Chemical Modelling

Volume 13

A Specialist Periodical Report

Chemical Modelling

Volume 13

Editors
J.-O. Joswig, *Technische Universität Dresden, Dresden, Germany*
M. Springborg, *University of Saarland, Saarbrücken, Germany*

Authors
Benjamin Fiedler, *Technische Universität Chemnitz, Germany*
Joachim Friedrich, *Technische Universität Chemnitz, Germany*
Abhisek Ghosal, *Indian Institute of Science Education and Research (IISER) Kolkata, India*
K. D. Jordan, *University of Pittsburgh, PA, USA*
Mathieu Linares, *Linköping University, Sweden*
Sougata Pal, *University of Gour Banga, India*
Amlan K. Roy, *Indian Institute of Science Education and Research (IISER) Kolkata, India*
Supriya Saha, *The University of Newcastle, Australia*
Pranab Sarkar, *Visva-Bharati University, India*
Sunandan Sarkar, *Visva-Bharati University, India*
Tobias Schwabe, *University of Hamburg, Germany*
K. Sen, *University of Pittsburgh, PA, USA*
J. K. Singh, *Indian Institute of Technology Kanpur, India*
Roy Tasker, *Purdue University, IN, USA*
Riccardo Volpi, *Linköping University, Sweden*

ROYAL SOCIETY
OF CHEMISTRY

THE QUEEN'S AWARDS
FOR ENTERPRISE:
INTERNATIONAL TRADE
2013

A catalogue record for this book is available from the British Library

Preface

Michael Springborg[a] and Jan-Ole Joswig[b]

You have opened the 13th volume of the *Specialist Periodical Reports* on *Chemical Modelling*. We, the editorial team, have selected reviews on present hot topics and active areas in computational chemistry and related fields. As you will see from the table of contents, we have accumulated eight chapters on four different scientific areas:

Volume 13 starts with two chapters on materials for solar cells highlighting different aspects. Whereas the first concentrates on methods for modelling the functioning of organic solar cells with a combination of classical and quantum mechanics, the second contribution focusses on a specific class of materials – carbon-based nanohybrids – and their electronic properties.

The second topic in this volume deals with liquids – in the broadest sense of understanding. We have one chapter that introduces the reader in recent developments of modelling fluids near surfaces on large scales. Here, special attention to wetting and prewetting is paid. It is a matter of opinion to see small water clusters as liquids as well. However, we present a contribution on modelling these clusters in neutral or charged states in this volume.

The third aspect covers quantum-chemical methods. Method development should be a key competence of every quantum chemist, because different problems need different modelling approaches. We present in this volume chapters on theory and applications of the incremental method, double hybrid density functional approximations combining concepts from (standard) hybrid density functionals with correlation corrections from wave function theory, and density-functional calculations on a Cartesian grid.

We close our selection of topics with a chapter of a rather special area that sometimes is forgotten, but nonetheless of utmost importance: education. Many students are able to draw chemical formulae, but when it comes to wave functions and orbitals, their imagination is put to the test. The final chapter starts with a citation of Peter W. Atkins, who said that the principle target of our education "should be to find a way to bridge the imagined to the perceived". This is an important aspect of science: explaining the findings to colleagues, money sources, or family.

We very much hope that you will enjoy the selection of topics from different areas, authors from different countries, and views from different sides. Forthcoming issues of SPR Chemical Modelling are planned already, but we will be grateful for additional suggestions with respect to authors or subjects.

[a]*Saarbrücken, Germany. E-mail: m.springborg@mx.uni-saarland.de*
[b]*Dresden, Germany. E-mail: jan-ole.joswig@chemie.tu-dresden.de*

CONTENTS
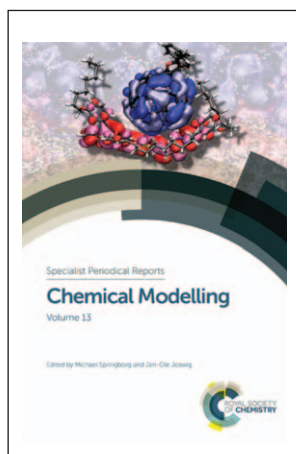
Cover

Image provided by Mathieu Linares, Linköping University, Sweden.

Specialist Periodical Reports

Chemical Modelling

Volume 13

Edited by Michael Springborg and Jan-Ole Joswig

DFT calculations of atoms and molecules in Cartesian grids   221

Abhisek Ghosal and Amlan K. Roy

Molecular-level visualisation for educational purposes   261

Roy Tasker

Organic solar cells

Riccardo Volpi[a] and Mathieu Linares*[a,b]

DOI: 10.1039/9781782626862-00001

1 Introduction

As climate change and energy sufficiency are progressively becoming more pressing issues, environmentally friendly and cheap energy sources need to be developed. Solar energy production, with inorganic solar cells, has progressed greatly in recent years and is already able to compete with traditional energy sources. Photovoltaic solar cells convert photons into electric current. Although traditional photovoltaic technology based on inorganic materials has become commercially successful, it faces some limitations that can be overcome by organic solar cells. Organic solar cells are low-cost and easy to process, furthermore they possess innovative properties being potentially lightweight, flexible, and transparent. However, organic solar cells still have lower efficiencies and shorter lifetimes than traditional inorganic solar cells. While the best organic solar cells have reached around 11% efficiency, the best single junction crystalline silicon solar cells and thin film CdTe cells have efficiencies of around 25% and 22%, respectively.[1,2] Furthermore, the lifetime of organic solar cells is still short in comparison to the lifetimes of inorganic solar cells, so stability challenges must be addressed for organic solar cells to compete with conventional photovoltaics on the market.[3–5] The efficiency of an organic solar cell is determined by the efficiency of the different steps from photon absorption to charge collection, as shown in Fig. 1. Photons are absorbed in either the acceptor or the donor phase with efficiency η_{abs}, generating excitons (1). These excitons then either diffuse toward the interface and form a Charge Transfer (CT) state (2) or decay (6). The CT state is defined as an electron–hole pair Coulombically bound, composed usually by an electron on an acceptor molecule and a hole on a donor molecule. In an organic solar cell CT states (3) are formed at the interface from exciton dissociation. The efficiency of the excitons forming CT states is η_{ex}, and it is high when the distance to the interface is small. The CT states then split (4) into free electrons and holes with efficiency η_{diss}. After the charge carriers are freed, they may still move back to the interface and recombine in a CT state (8), or even to the ground state (7) and (9). η_{trans} is the efficiency of the free charge carriers being collected at the electrodes (5). The product of these different efficiencies gives the external quantum efficiency (EQE), the ratio of charge carriers collected to the number of incoming photons.[6–12]

$$EQE(E) = \eta_{abs}\eta_{ex}\eta_{diss}\eta_{trans} \tag{1}$$

[a]Department of Physics, Chemistry and Biology (IFM), Linköping University, SE-581 83 Linköping, Sweden. E-mail: mathieu@ifm.liu.se
[b]Swedish e-Science Research Centre (SeRC), Linköping University, SE-581 83 Linköping, Sweden

