisfactory accuracy can often be obtained with several dispersion-corrected DFT methods, it is interesting to develop computationally inexpensive variants of these methods for purposes such as ab initio molecular dynamics simulations or crystal structure prediction.[13-18]

One way of reducing the computational cost of a dispersion-corrected DFT method is by representing the Kohn-Sham orbitals in a basis set that ensures asymptotic linear scaling with system size as in, for example, the SIESTA method.[19,20] In SIESTA, linear scaling is achieved through a combination of $\mathcal{O}(N)$ algorithms and basis sets composed of finitely supported atom-centered numerical orbitals.[21-23] The SIESTA method is implemented in the program of the same name,[19,20] whose fourth major version release allows the treatment of London dispersion interactions. Two dispersion methods are currently implemented in SIESTA:

Grimme's semiempirical dispersion model (DFT-D2),[24,25] and the non-local van-der-Waals density functionals developed by Langreth and co-workers (vdW-DF1[26] and vdW-DF2[27]). DFT-D2 uses an asymptotic energy expression with fixed empirical dispersion coefficients and it is cheap and relatively accurate, but the coefficients are empirical and do not change with the chemical environment, which is essential in certain systems.[28,29] The vdW-DF methods incorporate dispersion effects using a non-local correlation energy functional. Non-local vdW-DF functionals are non-empirical and "seamless", but their use increases the computational cost significantly.[9]

An alternative approach to include dispersion effects is the exchange-hole dipole moment (XDM) dispersion model.[10,30] In XDM, the interatomic dispersion coefficients are calculated from first principles using the self-consistent density and kinetic-energy density. This makes the dispersion coefficients sensitive to the chemical environment and non-empirical, while retaining the computational and conceptual simplicity of an asymptotic pairwise dispersion expansion. The XDM model has previously been implemented in the Quantum ESPRESSO (QE) solid-state code, which makes use of plane-wave basis sets,[28,31] and it has demonstrated excellent accuracy in the treatment of non-covalent interactions in the gas phase,[32,33] as well as for surfaces,[29] layered materials,[34,35] and molecular crystals.[6,7,28,36–39]

In this work, we present the first implementation of the XDM dispersion model in combination with a finite-support local-orbital method for periodic solids, as implemented in SIESTA. The resulting XDM implementation is then parameterized on the Kannemann-Becke set of molecular dimers,[40] and tested on solid-state systems such as graphite and the X23 set of molecular solids. The implications of using local basis sets in SIESTA as opposed to delocalized plane-waves for the treatment of non-covalent interactions are discussed.

3 The XDM dispersion model: implementation, parametrization, and testing

3.1 Theory and implementation details

The XDM dispersion model supplements the total DFT energy with a dispersion energy term:

$$E = E_{\text{DFT}} + E_{\text{disp}} = E_{\text{DFT}} - \frac{1}{2} \sum_{\mathbf{L}} \sum_{n} \sum_{i \neq j} \frac{C_{ij}^n f(R_{ij\mathbf{L}})}{R_{ij\mathbf{L}}^n}, \tag{1}$$

where C_{ij}^n are the n-th order interatomic dispersion coefficients, $f(R_{ij\mathbf{L}})$ is a damping function that deactivates the dispersion interaction at short range, and

$$R_{ij\mathbf{L}} = |\mathbf{R_i} - \mathbf{R_j} + \mathbf{L}| \tag{2}$$

is the distance between atom i and j in cells separated by lattice vector $\mathbf{L}$. In practice, the sum over lattice vectors is truncated at a point such that all remaining interatomic contributions to the total dispersion energy fall below a specified energy threshold. The Becke-Johnson damping function [10,41] has the form

$$f(R_{ij\mathbf{L}}) = \frac{1}{R_{\text{vdW},ij}^n + R_{ij\mathbf{L}}^n}, \tag{3}$$

where

$$R_{\text{vdW},ij} = a_1 R_{c,ij} + a_2, \tag{4}$$

$$R_{c,ij} = \frac{1}{3} \left[\left(\frac{C_{8,ij}}{C_{6,ij}} \right)^{1/2} + \left(\frac{C_{10,ij}}{C_{6,ij}} \right)^{1/4} + \left(\frac{C_{10,ij}}{C_{8,ij}} \right)^{1/2} \right]. \tag{5}$$

The sum of van der Waals radii ($R_{\text{vdW},ij}$) is constructed from a critical radius, $R_{c,ij}$, corresponding to the point where dispersion contributions from the first three leading-order pairwise dispersion coefficients, C_{ij}^6, C_{ij}^8, and C_{ij}^{10}, are equal. The a_1 and a_2 parameters are

4

found by minimizing the residual errors between computed and reference binding energies for a benchmark set of non-covalently bound dimers (the Kannemann-Becke set[40]). These parameters depend on the functional and they serve to match the long-range dispersion and short-range exchange-correlation contributions.

The XDM dispersion coefficients are determined using second-order perturbation theory.[10,30,42] Dispersion forces are calculated as interactions between instantaneous atomic multipole moments, which originate from the distribution of electron plus exchange hole dipoles.[30] The first three leading-order pairwise dispersion coefficients are

$$C_6 = \frac{\alpha_i \alpha_j \langle M_1^2 \rangle_i \langle M_1^2 \rangle_j}{\langle M_1^2 \rangle_i \alpha_j \langle M_1^2 \rangle_j \alpha_i}, \tag{6}$$

$$C_8 = \frac{3}{2} \frac{\alpha_i \alpha_j \left(\langle M_1^2 \rangle_i \langle M_2^2 \rangle_j + \langle M_2^2 \rangle_i \langle M_1^2 \rangle_j \right)}{\langle M_1^2 \rangle_i \alpha_j \langle M_1^2 \rangle_j \alpha_i}, \tag{7}$$

$$C_{10} = 2 \frac{\alpha_i \alpha_j \left(\langle M_1^2 \rangle_i \langle M_3^2 \rangle_j + \langle M_3^2 \rangle_i \langle M_1^2 \rangle_j \right)}{\langle M_1^2 \rangle_i \alpha_j \langle M_1^2 \rangle_j \alpha_i} + \frac{21}{5} \frac{\alpha_i \alpha_j \langle M_2^2 \rangle_i \langle M_2^2 \rangle_j}{\langle M_1^2 \rangle_i \alpha_j \langle M_1^2 \rangle_j \alpha_i}. \tag{8}$$

Combined, the three dispersion energy terms corresponding to the these coefficients have been shown to describe long-range interactions in solids accurately.[10,43] Thus, in the canonical XDM implementation, the summation in Eq. 1 over n is truncated at the $n = 10$ term, and only takes into consideration atomic-pairwise contributions (i.e., $n = 6, 8, 10$). Nevertheless, it is important to note that the XDM dispersion model does take into account electronic many-body effects to all orders by way of the construction of the dispersion coefficients from the exchange hole, which is evaluated using the fully-interacting electron density.

In order to compute the XDM dispersion coefficients (Eqs. 6-8), the l-th order exchange-hole multipole moments, M_l ($l = 1, 2, 3, ...$), and the atomic polarizabilities, α_i, are needed:

$$\langle M_l^2 \rangle_i = \sum_\sigma \int \omega_i(\mathbf{r}) \rho_\sigma^{\mathrm{ae}}(\mathbf{r}) \left[r_i^l - (r_i - d_{X\sigma}(\mathbf{r}))^l \right]^2 \mathrm{d}\mathbf{r}, \tag{9}$$

$$\alpha_i = \frac{\int r^3 \omega_i(\mathbf{r}) \rho_\sigma^{\mathrm{ae}}(\mathbf{r}) \mathrm{d}\mathbf{r}}{\int r^3 \rho_i^{\mathrm{at}}(\mathbf{r}) \mathrm{d}\mathbf{r}} \alpha_i^{\mathrm{at}}. \tag{10}$$

Here, $\rho_\sigma^{\mathrm{ae}}$ is the all-electron spin-density and r is the distance to atom i. The $\rho_{i,\mathrm{at}}$ and α_i^{at} are the reference free-atom densities and polarizabilities and ω_i is the weight of that atom's contribution to the spin-density. The weights can, in principle, be constructed using any partitioning method. In XDM, we use the Hirshfeld partitioning scheme:[44,45]

$$\omega_i(\mathbf{r}) = \frac{\rho_i^{\mathrm{at}}(\mathbf{r})}{\sum\limits_j \rho_j^{\mathrm{at}}(\mathbf{r})}, \tag{11}$$

which is relatively simple to implement. The atomic polarizabilities (Eq. 10) are calculated from their *in vacuo* (free) counterparts (α_i^{at}) by exploiting their proportionality with the atomic volumes.[46] The in-the-solid and free atomic volumes are the numerator and denominator in Eq. 10, respectively. The all-electron spin-density, $\rho_\sigma^{\mathrm{ae}}$, is approximated from the valence spin-density by adding the core electron density.