(1) Light is absorbed and an exciton is formed
(2) Diffusion of the exciton to the interface
(3) CT state
(4) CT state dissociation
(5) Transport of the charges to the electrodes
(6) Exciton decay
(7) CT state decay (geminate)
(8) Free carriers recombination into a CT state
(9) CT state decay (non geminate)
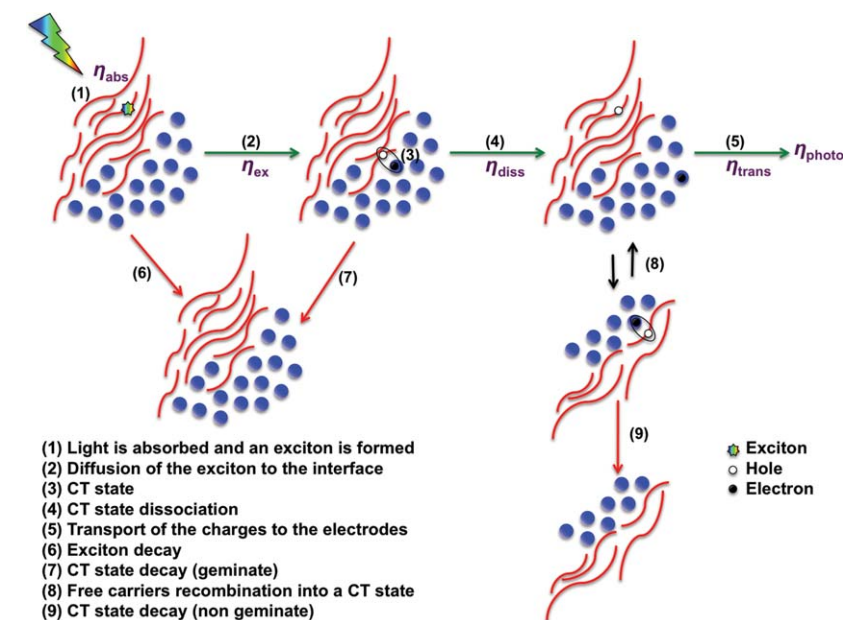
Fig. 1 Steps from photon absorption to collection of free charge carriers at the electrodes. Adapted from ref. 6. [colors online]

The efficiency of organic solar cells is greatly impacted by the processes that occur at the interface between the donor and acceptor. The electron and hole have a strong Coulombic attraction, which must be overcome for the charges to separate. Organic solar cells have much lower dielectric constants and more localized electronic states than inorganic solar cells, so excited states are localized and the Coulombic barrier to overcome for the charges to dissociate is large.

The amorphous nature of these materials make analytical studies difficult and theoretical investigations of mobility and charge transport are to a large extent based on simulations: drift diffusion and Kinetic Monte-Carlo simulations (KMC). Drift–diffusion models[13–16] employ a classical picture, they are very suitable to simulate the whole solar cell and give some macroscopic information. Kinetic Monte-Carlo simulations[17–28] have potentially the capability to model the quantum phenomena happening at the nanoscale, but the more details are included in the KMC scheme the more computational effort is required. Detailed KMC schemes are thus mainly limited in studying some interesting portions of a solar cell. Several studies employing both approaches showed the significant contribution of morphology to the efficiency of solar cells.[14–16,18] Different structures for organic solar cells have been researched, the simplest consisting of a single layer of an organic semiconductor between two electrodes. However, solar cells with this structure have low efficiency, and the performance of the device is improved by a bilayer structure of two organic materials: a donor and an acceptor materials.[29] With this design, a trade-off must be made between light

absorption and exciton dissociation. The thicker the layers are, the more light will be absorbed, increasing the efficiency of the solar cell. However, the excitons formed in the single material, will have in average a greater distance to the donor–acceptor interface, where they can dissociate in a CT state. If the domain size is too large, the exciton will decay before reaching the interface, thereby lowering the efficiency of the solar cell. A solution to this problem is to combine the donor and acceptor phases so that the distance to the interface is short even with thick films. This can be best achieved with interdigitated structures,[14,16] providing good exciton dissociation and at the same time a clear path to the electrodes for the free charge carriers. The interdigitated interface is very promising but also very difficult to obtain in practice. Another type of interface commonly used in the lab employs a blend of donor and acceptor materials forming a three-dimensional interpenetrated structure called bulk heterojunction (BHJ) solar cell.[30]

In the present chapter we will focus on the KMC approach used to model charge transport in organic materials. The properties of organic materials are different from those of crystals; their intrinsic disorder tends to localize the charge carriers on one or few molecules. The conduction in these materials is therefore temperature activated, in contrast to the band conduction of crystalline materials. Two methods are predominantly used in literature to study charge transport using KMC simulations: the Miller-Abrahams and the Marcus hopping rates. The Miller-Abrahams[31] formula is one of the simplest ways to couple temperature to the hopping rate. This is achieved by means of a Boltzmann factor helping to overcome the energy barrier for the transport. The Marcus hopping rate[32] can be derived from the Fermi golden rule and takes into account also the reorganization energy after each hop.

In this book chapter, we will show how to model the steps (3), (4) and (5) of Fig. 1, namely, how the CT state split and the charges are subsequently transported to the electrodes. We will thus not consider the transport of excitons[12] or the coupling between the excited state, the CT state and the ground state[33–35] and we refer the reader to the appropriate literature. After a brief discussion on the method to obtain realistic interfaces with atomistic details, we will focus on the transport through Marcus equation and on the way to calculate the different parameters involved in it. We will also analyze the effect of the electric field on CT state splitting and conduction of free charge carriers. In particular, in order to optimize the device efficiency, we will illustrate how to choose the electric field to favor simultaneously both processes, if this is possible.

2 Morphology

The morphology of the donor–acceptor interface in a BHJ organic solar cell has a large impact on the efficiency of the solar cell. Excitons can only diffuse 10–20 nm before decaying, so the donor and acceptor should be

sufficiently mixed so that the excitons can reach the interface before decaying. However, once separated, the charge carriers need pathways to their respective electrodes. If, for example, an electron is in an acceptor domain that is completely surrounded by the donor, then the electron will have no path to the electrode and it will eventually decay through recombination. In a bilayer structure, there is always a pathway for free charges to travel to the electrode, but many excitons will decay before they can reach the interface and thus have a chance to split. In the bulk heterojunction structure, however, it is much easier for excitons to reach the interface before decaying, but some of the charges will get trapped in domains that have no path to the electrode or recombine with other charges when traveling by an interface. Two types of recombination can occur: *geminate* or *nongeminate*. Geminate recombination is when two charge carriers resulting from the absorption of the same photon recombine (7 in Fig. 1). Nongeminate recombination (9 in Fig. 1) occurs when two free charge carriers originating from different photons recombine with each other at the interface. When there is high inter-penetration, free charge carriers are likely to travel close to an interface and recombine with other free charge carriers, increasing the non-geminate recombination rate. It is important to optimize the inter-penetration of the donor and acceptor materials in order to extract as many charge carriers as possible from the photogenerated excitons.

Charges hop between molecules or polymer segments, and this hopping occurs with a certain probability, making probabilistic methods such as KMC appropriate for modeling charge transport. In early works[17,18] the molecules of an organic material were modeled only as a lattice of molecular sites, and the structure of the molecules was neglected. More recent works have improved upon these models by including an atomistic description of the molecules.[19–28,36–39] To study the complex interplay of geminate and non-geminate recombination and thus obtain a meaningful simulation of charge transport in the solar cell, a realistic morphology is needed. Several methods can be used to model the interface between donor and acceptor materials. As a first approxi-mation, the crystal structure of the two materials can be brought into contact and the interdistance can be optimized without relaxing the molecule inside each phase.[38,39] In order to add more disorder, it is then possible to relax the interface by using molecular dynamics simulations like it has been done in several studies.[35,40] However, if this technique will allow to introduce disorder and create mixing at the interface, this latter will remain quite planar and it will not reach a real three-dimensional interpenetrated structure. In order to obtain a realistic BHJ interface we are currently working on another approach. The idea is to start from volumes obtained with the Ising model developed by Heiber *et al.* with different degrees of intermixing,[41] and subsequently fill these volumes with donor and acceptor molecules as illustrated in Fig. 2. From those initial structures, molecular dynamics in the NPT ensemble can be performed to relax the different boxes that can be then used for Kinetic Monte Carlo simulations.
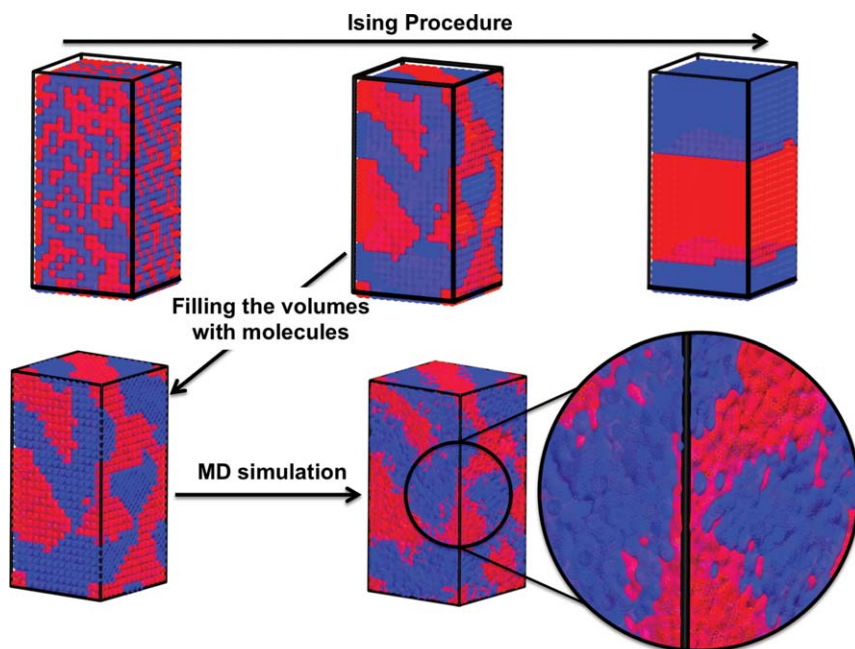
Fig. 2 Strategy to build an interpenetrated network. Acceptor and donor sites are represented in dark gray [blue] and light gray [red] respectively. The filling has been done with C_{60} (acceptor) and anthracene (donor). [colors online]

3 Transport through Marcus equation: kinetic Monte Carlo

Kinetic Monte Carlo simulations are one of the prominent tools for the simulation of charge carriers (single or multiple) in organic materials. In the notation used in this chapter, a particular molecular orbital level in the system is identified by an upper case letter $M = (i, m)$, *i.e.* the orbital level m of the molecule i. The molecular orbital M will always implicitly belong to the molecule i, unless explicitly stated. In the same way, the orbital $N = (j, n)$ belongs to the molecule j (where i and j are different molecules). The flowchart of the KMC algorithm is presented in Fig. 3. Since organic materials exhibit usually a disordered amorphous structure, it is common practice in the literature to obtain a realistic structure as outcome of a MD simulation (see previous Section 2). Once the structure has been defined, every charge carrier a involved in the simulation is placed in its respective initial positions (step 1). At every time step the hopping rates for all the charge carriers are calculated with Marcus formula[32] (step 2). If a charge carrier is situated on the molecular orbital M, its probability to hop to the orbital N is

$$w_{MN} = \frac{2\pi}{\hbar} |H_{NM}|^2 \sqrt{\frac{1}{4\pi k_B T \lambda_{NM}}} \exp\left(-\frac{(\Delta E_{NM} + \lambda_{NM})^2}{4\lambda_{NM} k_B T}\right). \qquad (2)$$
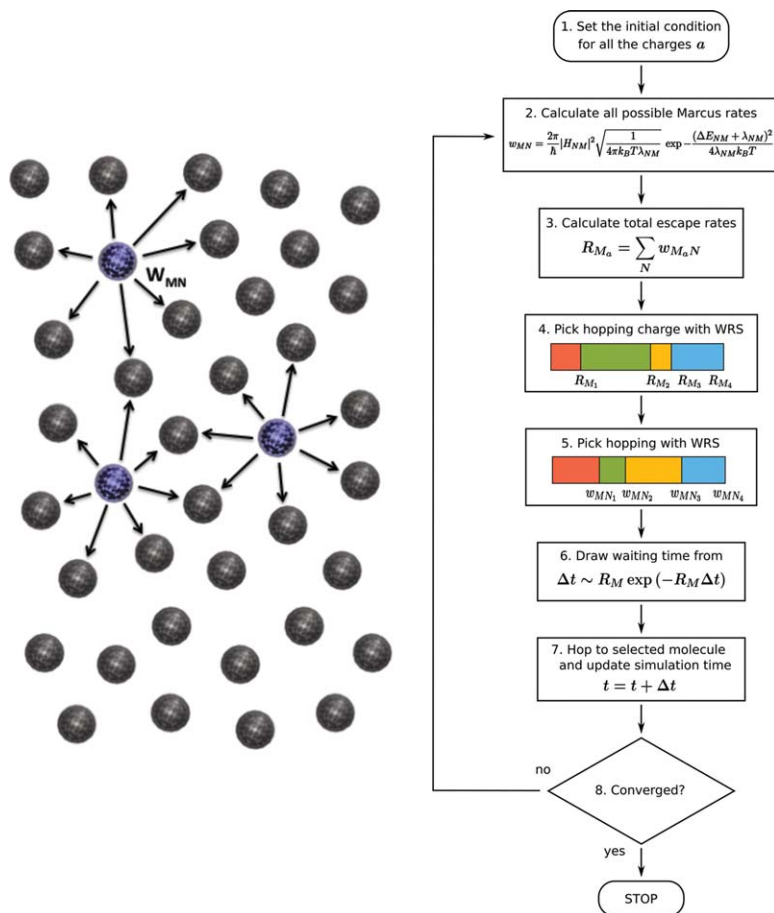
Fig. 3 Flowchart of the KMC simulation. [colors online]