In Eq. 9, $d_{X\sigma}$ is the dipole moment between the electron at the reference point $\mathbf{r}$ and its associated exchange-hole ($h_{X\sigma}$) given by

$$d_{X\sigma}(\mathbf{r}) = \int \mathbf{r}' h_{X\sigma}(\mathbf{r}, \mathbf{r}') \mathrm{d}\mathbf{r}' - \mathbf{r}. \tag{12}$$

The exact expression for $h_{X\sigma}$ is computationally prohibitive in solids because it involves a double sum over occupied states. Instead, we use the Becke-Roussel (BR) semi-local model for the spherically averaged exchange hole.[47] The BR hole provides a better approximation to the full exchange-correlation hole than $h_{X\sigma}$ and, consequently, it results in improved performance of the resulting dispersion-corrected energies.[10,48]

The BR model hole has the form of an off-centered exponential function ($-Ae^{-ar}$) displaced from the electron's reference point by a distance b.[28] The three parameters (A, a, b) are determined by enforcing three exact constraints related to the hole normalisation and its

value and curvature at the reference point. This leads to $b = d_{X\sigma}$, where b is calculated as

$$b^3 = \frac{x^3 e^{-x}}{8\pi \rho_\sigma^{\text{val}}}, \tag{13}$$

where $x = ab$ is the solution to the non-linear equation

$$\frac{xe^{-2x/3}}{x-2} = \frac{2}{3}\pi^{2/3}\frac{\rho_\sigma^{\text{val}\,5/3}}{Q_\sigma}, \tag{14}$$

which is solved numerically using Newton's method. The hole curvature (Q_σ) is

$$Q_\sigma = \frac{1}{6}(\nabla^2 \rho_\sigma^{\text{val}} - 2D_\sigma), \tag{15}$$

with

$$D_\sigma = \tau_\sigma - \frac{1}{4}\frac{|\nabla \rho_\sigma^{\text{val}}|^2}{\rho_\sigma^{\text{val}}}, \tag{16}$$

and where τ_σ is the positive-definite valence spin kinetic-energy density,

$$\tau_\sigma = \sum_i |\nabla \psi_{i\sigma}|^2. \tag{17}$$

It is important to make sure that the calculated value of b does not unphysically overshoot the distance to the closest nucleus.[28,30] For this reason, $d_{X\sigma}$ is set to $\min(b, r_i)$ instead of simply b in Eq. 9.

The XDM equations (Eq. 1 to 17) are straightforward to implement in SIESTA. In contrast to the plane-wave (QE) implementation,[31] the electron density and kinetic energy density are written as a sum over local orbitals:

$$\rho_\sigma = \sum_i |\psi_{i,\sigma}|^2 = \sum_i \left|\sum_{\mu,\nu} c_{\mu,\nu}\phi_\mu\phi_\nu\right| \tag{18}$$

and

$$\tau_\sigma = \sum_i |\nabla \psi_{i,\sigma}|^2 = \sum_i \left| \sum_{\mu,\nu} c_{\mu,\nu} \left(\nabla \phi_\mu \phi_\nu + \phi_\mu \nabla \phi_\nu \right) \right|^2, \tag{19}$$

where the $c_{\mu,\nu}$ are the density matrix elements. The sum extends only over the atomic orbitals, ϕ_μ and ϕ_ν, that give a non-zero contribution at the reference point. The moments and volumes are integrated via summation over a uniform grid.

The value of b is calculated at every point on the integration grid from the corresponding values for the valence spin density ($\rho_\sigma^{\mathrm{val}}$), its gradient norm ($|\nabla \rho_\sigma^{\mathrm{val}}|^2$), its Laplacian ($\nabla^2 \rho_\sigma^{\mathrm{val}}$), and the spin kinetic-energy density (τ_σ). To validate the implementation, these quantities and the resulting b parameter for argon are compared in Figure 1 against the same values using the XDM implementations in Quantum ESPRESSO (plane waves),[31] and the numerical-orbital Numol program,[56,57] as shown in Figure 1. The SIESTA density, its derivatives, and the kinetic-energy density are in good agreement with Numol and Quantum ESPRESSO, except close to the core regions, where the three methods differ in their treatment (Numol is an all-electron code, and SIESTA and QE use different pseudopotentials). Figure 1 confirms that the use of finite-support numerical orbitals in SIESTA does not adversely affect the computed exchange-hole dipole moment values, when compared to either plane-wave or numerical-orbital implementations of XDM.

We also compared the XDM molecular dispersion coefficients calculated using SIESTA and Quantum ESPRESSO. The C_6 dispersion coefficients were calculated for the isolated molecules of the X23 lattice-energy benchmark set.[28,58] The mean percent errors obtained with the DZP and TZP finite-support basis sets were 5.6 and -4.9%, respectively, relative to the B86bPBE plane-wave calculations. Again, this indicates that the use of relatively compact, atom-centered basis functions does not strongly affect the dispersion energies calculated using the XDM method.

Lastly, from the dispersion energy expression (Eq. 1), it is possible to determine the

Figure 1: Quantities used to calculate the XDM dispersion coefficients along the internuclear coordinate in argon. (a) the valence spin density, ρ_σ^{val}; (b) the Weizsaecker term, $|\nabla\rho_\sigma^{\text{val}}|^2/\rho_\sigma^{\text{val}}$; (c) the Laplacian of the spin density, $\nabla^2\rho_\sigma^{\text{val}}$; (d) the spin kinetic-energy density, τ_σ; and (e) the exchange-hole dipole moment, $b = d_{X\sigma}$. All values are in atomic units. The SIESTA (red circles) and Quantum ESPRESSO[49] (blue boxes) calculations used an argon atom centered at the origin of a 3.760 Å cubic box, the PBE functional,[50] and only one **k**-point at Γ. Troullier-Martins[51–53] and Goedecker/Hartwigsen/Hutter/Teter[54,55] norm-conserving pseudopotentials were used, respectively. The Numol (black lines) calculations used the local density approximation (LDA) and an isolated argon dimer. The plots represent the values along the shortest argon-argon contact ($d_{\text{Ar−Ar}} = 3.76$ Å).
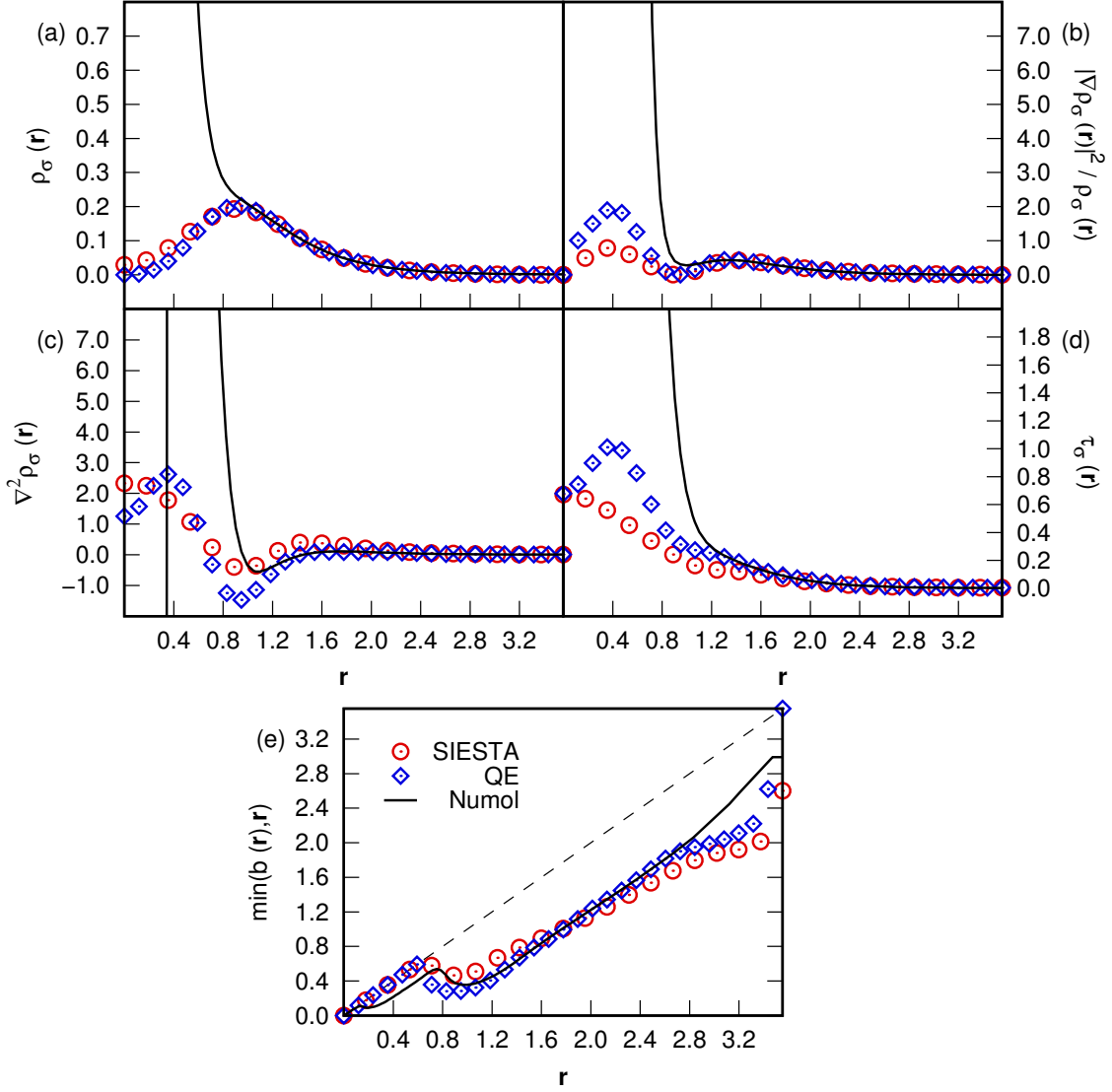
dispersion contribution to the atomic forces and the stress tensor. For atom i, the force is

$$\mathbf{F}_{\text{disp},i} = \sum_{\mathbf{L}}\sum_{j}\sum_{n} \frac{nC_{n,ij}R_{ij\mathbf{L}}^{n-2}}{(R_{\text{vdW},ij}^n + R_{ij\mathbf{L}}^n)^2}\mathbf{R}_{\mathbf{ijL}}, \tag{20}$$

9

and the components of the stress tensor are

$$\sigma_{\text{disp},\eta\zeta} = -\frac{1}{2V} \sum_{\mathbf{L}} \sum_{i\neq j} \sum_n \frac{nC_{n,ij}R_{ij\mathbf{L}}^{n-2}(R_{ij\mathbf{L}})_\eta(R_{ij\mathbf{L}})_\zeta}{(R_{\text{vdW},ij}^n + R_{ij\mathbf{L}}^n)^2}, \tag{21}$$

where $\eta, \zeta = x, y, z$ and V is the unit-cell volume. Note that, in these expressions, the dispersion coefficients are assumed to be constant with respect to changes in the crystal geometry, which is not strictly correct. However, practise has shown that this approximation does not noticeably affect the geometry minimization, in most cases. In fact, it is more computationally efficient to calculate the dispersion coefficients at the first ionic step and then keep them constant throughout the geometry minimization. "Relaxed" geometries are then subject to additional geometry optimization calculations, until the newly computed dispersion coefficients, and consequently the total energy, cease to to change. The effect of fixing the dispersion coefficients at the first ionic step, as opposed to recalculating them at every step, has been tested and shown to yield equivalent geometries in both molecules and solids.

3.2 Computational Methods

SIESTA calculations: The B86bPBE[50,59] functional was implemented in our in-house version of the 4.0b-485 SIESTA code, as it typically yields the best results when paired with the XDM dispersion model.[10] Additional calculations were performed using the PBE[50] functional, with either XDM or Grimme's D2 dispersion correction.[24] Both double-ζ, and triple-ζ, plus polarization (DZP and TZP) basis sets were considered. The D2 damping parameters were set to $s_r = 1.1$ and $s_6 = 0.50$ or 0.64, for DZP and TZP, respectively.[25] DZP is the standard basis set implemented in SIESTA, whereas TZP was constructed and optimized for H, N, and O atoms by Louwerse and Rothenberg,[60] and further extended to C atoms by Carter and Rohl.[61]

The confinement radius of the finite numerical orbitals was set by an "energy shift"

parameter of 0.001 Ry, found to be sufficient in reducing basis set superposition error and yielding converged energies.[18,61,62] The real-space integration grid cutoff value for charge densities and potentials was set to 200 Ry, consistent with our previous work.[18] Troullier-Martins-type[51,52] norm-conserving pseudopotentials[63,64] were generated and tested for both density functionals with the ATOM code.[53] These pseudopotentials included nonlinear core corrections.[65]

Plane-wave calculations: Reference benchmark calculations were also performed with B86bPBE-XDM, PBE-XDM, and PBE-D2 as implemented in Quantum ESPRESSO[49] v. 5.1, using plane-waves/pseudopotentials[66,67] within the projector-augmented wave (PAW) formalism.[68] The damping function parameters, a_1 and a_2, were set to 0.6512, and 1.4633 Å for B86bPBE-XDM and to 0.3275, and 2.7673 Å for PBE-XDM.[10] The s_6 damping parameter for PBE-D2 was set to a value of 0.75. Wave-function and density cutoffs were set to 80 Ry and 800 Ry, respectively. Structure relaxations were performed with tighter thresholds for convergence of the energies and forces, i.e., 10^{-5} Ry and 10^{-4} Ry/bohr, respectively.

k-point grid sampling and structure relaxation: For all calculations, a Brillouin-zone sampling with a $4 \times 4 \times 4$ k-point Monkhorst-Pack (MP) scheme was used to treat crystal structures, whereas isolated molecules were studied at the Γ-point only. During structure optimization, unit-cell parameters and atomic positions for crystalline systems were allowed to fully relax, while for molecules in a large-vacuum simulation box, only the atomic coordinates were allowed to vary.

Geometry relaxations were carried out using the conjugate-gradient (CG) algorithm. However, in the last stages of this study we found that the modified Broyden algorithm[69] implemented in SIESTA is much more efficient than CG, and yields the same results with tight convergence criteria (see below). Therefore, we recommend its use for geometry relaxations in molecular crystals.

Similarity index: As a tool to measure similarity between crystal structures, the POWDIFF utility in CRITIC2[70] was used. This tool is based on the comparison of powder

diffraction patterns using a cross-correlation function,[71] and ranges in value between zero and one. A result of zero indicates an exact match, while a result of one indicates maximum dissimilarity between two crystal structures.

3.3   Parametrization of the XDM dispersion model

Table 1: XDM damping parameters for the PBE and B86bPBE functionals and selected basis sets, along with the resulting error statistics for the fit set.