Marcus formula involves temperature T (through the Boltzmann constant k_B) and three main physical quantities depending on the particular hop: the transfer integral H_{NM}, the site energy difference ΔE_{NM} and the reorganization energy λ_{NM}. The evolution of the system is determined thanks to a weighted random selection algorithm (WRS), in which the probability of selecting a particular hop depends on its associated weight. First, the hopping charge carrier is selected using as weight for every charge carrier a the total escape rates R_{M_a} (step 3). This selects the charge that is jumping, thus identifying the origin of the hop (step 4). Let us call c the selected hopping charge and let us assume such charge is placed on the orbital M. Now, among all the possible hops that charge c can make, we will select the one determining the evolution of the system through another WRS using as weights the Marcus rates w_{MN} (step 5). This hop will be considered in the simulation as happening after a time Δt sampled from an exponential distribution $\sim R_M \exp(-R_M \Delta t)$, with R_M being the total escape rate from the orbital M, *i.e.* $R_M = \sum_N w_{MN}$ (step 6).

At step 7 the system is updated with the selected hop and the time of the

simulation is increased by the estimated time for such hop calculated at step 6. The simulation can be stopped due to several stopping criteria. A stopping criterion can be the simulation time or the number of hops. Also positional criteria can be used, like for example checking if a charge carrier is passing through a particular plane in space. If no stopping criterion is respected (step 8), we go back at step 2 and continue the simulation. Periodic boundary conditions can be used in x, y and/or z. In the following of this section, we will explain how the three main parameters in Marcus equation (the transfer integral H_{NM}, the site energy difference ΔE_{NM} and the reorganization energy λ_{NM}) can be calculated with a combination of quantum chemistry and classical mechanics tools.

3.1 Site energy difference

The site energy difference is the difference in energy of the charge carrier before and after the jump. The energy of a charge carrier can be calculated with several methods. For instance, it can be sampled from a Gaussian distribution with a certain width reflecting the energetical disorder of the system. This technique was initially proposed in the pioneering work by Bässler[17] and it allows to phenomenologically recover the broadening of the density of states (DOS) of a charge carrier in the bulk. This approach allowed to gain great insight into the charge transport mechanisms of organic materials and also to achieve qualitatively the correct mobility temperature dependence. However, to achieve the correct mobility field dependence instead, the correlation of the energy landscape is mandatory as demonstrated by Novikov and Vannikov[42,43] and subsequently showed in a KMC scheme.[44]

To model a realistic correlation of the energy landscape in the bulk, it is possible to consider in a more sophisticated way the energy of a charge placed on the molecular orbital M as composed by four terms

$$u_M = \pm \epsilon_M + E_i^{\text{perm-el}}(\underline{s}) + E_i^{\text{ind-el}}(\underline{s}) + E_i^{\text{field}}(s_i). \tag{3}$$

ϵ_M is the energy of the orbital $M = (i, m)$ and the $\pm$ sign depends on the type of charge carrier, if the energy of the orbital has to be added (electron) or subtracted (hole). This can be evaluated either at the semi-empirical or quantum mechanical level for each and every molecule in the box in order to consider the energetic shift due to specific conformation. The other terms represent the effect of the environment and they are depending on the positions and the quantum states of the molecules around. The array $\underline{s} = (s_1, s_2, \ldots, s_i, s_j, \ldots)$ represents the state of the system and it is composed by the quantum states of all the molecules. After every charge hopping, the state of the system will change. If for example an electron jumps from the LUMO (Lowest Unoccupied Molecular Orbital) of molecule i to the LUMO of molecule j, the state passes from $\underline{s_1} = (0, 0, \ldots, \text{LUMO}, 0, \ldots)$ before the jump, to $\underline{s_2} = (0, 0, \ldots, 0, \text{LUMO}, \ldots)$ after the jump. The term $E_i^{\text{perm-el}}(\underline{s})$ considers the Coulombic interactions with permanent atomic charges within a cut-off distance, so that a charge carrier located on molecule i interacts only with the molecules inside its interaction set $I(i)$, i.e. the set composed by the
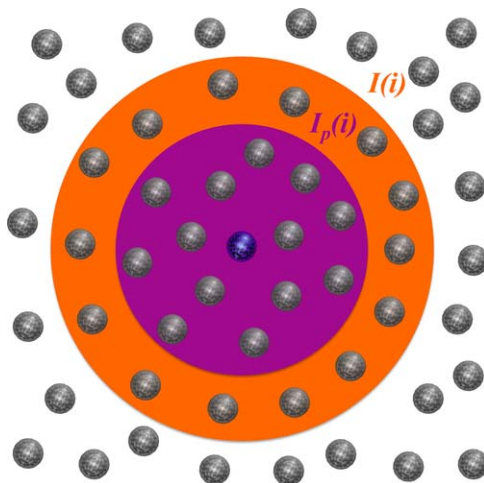
Fig. 4 The interaction sets used for Coulombic interactions $I(i)$ and polarization $I_p(i)$, for a charge carrier situated on the molecular orbital $M = (i, m)$. [colors online]

molecules nearer than the cut-off (Fig. 4). The interaction of the charge carrier with the permanent atomic charges of the surrounding molecules can be expressed as

$$E_i^{\text{perm-el}}(\underline{s}) = \frac{1}{4\pi\epsilon_0} \sum_{a \in i} \sum_{\substack{b \in h \\ h \in I(i)}} \frac{q_a(s_i) q_b(s_h)}{|\vec{r}_a - \vec{r}_b|}, \qquad (4)$$

where ϵ_0 is the vacuum dielectric constant, h is a molecule belonging to the interaction set $I(i)$ and a, b are atoms. Every atom a has a position $\vec{r}_a$ and a permanent atomic charge $q_a(s_i)$ (which depends on the quantum state of molecule i).

For the polarization several approaches have been considered. For instance, it is possible to retrieve polarizabilities from polarizable force fields, in particular from works of van Duijnen and Swart (ref. 45) and the more recent work of Wang, Cieplak *et al.* (ref. 46) In our approach, we use another cut-off, leading to a different interaction set denoted by $I_p(i)$, see Fig. 4. Atoms belonging to the molecules in this set will be polarized according to Thole model,[47] using linear atomic polarizabilities. These polarizabilities can for example be calculated[28] at the QM level using the LoProp localization method.[48] The charge carrier interaction with the nearby induced atomic dipoles is computed

$$E_i^{\text{ind-el}}(\underline{s}) = \frac{1}{4\pi\epsilon_0} \sum_{a \in i} \sum_{\substack{b \in h \\ h \in I_p(i)}} \frac{q_a(s_i)\, \vec{p}_b(\underline{s}) \cdot (\vec{r}_a - \vec{r}_b)}{|\vec{r}_a - \vec{r}_b|^3}, \qquad (5)$$

where $p_b(\underline{s})$ represents the induced dipole at the atom b, given that the molecules of the system are in the state $\underline{s}$. The dipoles inside $I_p(i)$ are calculated depending on the permanent charges of all the molecules inside the interaction set $I(i)$. To avoid polarization border effects, the

cut-off chosen for polarization has to be significatively smaller than the cut-off for Coulombic interactions.

The last term of eqn (3) considers the field contribution to the site energies

$$E_i^{\text{field}}(s_i) = -\vec{E}_{\text{ext}} \cdot \sum_{a \in i} q_a(s_i)\vec{r}_a, \tag{6}$$

where $\vec{E}_{\text{ext}}$ is the external field considered.

3.2 Transfer integrals

Transfer integrals can be evaluated in several different ways. We will present in this subsection a non-exhaustive list of methods to calculate them.

3.2.1 Distance dependent. The simplest method to calculate the transfer integral is the exponential decay or distance dependent (DD) transfer integral

$$H_{NM} = H_0 \exp(-\alpha r_{NM}), \tag{7}$$

where the transfer integral is considered to be exponentially decaying with the distance r_{NM} between the two molecular orbitals M and N. H_0 and α are two parameters that need to be chosen carefully for the molecules at hand. This method was widely used in the early studies where molecules were represented only by a site. An obvious limitation of this method is of course the fact that the relative orientation of the two molecules does not influence the value of the transfer integral.

3.2.2 Distance orientation dependent. It is possible to improve the Distance Dependent model through a weighted Mulliken formula[49] for carbon atoms[27] to take into account the relative orientation of the molecules.

$$H_{NM} = \sum_{a \in i} \sum_{b \in j} C_{ma}^i C_{nb}^j S_{ab} \tag{8}$$

where S_{ab} is the overlap between two $2p\pi$ orbitals and C_{ma}^i is the $2p\pi$ expansion coefficient related to the atom a for the orbital level m of molecule i. S_{ab} can be calculated with a Mulliken formula modified to consider the relative orientation of the two $2p\pi$ atomic orbitals on the atoms a and b.[27,50] This approach is based on the approximation that the transfer integral is proportional to the molecular wavefunctions overlap where only the p electrons contributes to the transfer integral.

3.2.3 Energy splitting dimer. This method is valid for a symmetric dimer. In this case the extra charge can be considered roughly equally distributed over the two molecules (having equal conformation) and the transfer integral H_{MN} is equal to half the energy difference between the energetic levels of molecules i and j. Therefore, once the geometry

of both the molecules is known, a single point QM calculation can be performed and the transfer integral can be calculated as

$$H_{NM} = \frac{\epsilon_{L+1[H]} - \epsilon_{L[H-1]}}{2} \, , \tag{9}$$

respectively for electron (LUMO) and hole (HOMO) transport. This method is called the energy-splitting-dimer (ESD) method.[51] It has been used widely because of its simplicity and low computational cost. However, the fact that this method breaks for non-symmetric dimers makes it difficult to use in real 3D bulk case studies.

3.2.4 Projection method. It is of course also possible to calculate the transfer integral directly from its definition, as non-diagonal element of the QM Hamiltonian. For every pair of nearby molecules (nearer than a chosen cut-off), three QM calculations are performed: one for the pair and two for the single molecules. The Fock matrix of the pair is then projected on the single molecular orbitals basis (the basis of the molecular orbitals of the two single molecules) and the transfer integrals $|H_{NM}|$ can be extracted as the non-diagonal elements of this projected Fock matrix in the single molecular orbitals basis. This is known in literature as the projection method. It was originally formulated for ZINDO[52] and it has been subsequently extended[53] for more sophisticated QM calculations that do not assume orthogonality of atomic orbitals. In its quite general formulation it considers three QM calculations: for the molecule i, for the molecule j and for the dimer ij (also indicated as d). For each of these calculations we obtain a set of wavefunctions

$$|\psi_m\rangle = \sum_\nu \varphi_\nu C_{\nu m} \tag{10}$$

expanded in the atomic basis set $\{\varphi_\nu\}$ and identified by the quantum numbers of the particular orbital level m. The transfer integral from $M = (i, m)$ to $N = (j, n)$ can be written as

$$H_{NM} = \langle \psi_m^i | H | \psi_n^j \rangle = \sum_{r,s} \langle \psi_m^i | \psi_r^d \rangle \langle \psi_r^d | H | \psi_s^d \rangle \langle \psi_s^d | \psi_n^j \rangle, \tag{11}$$

where we have used the completeness of the eigenfunctions of the dimer $\{|\psi_r^d\rangle\}$. If we now assume the dimer system d in the void, we can write

$$\langle \psi_r^d | H | \psi_s^d \rangle = \epsilon_r^d \delta_{rs}, \tag{12}$$

where ϵ_r^d is the eigenvalue associated to the eigenfunction $|\psi_r^d\rangle$ and δ_{rs} is the Kronecker delta. Substituting eqn (10) and (12) in eqn (11) we arrive at

$$H_{NM} = \Pi^{id} F^d (\Pi^{jd})^T, \tag{13}$$

where F^d is the Fock matrix of the dimer, composed by the energies $\{\epsilon_r^d\}$. Let us assume there are n_i atomic basis functions $\{\varphi_\nu^i\}$ for molecule i and n_j atomic basis functions $\{\varphi_\mu^j\}$ for molecule j, then there will be $n_d = n_i + n_j$ atomic basis functions $\{\varphi_\chi^d\}$ for the dimer and these basis