Method	N^a	a_1	a_2 (Å)	MPE^b	MAPE^c	MAX^d
SIESTA						
PBE/DZP	49	–	–	-17.0	39.5	232.2 (c6h6-c6h6-stack)
PBE-D2/DZP	49	–	–	20.8	33.5	130.6 (cf4-cf4)
PBE-D2/TZP	34	–	–	25.9	26.5	62.0 (ch4-nh3)
PBE-XDM/DZP	49	1.4588	0.0000^e	14.7	30.3	132.0 (cf4-cf4)
PBE-XDM/DZP	34	1.4025	0.0000^e	11.2	22.5	94.3 (ch4-nh3)
PBE-XDM/DZP+CP^f	34	1.2901	0.0000^e	-3.6	19.2	73.9 (ch4-c2h4)
PBE-XDM/TZP	34	0.7086	2.3542	1.1	**12.3**	56.0 (ch4-c2h4)
PBE-XDM/TZP+CP^f	34	1.2480	0.0000^e	-3.1	18.6	74.0 (ch4-c2h4)
B86bPBE/DZP	49	–	–	-32.7	43.3	158.6 (c6h6-c6h6-stack)
B86bPBE-XDM/DZP	49	0.2307	3.4210	8.7	21.7	62.4 (ch4-nh3)
B86bPBE-XDM/DZP	34	0.5000	2.5556	4.9	18.2	65.2 (ch4-nh3)
B86bPBE-XDM/DZP+CP^f	34	1.2343	0.0000^e	-9.7	20.4	80.4 (c6h6-ch4)
B86bPBE-XDM/TZP	34	1.3543	0.0000^e	-1.2	**11.2**	40.4 (ch4-c2h4)
B86bPBE-XDM/TZP+CP^f	34	1.1874	0.0000^e	-8.4	19.2	82.7 (c6h6-ch4)
QE						
PBE-D2	49	–	–	13.2	18.4	69.7 (ch4-hf)
PBE-XDM	49	0.3275	2.7673	3.9	**13.7**	37.8 (h2s-h2s)
B86bPBE-XDM	49	0.6512	1.4633	2.6	**11.4**	23.1 (ch4-nh3)

^{a} Number of molecular dimers contained in the parameterisation set. ^{b} Mean percent error; a negative (positive) sign indicates underbinding (overbinding) with respect to the benchmark data. ^{c} Mean absolute percent error. ^{d} Maximum absolute percent error. Labels in parentheses identify the dimer that gives the maximum error. ^{e} In order to avoid unphysical (negative) values, a_2 was set to zero in the fit. ^{f} Counterpoise corrections were applied to the computed dimer binding energies.

The XDM dispersion model was parametrized for both the (newly implemented) B86bPBE and PBE functionals by minimizing residual errors with respect to high-level benchmark data from the Kannemann-Becke (KB49) set of 49 weakly bound molecular dimers.[40] In all cases, single-point energy calculations were performed at the fixed benchmark set geometries. The optimal parameters obtained by least-squares fit and the performance of the XDM-corrected methods are shown in Table 1 for the standard DZP basis set and the extended TZP basis set.[60,61] Because the TZP basis is not available for all elements contained in the KB49 data (namely, Si, S and F), the XDM model was parametrized on a subset of the KB set containing 34 dimers (of a total of 49, not counting the noble gas dimers). For consistency, the parameters obtained by fitting to the 34-dimer set will be used for both the DZP and TZP

basis sets in the rest of the article.

The B86bPBE average errors in Table 1 are consistently lower than those from PBE, when paired with XDM. This has been observed in previous studies,[31] and is expected from the large-gradient-limit behavior of the exchange enhancement factor in these functionals.[10,72,73] Using the largest basis set in this study (TZP), the MAPE of both functionals (12.3% for PBE and 11.2% for B86bPBE) are very similar to those obtained using QE (13.7% and 11.4%) and also to the near-complete-basis-set values using Gaussian basis sets reported in a previous study[32] (14.3% and 13.1% for the full KB49 set, with the latter value obtained using the psi4 program[7,74]). For comparison, PBE-D2 gives MAPEs of 33.5% (DZP), 26.5% (TZP), and 18.4% using the QE implementation. However, it must be noted that the PBE-D2 was not specifically fitted to the KB set.

The performance of B86bPBE-XDM and PBE-XDM suffers considerably from basis set incompleteness when the smaller DZP basis set provided in the SIESTA package are used, with MAPE of 18.2% (B86bPBE) and 22.5% (PBE) (21.7% and 30.3%, respectively, if the 49 dimers in the KB set are considered). These values are consistent with the results for double-ζ Gaussian basis sets lacking diffuse functions, which have been shown in the past to be inadequate for non-covalent interactions.[32,75] For instance, the MAPE of the PBE-XDM parametrization to the 49-dimer KB set with non-diffuse double-ζ Gaussian basis sets is 37.2% (6-31G∗) and 45.9% (cc-pVDZ). In contrast, the MAPE of 6-31+G∗, which contains one set of diffuse functions, is 17.8%. SIESTA's DZP basis set, with a 30.3% MAPE (Table 1), is intermediate between these two results and the TZP basis sets benefits from the increased cutoff radii compared to DZP. Other functionals, such as PBE-D2, are similarly affected by basis-set incompleteness, with MAPEs of 26.5% (TZP) and 33.5% (DZP). Therefore, it is clear that reliable calculation of non-covalent binding energies in SIESTA necessitates the use of larger basis sets than the standard DZP, or the design of new basis sets with increased cutoff radii.

The use of counterpoise (CP) corrections[76] to account for basis set superposition error

(BSSE), in conjunction with the DZP or TZP basis sets, does not improve the performance of any method other than PBE-XDM/DZP. The CP correction tends to over-compensate, causing the dimers to be underbound on average, in agreement with previous reports.[77] The size of the CP correction indicates that, even if the average error from its application is lower, there is still significant basis-set incompleteness error in the TZP results, probably stemming from the finite-support nature of the basis sets.

3.4   Graphite exfoliation

We now calculate the exfoliation of graphite using the new XDM-corrected functionals in SIESTA.[31] Graphite exfoliation is a simple test of the accuracy in the treatment of non-covalent interactions for which high-level experimental reference data exists.[78] For each method in Table 1, a scan was performed by systematically varying the interlayer distance between graphene sheets in graphite, while the intralayer hexagonal lattice parameter was kept fixed at 2.456 Å. The resulting potential energy curves are shown in Figure 2.

In all cases, the energy curves for the uncorrected functionals (PBE and B86bPBE) are very slightly binding, or non-binding, and the XDM dispersion correction brings the results into close agreement with experiment. The equilibrium interlayer separations using DZP and TZP are slightly higher ($\approx$0.1–0.2 Å) than the plane-wave results and the experimental reference data. The exfoliation energies are in excellent agreement with experiment and the QE calculations, although slightly underestimated with the DZP basis set. Strangely, TZP gives greater interlayer binding than DZP with all functionals, with or without dispersion correction. In particular, uncorrected B86bPBE and PBE show between 10–20 meV/atom of spurious binding, another indication that there is significant basis set incompleteness even at the TZP level. Overall, the DFT-XDM methods perform slightly better at reproducing the experimental graphite interlayer distance and exfoliation energy than PBE-D2, regardless of the choice of local or plane-wave basis sets.