functions can be rearranged such that the first n_i functions of $\{\varphi_\chi^d\}$ are equal to $\{\varphi_\nu^i\}$, while the remaining n_j atomic basis functions for the dimer are equal to $\{\varphi_\mu^j\}$. The matrices Π^{id} and Π^{jd} have dimensions $n_i \times n_d$ and $n_j \times n_d$ respectively, and they take care of the projection of the eigenfunction of the dimer on the eigenfunctions of the single molecules and *vice versa*. They can be defined element-wise as

$$\Pi_{mr}^{id} = \langle \psi_m^i | \psi_r^d \rangle \tag{14}$$

$$\Pi_{nr}^{jd} = \langle \psi_n^j | \psi_r^d \rangle. \tag{15}$$

As already pointed out, in all the methods introduced in this section the transfer integrals are evaluated excluding the surrounding molecules. However, several studies showed that the polarization can have a non negligible effect on the transfer integral.[54–59] For instance, Castet *et al.* showed by using a valence bond approach that the inclusion of polarization can decrease the value of transfer integrals by 10–15% for a hole in anthracene.[59]

3.3 Reorganization energy
The reorganization energy λ is the energy required for all the structural adjustments of the system, after the charge hopping has occurred. It can be decomposed in two parts, an internal and an external reorganization contribution

$$\lambda = \lambda_{\text{int}} + \lambda_{\text{out}}.$$

3.3.1 Internal reorganization. The internal reorganization energy (λ_{int}) is related to the internal geometry rearrangement of the molecule, after a charge is added or subtracted to the molecule. Let us consider a charge hopping from molecule i to molecule j (Fig. 5). The internal reorganization energy is the sum of the reorganization energy of the
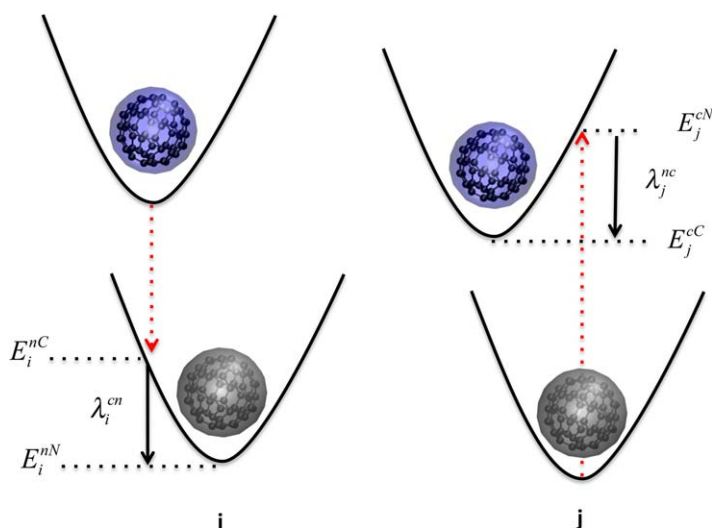

Fig. 5 Internal reorganization energy. [colors online]

molecule i that is passing from charged to neutral state and the reorganization energy of the molecule j that is passing from neutral to charged state

$$\lambda_{\text{int}} = \lambda_i^{cn} + \lambda_j^{nc}, \tag{16}$$

where

$$\lambda_i^{cn} = E_i^{nC} - E_i^{nN}$$
$$\lambda_j^{nc} = E_j^{cN} - E_j^{cC}. \tag{17}$$

E_i^{nC}, for example, is the energy of molecule i in the geometrical conformation of the charged molecule (C), but in the neutral electronic configuration (n). All the other terms of eqn (17) can be interpreted analogously. These energies can be calculated by performing quantum mechanical geometry optimizations for the neutral and charged molecules followed by two single point calculations.[25,51]

3.3.2 External reorganization. The external reorganization energy (λ_{out}) is related to the rearrangement of the environment after the charge carrier's jump. It can be calculated for example from the electric displacement fields created by the charge transfer complex[26]

$$\lambda_{ij} = \frac{c_p}{2\epsilon_0} \int_{V^{\text{out}}} (\vec{D}_I(\vec{r}) - \vec{D}_F(\vec{r}))^2 \mathrm{d}V. \tag{18}$$

ϵ_0 is the permittivity of free space, $\vec{D}_I(\vec{r})$ and $\vec{D}_F(\vec{r})$ are the electric displacement fields created by the charge transfer complex in the initial (charge on molecule i) and final (charge transferred to molecule j) states, respectively. V^{out} is the volume outside the complex, and

$$c_p = \frac{1}{\epsilon_{\text{opt}}} - \frac{1}{\epsilon_s} \tag{19}$$

is the Pekar factor, determined by the low (ϵ_s) and high (ϵ_{opt}) frequency dielectric permittivities.

Another possibility to calculate the external reorganization energy is by using atomic induced dipoles and computing the energy needed for a re-polarization of the environment.[28] In these calculations we neglect the rearrangement of the permanent nearby multipoles, *i.e.* we do not perform molecular dynamics during the KMC simulation. The MD structural rearrangement is much slower than both the wavefunctions polarization and the internal reorganization of the molecule (fast reorganization of covalent bonds), thus it is happening on a different timescale respect to charge conduction. The polarization contribution is taken into account by means of induced dipoles on the nearby molecules. Let us consider the energy contribution

$$E_i^{\text{ind-el}}(\underline{s_1}, \underline{s_2}) = \frac{1}{4\pi\epsilon_0} \sum_{a \in i} \sum_{\substack{b \in h \\ h \in I_p(i)}} \frac{q_a(\underline{s_{1i}}) \, \vec{p}_b(\underline{s_2}) \cdot (\vec{r}_a - \vec{r}_b)}{|\vec{r}_a - \vec{r}_b|^3} , \tag{20}$$
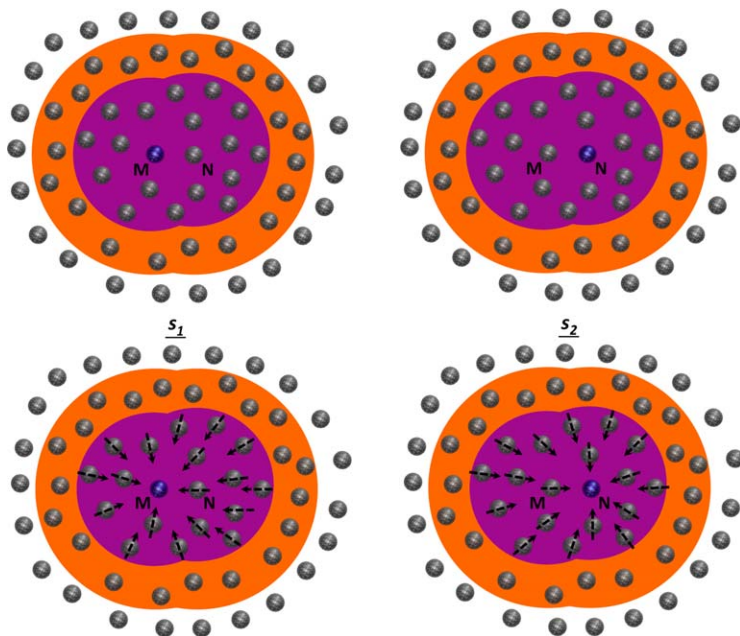
Fig. 6 On top, the description of two quantum states: $\underline{s_1}$ where an extra charge is on molecular orbital M and $\underline{s_2}$ where an extra charge is on molecular orbital N. On the bottom, the polarization induced by the two different states. [colors online]