There is an important point to note about geometry optimizations in weakly bound crys-

Figure 2: Graphite exfoliation curves calculated with plane-wave/pseudopotentials in (a) QE, (b) SIESTA using a DZP basis set, and (c) SIESTA using a TZP basis set, compared to experimental data.[78]
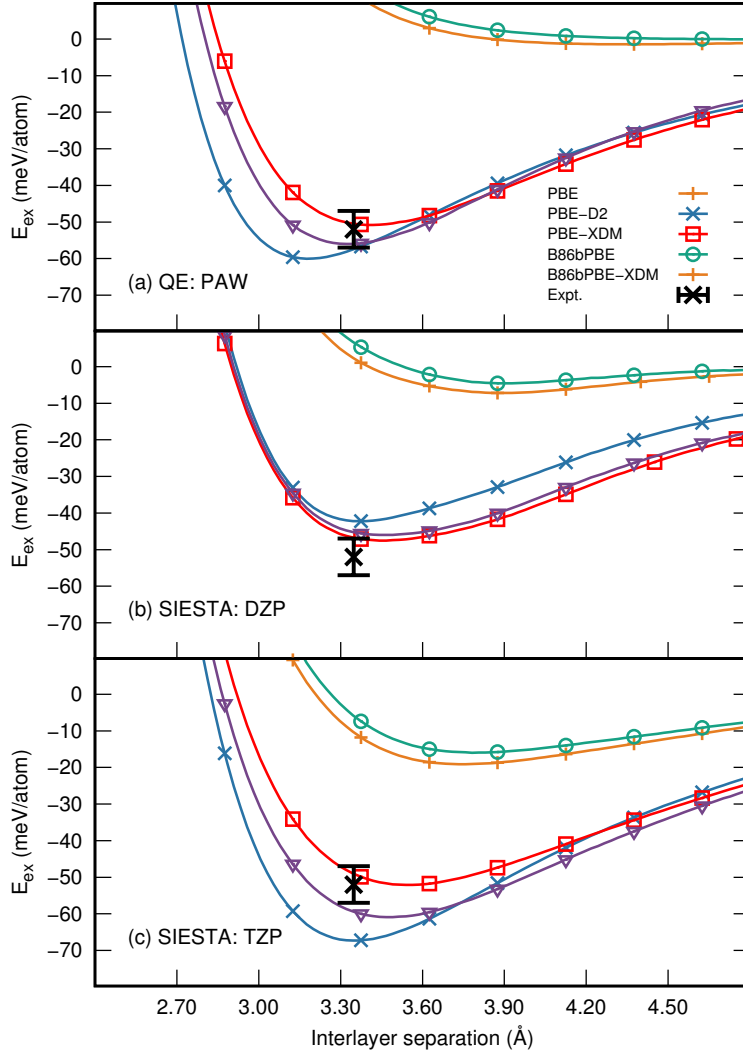
tals using SIESTA. The default convergence thresholds for maximum atomic force and maximum stress components in SIESTA are 0.04 eV/Å and 1.0 GPa, respectively. While these are suitable for hard solids, they lead to unfinished geometry optimizations in molecular crystals, which makes the energy landscapes dependent on the choice of input geometries.[18] This is illustrated by varying maximum stress component threshold for calculations on graphite (Table 2). The default convergence threshold (1.0 GPa) leads to incomplete geometry optimizations and erroneous equilibrium geometries. From these tests and previous work,[18] maximum force and stress thresholds of 0.01 eV/Å and 0.02 GPa, which are tighter than the

Table 2: Results of geometry optimizations on graphite performed with DFT-XDM/DZP and DFT-D2/DZP using various input values of the c lattice parameter (specified in parentheses), as a function of the stress convergence threshold.[a]

Stress thr.	N_{opt}	c (6.0)	N_{opt}	c (7.0)	N_{opt}	c (8.0)
		B86bPBE-XDM				
1.0	21	6.939	6	6.998	31	7.315
0.02	78	6.936	58	6.937	146	6.940
0.002	126	6.936	179	6.937	248	6.937
		PBE-XDM				
1.0	21	6.988	6	6.999	30	7.069
0.02	47	6.996	23	6.998	163	6.995
0.002	309	6.976	291	6.977	314	6.991
		PBE-D2				
1.0	13	6.871	2	6.994	19	6.820
0.02	19	6.796	46	6.797	62	6.798
0.002	66	6.795	105	6.795	137	6.795

[a] The units are GPa for the cell stresses, and Å for the c lattice parameter. N_{opt} is the number of optimization steps.

defaults, seem to be sufficient. These convergence thresholds will be used in the rest of the article. It is important to note that excessively tight stress thresholds are also problematic because of numerical instabilities that arise near the cutoff radii of the finite-support atomic orbitals when performing numerical integration.[19,20] This can lead to "endless" geometry optimizations, as pointed out by the SIESTA developers.

3.5   Lattice energies of molecular crystals

We now assess the SIESTA implementation of XDM for molecular crystals. The statistics on the X23 benchmark set are shown in Table 3. The X23 set [28,58] contains reference lattice energies for 23 small-molecule crystals. With the DZP basis set, B86bPBE-XDM performs best overall, yielding an MAE of 8.2 kJ/mol, followed by PBE-D2 and PBE-XDM. However, the use of the larger TZP basis set does not necessarily improve the quality of the results. Similar to Carter and Rohl,[61] we observe that the lattice energy error statistics with finite-support basis sets are somewhat poorer than using plane waves. Carter and Rohl obtained a MAE of ca. 23 kJ/mol on the C21 subset of X23 (with 21 lattice energies[28]) using the

16

Table 3: Statistics for the X23 set of lattice energies using DFT-XDM and DFT-D2 methods, in kJ/mol per molecule, relative to back-corrected, experimental sublimation enthalpies.[28,58]

Method	ME[a]	MAE[b]	MAX[c]
SIESTA			
PBE-D2/DZP	7.5	11.4	25.1 (ant)
PBE-D2/TZP	3.9	5.7	14.5 (ada)
PBE-XDM/DZP	11.8	14.1	29.9 (suc)
PBE-XDM/TZP	-9.4	10.0	23.0 (cyt)
B86bPBE-XDM/DZP	3.1	8.2	24.3 (ant)
B86bPBE-XDM/TZP	-8.4	8.9	21.3 (cyt)
PBE-XDM/DZP+CP[d]	-0.8	4.4	14.4 (ant)
PBE-XDM/TZP+CP[d]	-2.1	4.4	18.2 (ant)
B86bPBE-XDM/DZP+CP[d]	-3.2	4.7	15.8 (ant)
B86bPBE-XDM/TZP+CP[d]	-3.6	4.7	17.5 (ant)
QE			
PBE-D2	3.7	5.8	18.4 (ada)
PBE-XDM	-3.2	4.7	17.9 (cyt)
B86bPBE-XDM	0.5	3.6	13.4 (cyt)
QE//SIESTA[e]			
PBE-D2/DZP	2.9	5.5	13.7 (cya)
PBE-XDM/DZP	-2.5	4.3	17.2 (cyt)
B86bPBE-XDM/DZP	-0.1	3.7	12.7 (cyt)

[a] Mean error; a negative (positive) ME indicates underbinding (overbinding) with respect to the benchmark data. [b] Mean absolute error. [c] Maximum absolute error; labels in parentheses identify the crystal responsible (ada: adamantane, ant: anthracene, cya: cyanamide, cyt: cytosine, suc: succinic acid). [d] Counterpoise corrections were applied to the crystal lattice energies, following the approach of Carter and Rohl.[61] [e] Composite methods using plane-wave single-point energies evaluated using the geometries obtained from SIESTA with the same functional and the DZP basis.

non-local vdW-DF methods and a DZP basis set. For comparison, plane-wave calculations with the same functionals yield MAEs of ca. 10 kJ/mol.[28] Thus, for molecular crystals, B86bPBE-XDM/DZP outperforms the vdW-DF methods, even when employed with plane-waves, at a considerably reduced computational cost. (Although it is possible that some of the more recent exchange functionals used in combination with vdW-DF work better for lattice energies of molecular crystals.)

Carter and Rohl also demonstrated that the application of counterpoise corrections greatly reduced the errors in the computed lattice energies with SIESTA, restoring the perfor-

mance of the vdW-DF functionals to the same quality as their plane-wave implementation.[61] Upon applying CP corrections, all of the DFT-XDM methods yield MAEs of 4.4–4.7 kJ/mol, approaching those of the plane-wave reference calculations. However, as argued previously,[61] although CP corrections significantly improve lattice energies and return values similar to plane-wave calculations, they are not straightforward to apply.