where as above a and b are atoms, orbital M belongs to molecule i, and h is a molecule in the polarization interaction set of i. This is the interaction energy between the charges on i as given by the quantum state $\underline{s_1}$ and the nearby dipoles induced by the quantum state $\underline{s_2}$. When a jump from $M = (i, m)$ to $N = (j, n)$ occurs, the quantum states are modified from $\underline{s_1}$ to $\underline{s_2}$, *i.e.* the state of i changes from charged to uncharged and the state of j changes from uncharged to charged. An illustration of the states $\underline{s_1}$ and $\underline{s_2}$ with their respective induced dipoles are presented in Fig. 6. The external reorganization energy can be expressed as the difference between the energy of the new charge configuration in the old dipoles arrangement (the jump happened but the structure did not relax yet) and the energy of the new charge configuration in the new dipoles arrangement (system has re-polarized through dipoles reorganization)

$$\lambda_{\text{out}} = E_i^{\text{ind-el}}(\underline{s_2}, \underline{s_1}) + E_j^{\text{ind-el}}(\underline{s_2}, \underline{s_1}) - E_i^{\text{ind-el}}(\underline{s_2}, \underline{s_1}) - E_j^{\text{ind-el}}(\underline{s_2}, \underline{s_1}). \quad (21)$$

The energetic contribution between the electric field and the induced dipoles is not considered. For more details on this method of calculating the external reorganization energy we refer the reader to ref. 28.

4   CT state splitting diagram

The CT state splitting into free charge carriers is one of the main processes determining the efficiency of an organic solar cell. Kinetic Monte Carlo simulations can be a very useful tool for the study of such a system.

Nevertheless, even once all the material parameters have been calculated, the electric field has to be given as input of the simulation. An iterative process based on trial and error is needed to establish which are the fields at which the CT state splits. This process of determination of the meaningful set of electric fields for the specific case study is slow to perform since many simulations are needed at each field value to obtain a statistically significant ensemble to analyze. Here we describe a method allowing to analyze the qualitative behaviour of a bilayer interface, based on the parameters calculated for Marcus formula, but without having to run KMC simulations. It is therefore possible to determine *a priori* the interesting sets of fields for the system under consideration and if the two chosen donor and acceptor materials are suitable to be used together in an OPV device.

The hopping probabilities used in the KMC simulation are calculated with Marcus formula (eqn (2)). As detailed in Section 3 both the transfer integrals and the internal reorganization energies are calculated without any external field or extra charges. Theoretical studies support the independence of transfer integrals and internal reorganization energy from the external electric field.[36,60] Also the external reorganization energy (eqn (21)) is independent of the external field and additional charges, due to the linear polarization method used. The only parameter influenced by the field is thus the energy difference ΔE, at least in a first approximation.

Since we are interested in the role of the electric field $\vec{E}$ on the hopping rates, we study the Marcus rate w_{MN} (eqn (2)) while everything is kept constant except ΔE_{NM}, the only parameter affected by the electric field. Marcus rate w_{MN}, reported here for the reader's convenience,

$$w_{MN} = \frac{2\pi}{\hbar} |H_{NM}|^2 \sqrt{\frac{1}{4\pi k_{\mathrm{B}} T \lambda_{NM}}} \exp\left(-\frac{(\Delta E_{NM} + \lambda_{NM})^2}{4\lambda_{NM} k_{\mathrm{B}} T}\right)$$

has the form of a Gaussian with standard deviation $\sigma_{NM} = \sqrt{2\lambda_{NM} k_{\mathrm{B}} T}$ and the maximum rate for the electron movement is obtained when $\Delta E_{NM} + \lambda_{NM} = 0$. Rates significantly different from 0 are obtained when $\Delta E_{NM} \in (-\lambda_{NM} - 3\sigma_{NM}, -\lambda_{NM} + 3\sigma_{NM})$. At the extremes of this interval the electron hopping rate in the z direction is reduced roughly by 99.3% of its initial value, thereby drastically reducing conduction. If ΔE_{NM} belongs to the interval $(-\lambda_{NM} - 3\sigma_{NM}, -\lambda_{NM} + 3\sigma_{NM})$ we will express it as

$$\Delta E_{NM} \simeq -\lambda_{NM}. \tag{22}$$

or even

$$\Delta E_{NM} + \lambda_{NM} \simeq 0. \tag{23}$$

The symbol $\simeq$ assumes thus a precise meaning in this context, it is used to specify that the first member of the equation belongs to the interval identified by the second member $\pm 3\sigma_{NM}$. In this way we will obtain shorter formulas, but the reader should keep in mind that the standard deviation σ_{NM} is implicit in the notation used and it depends on the
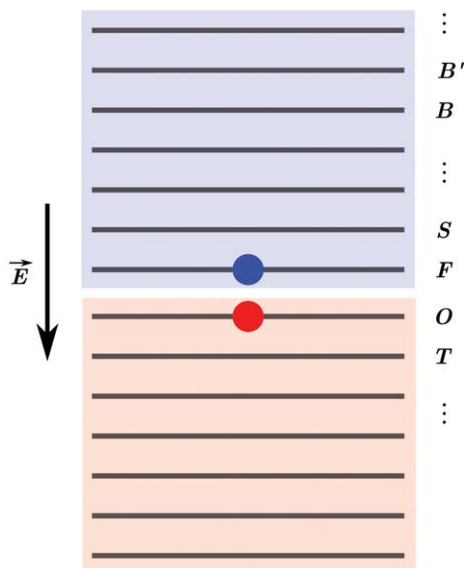
Fig. 7 Illustration of the layers of the system. At the interface, for the acceptor (in dark gray [blue]), we note the first (F) and second (S) layer. For the donor (in light gray [red]), we note layer minus one (O) and minus two (T). Two consecutive bulk layers are noted B and B'. [colors online]

particular hopping considered. The fields maintaining the energy difference inside the interval specified by eqn (23) favor conduction in Marcus theory.