The impact of basis set incompleteness error on the lattice energies in Table 3 raises the question of whether the equilibrium geometries are similarly affected. This question is also important in the context of composite methods, in which high-level single-point calculations (using plane-waves) are used at geometries obtained using a low-level calculation (SIESTA). Composite methods have been proposed as a computationally efficient alternative to plane waves for the calculation of lattice energies in molecular crystals,[18,61] but their performance relies on whether the low-level method offers equilibrium geometries close to those of the high-level method.

Table 4: Statistics for the X23 set of lattice energies using composite methods, in kJ/mol per molecule, relative to plane-wave calculations with the same density functional.

	SIESTA	QE//SIESTA			
Method	MAE^a	ME^b	MAE^a	MAX^c	POWDIFF^d
PBE-D2/DZP	10.8	-0.8	1.2	5.4 (ada)	0.2935
PBE-XDM/DZP	15.0	0.8	0.8	2.0 (pyr)	0.1578
B86bPBE-XDM/DZP	7.0	0.5	0.6	1.2 (eth)	0.1975

^{a} Mean error; a negative (positive) ME indicates underbinding (overbinding) with respect to the benchmark data. ^{b} Mean absolute error. ^{c} Maximum absolute error; labels in parentheses identify the crystal responsible (ada: adamantane, eth: ethylcarbamate, pyr: pyrazole). ^{d} The deviation between the SIESTA and QE equilibrium geometries is quantified by the powder diffraction similarity measure.

Table 4 evaluates the ability of DZP calculations with SIESTA to reproduce plane-wave equilibrium geometries, using the same density functional and dispersion correction. The plane-wave geometries are close to the complete-basis-set limit and, in our experience, the choice of PAW dataset/pseudopotential has very little effect on non-covalent equilibrium geometries, provided enough plane waves are used in the calculation. The powder similarity measure (POWDIFF) indicates that there are significant differences between the SIESTA

and plane-wave geometries, and that this deviation is somewhat smaller for XDM than for D2. However, the lattice energies calculated using the corresponding composite methods are excellent, with MAEs being within at most 1.2 kJ/mol from the pure plane-wave calculations. Therefore, even though the lattice energies calculated using the DZP basis set are significantly affected by basis-set incompleteness error (Table 3), the corresponding equilibrium geometries seem to be reasonably close to the plane-wave reference and, consequently, the composite methods built with a DFT-XDM/DZP low-level approach are quite accurate and significantly cheaper than a pure plane-wave optimization.

4   Conclusions

In this article, we presented the first implementation of the XDM dispersion model with the numerical finite-supported orbital method in the SIESTA software package. The new SIESTA/XDM code was verified by comparing the calculated dispersion coefficients to other XDM implementations using plane waves (QE) and numerical orbitals in the gas phase (Numol). The XDM method in SIESTA was then parametrized for the PBE and the B86bPBE functionals with double-ζ (DZP) and triple-ζ (TZP) basis sets against the Kannemann-Becke (KB) set of binding energies of non-covalently bound gas-phase dimers. DZP is the standard basis set implemented in SIESTA, while TZP, which contains more basis functions, has been recently proposed for C, H, O, and N by Louwerse and Rothenberg[60] and Carter and Rohl.[61] The performance of the new XDM-corrected methods for the molecular dimers in the KB set is of similar quality to the results using plane-waves, provided that the TZP basis set is employed. DZP, on the other hand, suffers from significant basis-set incompleteness, and its performance is similar to double-ζ Gaussian basis sets without diffuse functions.

The new XDM implementation was tested in two cases: graphite exfoliation and the calculation of lattice energies of molecular crystals. The application of the XDM correction brings the potential energy curves for graphite exfoliation into close agreement with exper-

imental data, although care needs to be taken to use tighter stress convergence thresholds than the SIESTA defaults.

For molecular crystal lattice energies, the XDM-corrected SIESTA methods also give good results. In particular, B86bPBE-XDM/DZP showed promise as an excellent balance between accuracy and efficiency when determining lattice energies. The performance of this method is a significant improvement over the currently implemented post-SCF dispersion corrections (DFT-D2) and non-local functionals (vdW-DF1 and vdW-DF2) in SIESTA. However, both the DZP and TZP basis sets display considerable basis set incompleteness effects, which reveals the necessity of designing specialized SIESTA basis sets for non-covalent interactions. Because of this, SIESTA's XDM-corrected results are inferior to their plane-wave equivalents for the lattice energies of molecular crystals unless counterpoise corrections are applied, which is undesirable in practice. The equilibrium geometries obtained from all XDM methods in SIESTA, however, are quite close to the plane-wave results. It was shown that composite approaches, in which a single-point plane-wave calculation is performed at SIESTA's DZP equilibrium geometries, are quite accurate and computationally efficient.

Acknowledgement

Supporting Information Available

Computed and benchmark lattice energies for the X23 set crystal structures. This material

is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Otero-de-la-Roza, A.; DiLabio, G. *Non-Covalent Interactions in Quantum Chemistry and Physics*; Elsevier: Amsterdam, Netherlands, 2017.

(2) Dale, S. G.; Johnson, E. R. The ionic *versus* Metallic Nature of 2D Electrides: A Density-Functional Description. *Phys. Chem. Chem. Phys.* **2017**, *19*, 27343–27352.

(3) Lorenzo, A.; Escribano, B.; Akhmatskaya, E.; Carrasco, J. Assessment of van der Waals Inclusive Density Functional Theory Methods for Layered Electroactive Materials. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10133–10139.

(4) Tawfik, S. A.; Gould, T.; Stampfl, C.; Ford, M. J. Evaluation of can der Waals Density Functionals for Layered Materials. *Phys. Rev. Materials* **2018**, *2*, 034005.

(5) Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg, J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; Case, D. H.; Chadha, R.; Cole, J. C.; Cosburn, K.; Cuppen, H. M.; Curtis, F.; Day, G. M.; DiStasio Jr, R. A.; Dzyabchenko, A.; van Eijck, B. P.; Elking, D. M.; van den Ende, J. A.; Facelli, J. C.; Ferraro, M. B.; Fusti-Molnar, L.; Gatsiou, C.-A.; Gee, T. S.; de Gelder, R.; Ghiringhelli, L. M.; Goto, H.; Grimme, S.; Guo, R.; Hofmann, D. W. M.; Hoja, J.; Hylton, R. K.; Iuzzolino, L.; Janckiewicz, W.; de Jong, D. T.; Kendrick, J.; de Klerk, N.; Ko, H.-Y.; Kuleshova, L. N.; Li, X.; Lohani, S.; Leusen, F.; Lund, A. M.; Lv, J.; Ma, Y.; Marom, N.; Masunov, A. E.; McCabe, P.; McMahon, D. P.; Meekes, H.; Metz, M. P.; Misquitta, A. J.; Mohamed, S.; Monserrat, B.; Needs, R. J.; Neumann, M. A.; Nyman, J.; Obata, S.; Oberhofer, H.; Oganov, A. R.; Orendt, A. M.; Pagola, G. I.; Pantelides, C. C.; Pickard, C. J.; Podeszwa, R.; Price, L. S.; Price, S. L.; Pulido, A.; Read, M. G.; Reuter, K.; Schneider, E.; Schober, C.; Shields, G. P.;

Singh, P.; Sugden, I. J.; Szalewicz, K.; Taylor, C. R.; Tkatchenko, A.; Tucker-man, M. E.; Vacarro, F.; Vasileiadis, M.; Vazquez-Mayagoitia, A.; Vogt, L.; Wang, Y.; Watson, R. E.; de Wijs, G. A.; Yang, J. Z.; Zhu, Q.; Groom, C. R. Report on the Sixth Blind Test of Organic Crystal-Structure Prediction Methods. *Acta Crystallogr., Sect. B: Struct. Sci.* **2016**, *72*, 439–459.

(6) Whittleton, S. R.; Otero-de-la-Roza, A.; Johnson, E. R. Exchange-Hole Dipole Dispersion Model for Accurate Energy Ranking in Molecular Crystal Structure Prediction. *J. Chem. Theory Comput.* **2017**, *13*, 441–450.

(7) Whittleton, S. R.; Otero-de-la-Roza, A.; Johnson, E. R. Exchange-Hole Dipole Dispersion Model for Accurate Energy Ranking in Molecular Crystal Structure Prediction II: Non-Planar Molecules. *J. Chem. Theory Comput.* **2017**, *13*, 5332–5342.

(8) Beran, G. Modelling Polymorphic Molecular Crystals with Electronic Structure Theory. *Chem. Rev.* **2016**, *116*, 5567–5613.

(9) Grimme, S. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.

(10) Johnson, E. R. In *Non-Covalent Interactions in Quantum Chemistry and Physics*; Otero-de-la-Roza, A., DiLabio, G., Eds.; Elsevier: Amsterdam, Netherlands, 2017; Chapter 5, pp 169–192.

(11) Hermann, J.; DiStasio Jr., R. A.; Tkatchenko, A. First-Principles Models for van der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications. *Chem. Rev.* **2017**, *117*, 4714–4758.

(12) Berland, K.; Cooper, V. R.; Lee, K.; Schröder, E.; Thonhauser, T.; Hyldgaard, P.; Lundqvist, B. I. van der Waals Forces in Density Functional Theory: A Review of the vdW-DF Method. *Rep. Prog. Phys.* **2015**, *78*, 066501.

(13) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 Dispersion Coefficient Model. *J. Chem. Phys.* **2017**, *147*, 034112.

(14) Tao, J.; Zheng, F.; Gebhardt, J.; Perdew, J. P.; Rappe, A. M. Screened van der Waals Correction to Density Functional Theory for Solids. *Phys. Rev. Materials* **2017**, *1*, 020802(R).

(15) Caldeweyher, E.; Brandenburg, J. G. Simplified DFT Methods for Consistent Structures and Energies of Large Systems. *J. Phys.: Condens. Matter* **2018**, *30*, 213001.

(16) Kumar Prasad, V.; Otero-de-la-Roza, A.; DiLabio, G. A. Atom-Centered Potentials with Dispersion-Corrected Minimal-Basis-Set-Hartree-Fock: An Efficient and Accurate Computational Approach for Large Molecular Systems. *J. Chem. Theory Comput.* **2018**, *14*, 726–738.

(17) Dolgonos, G. A.; Loboda, O. A.; Boese, A. D. Development of Embedded and Performance of Density Functional Methods for Molecular Crystals. *J. Phys. Chem. A* **2018**, *122*, 708–713.

(18) LeBlanc, L. M.; Otero-de-la-Roza, A.; Johnson, E. R. Composite and Low-Cost Approaches for Molecular Crystal Structure Prediction. *J. Chem. Theory Comput.* **2018**, *14*, 2265–2276.

(19) Soler, J. M.; Artacho, E.; Gale, J. D.; García, A.; Junquera, J.; Ordejón, P.; Sánchez-Portal, D. The SIESTA Method for *Ab Initio* Order-N Materials Simulation. *J. Phys.: Condens. Matter* **2002**, *14*, 2745–2779.

(20) Artacho, E.; Gale, J. D.; García, A.; Junquera, J.; Martin, R. M.; Ordejón, P.; Pruneda, J. M.; Sánchez-Portal, D.; Soler, J. M. The SIESTA Method; Developments and Applicability. *J. Phys.: Condens. Matter* **2008**, *20*, 064208.

(21) Sánchez-Portal, D.; Ordejón, P.; Artacho, E.; Soler, J. M. Density-Functional Method for Very Large Systems with LCAO Basis Sets. *Int. J. Quantum Chem.* **1997**, *65*, 453–461.

(22) Artacho, E.; Sánchez-Portal, D.; Ordejón, P.; García, A.; Soler, J. M. Linear-Scaling ab-initio Calculations for Large and Complex Systems. *Phys. Stat. Sol.* **1999**, *215*, 809–817.

(23) Junquera, J.; Paz, O.; Sánchez-Portal, D.; Artacho, E. Numerical Atomic Orbitals for Linear-Scaling Calculations. *Phys. Rev. B* **2001**, *64*, 235111.

(24) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(25) Peverati, R.; Baldridge, K. K. Implementation and Performance of DFT-D with Respect to Basis Set and Functional for Study of Dispersion Interactions in Nanoscale Aromatic Hydrocarbons. *J. Chem. Theory. Comput.* **2008**, *4*, 2030–2048.

(26) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. Van der Waals Density Functional for General Geometries. *Phys. Rev. Lett.* **2004**, *92*, 246401.

(27) Lee, K.; Murray, E. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. Higher-Accuracy van der Waals Density Functional. *Phys. Rev. B* **2010**, *82*, 081101.

(28) Otero-de-la-Roza, A.; Johnson, E. R. A Benchmark for Non-Covalent Interactions in Solids. *J. Chem. Phys.* **2012**, *137*, 054103.

(29) Christian, M. S.; Otero-de-la-Roza, A.; Johnson, E. R. Surface Adsorption from the Exchange-Hole Dipole Moment Dispersion Model. *J. Chem. Theory Comput.* **2016**, *12*, 3305–3315.

(30) Becke, A. D.; Johnson, E. R. Exchange-hole Dipole Moment and the Dispersion Interaction Revisited. *J.Chem. Phys.* **2007**, *127*, 154108.

(31) Otero-de-la-Roza, A.; Johnson, E. R. Van der Waals Interactions in Solids Using the Exchange-Hole Dipole Moment Model. *J. Chem. Phys.* **2012**, *136*, 174109.

(32) Otero-de-la-Roza, A.; Johnson, E. R. Non-Covalent Interactions and Thermochemistry using XDM-Corrected Hybrid and Range-Separated Hybrid Density Functionals. *J. Chem. Phys.* **2013**, *138*, 204109.

(33) Otero-de-la-Roza, A.; Johnson, E. R. Predicting Energetics of Supramolecular Systems Using the XDM Dispersion Model. *J. Chem. Theory Comput.* **2015**, *11*, 4033–4040.

(34) Christian, M. S.; Otero-de-la-Roza, A.; Johnson, E. R. Adsorption of Graphene to Nickel (111) Using the Exchange-Hole Dipole Moment Model. *Carbon* **2017**, *118*, 184–191.

(35) Christian, M. S.; Otero-de-la-Roza, A.; Johnson, E. R. Adsorption of Graphene to Metal (111) Surfaces Using the Exchange-Hole Dipole Moment Model. *Carbon* **2017**, *124*, 531–540.

(36) Otero-de-la-Roza, A.; Cao, B. H.; Price, I. K.; Hein, J. E.; Johnson, E. R. Predicting the Relative Solubilities of Racemic and Enantiopure Crystals by Density-Functional Theory. *Angew. Chem. Int. Ed.* **2014**, *53*, 7879–7882.

(37) LeBlanc, L. M.; Otero-de-la-Roza, A.; Johnson, E. R. Evaluation of Shear-Slip Transitions in Crystalline Aspirin by Density-Functional Theory. *Cryst. Growth Des.* **2016**, *16*, 6867–6873.

(38) Yang, Y.; Rice, B.; Shi, X.; Brandt, J. R.; Correa da Costa, R.; Hedley, G.; Smilgies, D. M.; Frost, J. M.; Samuel, I.; Otero-de-la-Roza, A.; Johnson, E. R.; Jelfs, K. E.; Nelson, J.; Campbell, A. J.; Fuchter, M. J. Emergent Properties of an Organic Semiconductor Driven by its Molecular Chirality. *ACS Nano* **2017**, *11*, 8329–8338.

(39) Rice, B.; LeBlanc, L. M.; Otero-de-la-Roza, A.; Fuchter, M. J.; Johnson, E. R.; Nelson, J.; Jelfs, K. E. A Computational Exploration of the Crystal Energy and Charge Carrier Mobility Landscapes of the Chiral [6]helicene Molecule. *Nanoscale* **2018**, *10*, 1865–1876.

(40) Kannemann, F. O.; Becke, A. D. van der Waals Interactions in Density-Functional Theory: Intermolecular Complexes. *J. Chem. Theory Comput.* **2010**, *6*, 1081–1088.

(41) Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124*, 174104.

(42) Salem, L. The Calculation of Dispersion Forces. *Mol. Phys.* **1960**, *3*, 441–452.

(43) Otero-de-la-Roza, A.; Johnson, E. R. Many-Body Dispersion Interactions from the Exchange-Hole Dipole Moment Model. *J. Chem. Phys.* **2013**, *138*, 054103.

(44) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theoret. Chim. Acta* **1997**, *44*, 129–138.

(45) Heidar-Zadeh, F.; Ayers, P. W.; Verstraelen, T.; Vinogradov, I.; Vöhringer-Martinez, E.; Bultinck, P. Information-Theoretic Approaches to Atoms-in-Molecules: Hirshfeld Family of Partioning Schemes. *J. Phys. Chem. A* **2018**, *122*, 4219–4245.

(46) Kannemann, F. O.; Becke, A. D. Atomic Volumes and Polarizabilities in Density-Functional Theory. *J. Chem. Phys.* **2012**, *136*, 034109.

(47) Becke, A. D.; Roussel, M. R. Exchange Holes in Inhomogeneous Systems: A Coordinate Space Model. *Phys. Rev. A* **1989**, *39*, 3761–3767.

(48) Becke, A. D.; Johnson, E. R. A Density-Functional Model of the Dispersion Interaction. *J. Chem. Phys.* **2005**, *123*, 154101.

(49) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Buongiorno Nardelli, M.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M.; Colonna, N.; Carnimeo, I.; Dal Corso, A.; de Gironcoli, S.; Delugas, P.; DiStasio, R.; Ferretti, A.; Floris, A.; Fratesi, G.; Fugallo, G.; Gebauer, R.; Gerstmann, U.; Giustino, F.; Gorni, T.; Jia, J.; Kawamura, M.; Ko, H.-Y.; Kokalj, A.; Küçükbenli, E.; Lazzeri, M.; Marsili, M.; Marzari, N.; Mauri, F.; Nguyen, N. L.; Nguyen, H.-V.; Otero-de-la-Roza, A.; Paulatto, L.; Poncé, S.; Rocca, D.; Sabatini, R.; Santra, B.; Schlipf, M.; Seitsonen, A.; Smogunov, A.; Timrov, I.; Thonhauser, T.; Umari, P.; Vast, N.; Baroni, S. Advanced Capabilities for Materials Modelling with Quantum ESPRESSO. *J. Phys.: Condens. Matter* **2017**, *29*, 465901.

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

(51) Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B* **1991**, *43*, 1993–2006.

(52) Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. II. Operators for Fast Iterative Diagonalization. *Phys. Rev. B* **1991**, *43*, 8861–8869.

(53) *ATOM, a program for DFT calculations in atoms and pseudopotential generation, maintained by Alberto Garcia and distributed as part of the SIESTA package. See* `http://www.icmab.es/siesta/atom`

(54) Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710.

(55) Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic Separable Dual-Space Gaussian Pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58*, 3641–3662.

(56) Becke, A. D.; Dickson, R. M. Numerical-Solution of Schrodinger-Equation in Polyatomic-Molecules. *J. Chem. Phys.* **1990**, *92*, 3610–3612.

(57) Becke, A. D.; Johnson, E. R. Exchange-Hole Dipole Moment and the Dispersion Interaction. *J. Chem. Phys.* **2005**, *122*, 154104.

(58) Reilly, A. M.; Tkatchenko, A. Understanding the Role of Vibrations, Exact Exchange, and Many-Body van der Waals Interactions in the Cohesive Properties of Molecular Crystals. *J. Chem. Phys.* **2013**, *139*, 024705.

(59) Becke, A. D. On the Large-Gradient Behavior of the Density Functional Exchange Energy. *J. Chem. Phys.* **1986**, *85*, 7184–7187.

(60) Louwerse, M. J.; Rothenberg, G. Transferrable Basis Sets of Numerical Atomic Orbitals. *Phys. Rev. B* **2012**, *85*, 035108.

(61) Carter, D. J.; Rohl, A. L. Benchmarking Calculated Lattice Parameters and Energies of Molecular Crystals Using van der Waals Density Functional. *J. Chem. Theory Comput.* **2014**, *10*, 3423–3437.

(62) Chapman, C.; Ting, E. C.; Kereszti, A.; Paci, I. Self-Assembly of Cysteine Dimers at the Gold Surface: A Computational Study of Competing Interactions. *J. Phys. Chem. C* **2013**, *117*, 19426–19435.

(63) Hamann, D. R.; Schlüter, M.; Chiang, C. Norm-Conserving Pseudopotentials. *Phys. Rev. Lett.* **1979**, *43*, 1494–1497.

(64) Bachelet, G. B.; Hamann, D. R.; Schlüter, M. Pseudopotentials That Work: From H to Pu. *Phys. Rev. B* **1982**, *26*, 4199–4228.

(65) Louie, S. G.; Froyen, S.; Cohen, M. L. Nonlinear Ionic Pseudopotentials in Spin-Density-Functional Calculations. *Phys. Rev. B* **1982**, *26*, 1738–1742.

(66) Pickett, W. E. Pseudopotential Methods in Condensed Matter Applications. *Comput. Phys. Rep.* **1989**, *9*, 115–197.

(67) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. Iterative Minimzation Techniques for Ab Initio Total-Energy Calculations: Molecular Dynamics and Conjugate Gradients. *Rev. Mod. Phys.* **1992**, *64*, 1045–1097.

(68) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(69) Johnson, D. D. Modified Broyden's Method for Accelerating Convergence in Self-Consistent Calculations. *Phys. Rev. B* **1988**, *38*, 807–813.

(70) Otero-de-la-Roza, A.; Johnson, E. R.; Luaña, V. CRITIC2: A Program for Real-Space Analysis of Quantum Chemical Interactions in Solids. *Comput. Phys. Commun.* **2014**, *185*, 1007–1018.

(71) de Gelder, R.; Wehrens, R.; Hageman, J. A. A Generalized Expression for the Similarity of Spectra: Application to Powder Diffraction Pattern Classification. *J. Comput. Chem.* **2001**, *22*, 273–289.

(72) Lacks, D. J.; Gordon, R. G. Pair Interactions of Rare-Gas Atoms as a Test of Exchange-Energy-Density Functionals in Regions of Large Density Gradients. *Phys. Rev. A* **1993**, *47*, 4681–4690.

(73) Zhang, Y.; Pan, W.; Yang, W. Describing van der Waals Interaction in Diatomic Molecules with Generalized Gradient Approximations: The Role of the Exchange Functional. *J. Chem. Phys.* **1997**, *107*, 7921–7925.

(74) Parrish, R. M.; Burns, L. A.; Smith, D.; Simmonett, A. C.; DePrince, A. E.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197.

(75) Johnson, E. R.; Otero-de-la-Roza, A.; Dale, S. G.; DiLabio, G. A. Efficient Basis Sets

for Non-Covalent Interactions in XDM-Corrected Density-Functional Theory. *J. Chem. Phys.* **2013**, *139*, 214109.

(76) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553–566.

(77) Mackie, I. A.; DiLabio, G. A. Approximations to Complete Basis Set-Extrapolated, Highly Correlated Non-Covalent Interaction Energies. *J. Chem. Phys.* **2011**, *135*, 134318.

(78) Zacharia, R.; Ulbricht, H.; Hertel, T. Interlayer Cohesive Energy of Graphite from Thermal Desorption of Polyaromatic Hydrocarbons. *Phys. Rev. B* **2004**, *69*, 155406.

Graphical TOC Entry