Let us now consider an electron and a hole forming a CT state at a flat interface in the xy plane, with the electric field applied in the $-z$ direction (Fig. 7). We focus first on the movement of the electron, keeping the hole still. Two interesting situations can be analyzed: when the electron is in its initial position at the interface and when the electron is far from the interface. When the electron is at the interface, the difference in energy for the electron hop from the first to the second layer of molecules in the opposite direction of the field is

$$\Delta E^{e}_{F_0 S_0} = \Delta E^{e\text{-int}}_{F_0 S_0} + \Delta E^{e\text{-h}}_{F_0 S_0} - eE\Delta z^{A}_{FS}, \tag{24}$$

where $\Delta E^{e\text{-int}}_{F_0 S_0}$ can be regarded as the contribution of the electron–interface interactions in passing from the first to the second layer (Fig. 7), and the 0 at the subscript specifies that we are considering the ground state. $\Delta E^{e\text{-int}}_{F_0 S_0}$ is composed by the shift of the energy levels of the molecules due to their geometrical deformation at the interface ($\Delta\epsilon$), and by the electrostatic interaction of the interface with the charge carrier due to permanent atomic charges (ΔE^{perm}) and induced dipoles (ΔE^{ind}), see terms in eqn (3). $\Delta E^{e\text{-h}}_{F_0 S_0}$ is the change in energy due to the Coulombic interaction of the hole and the last term in eqn (24) is the contribution of the field, with Δz^{A}_{FS} being the distance between first and second layer of

the acceptor in the z direction (that of course does not depend on which molecular energy levels we are considering). Using eqn (24) we can write the Marcus rate for an electron passing from first to second layer as a function of the external electric field

$$
\begin{aligned}
w^{e}_{F_0 S_0} = \frac{2\pi}{\hbar} \, |H^{e}_{F_0 S_0}|^2 \sqrt{\frac{1}{4\pi k_B T \lambda^{e}_{F_0 S_0}}} \\
\times \exp\left(-\frac{(-eE\Delta z^{A}_{FS} + \Delta E^{e\text{-int}}_{F_0 S_0} + \Delta E^{e\text{-h}}_{F_0 S_0} + \lambda^{e}_{F_0 S_0})^2}{4\lambda^{e}_{F_0 S_0} k_B T}\right).
\end{aligned}
\tag{25}
$$

This is still a Gaussian but with standard deviation

$$
\sigma^{e}_{F_0 S_0} = \frac{\sqrt{2 k_B T \lambda^{e}_{F_0 S_0}}}{e\Delta z^{A}_{FS}}.
\tag{26}
$$

Eqn (25) represents the efficiency of the process considered (the electron splitting the CT state by passing from first to second layer) as a function of the external electric field. With the symbol defined in eqn (23), this can be written concisely as

$$
-eE\Delta z^{A}_{FS} \simeq -(\Delta E^{e\text{-int}}_{F_0 S_0} + \Delta E^{e\text{-h}}_{F_0 S_0}) - \lambda^{e}_{F_0 S_0}.
\tag{27}
$$

When the electron instead is far from the interface, the energy difference of a hop between two layers in the z direction can be simplified as

$$
\Delta E^{e}_{B_0 B'_0} = \Delta E^{e}_{\text{bulk}} \approx -eE\Delta z^{A}_{\text{bulk}},
\tag{28}
$$

representing the difference in energy between the ground states of two consecutive layers in the bulk B and B'. The average distance of two layers in the bulk in the z direction $\Delta z^{A}_{\text{bulk}}$ is in general different from Δz^{A}_{FS} due to the deformations at the interface. The interface will indeed usually introduce an extra stress on the molecule leading to an increased compression of the interfacial layers (for example see ref. 61). In this case, the hole is too far to interact with the electron, and the energy shifts due to molecular distortions ($\Delta\epsilon$), permanent charges (ΔE^{perm}) and induced dipoles (ΔE^{ind}) will be neglected since they average to 0 in the bulk. The Marcus rate for a free electron moving in the bulk is

$$
w^{e}_{\text{bulk}} = \frac{2\pi}{\hbar} \, |H^{e}_{\text{bulk}}|^2 \sqrt{\frac{1}{4\pi k_B T \lambda^{e}_{\text{bulk}}}} \exp\left(-\frac{(-eE\Delta z^{A}_{\text{bulk}} + \lambda^{e}_{\text{bulk}})^2}{4\lambda^{e}_{\text{bulk}} k_B T}\right)
\tag{29}
$$

whose standard deviation is

$$
\sigma^{e}_{\text{bulk}} = \frac{\sqrt{2 k_B T \lambda^{e}_{\text{bulk}}}}{e\Delta z^{A}_{\text{bulk}}}.
\tag{30}
$$

For the electron movement then we can finally obtain the pair of equations

$$-eE\Delta z_{FS}^{A} \simeq -(\Delta E_{F_0 S_0}^{\text{e-int}} + \Delta E_{F_0 S_0}^{\text{e-h}}) - \lambda_{F_0 S_0}^{\text{e}} \qquad (31)$$

$$-eE\Delta z_{\text{bulk}}^{A} \simeq -\lambda_{\text{bulk}}^{\text{e}}; \qquad (32)$$

these represent the system in the two situations when the electron is near to or far from the interface. Eqn (31) and (32) specifies the conditions for the field to obtain maximum conduction in the z direction, when the electron is near to and far from the interface, respectively. Eqn (31) is related to the efficiency of the CT state splitting initiated by the electron, while eqn (32) is related to the transport of the free electron to the electrode. If these two equations can be satisfied simultaneously, we will obtain both CT state dissociation (initiated by the electron) and free electron transport in the acceptor phase. Figure 8 illustrates these facts in what we will call the efficiency diagram for the electron. The 3σ tolerances are illustrated both on the energy and on the field axis.

Analogous considerations can be made for the hole, leading to other two equations regarding the efficient CT state splitting initiated by the hole and the transport of the free hole to the electrode

$$-eE\Delta z_{OT}^{D} \simeq -(\Delta E_{O_0 T_0}^{\text{h-int}} + \Delta E_{O_0 T_0}^{\text{h-e}}) - \lambda_{O_0 T_0}^{\text{h}} \qquad (33)$$

$$-eE\Delta z_{\text{bulk}}^{D} \simeq -\lambda_{\text{bulk}}^{\text{h}}; \qquad (34)$$
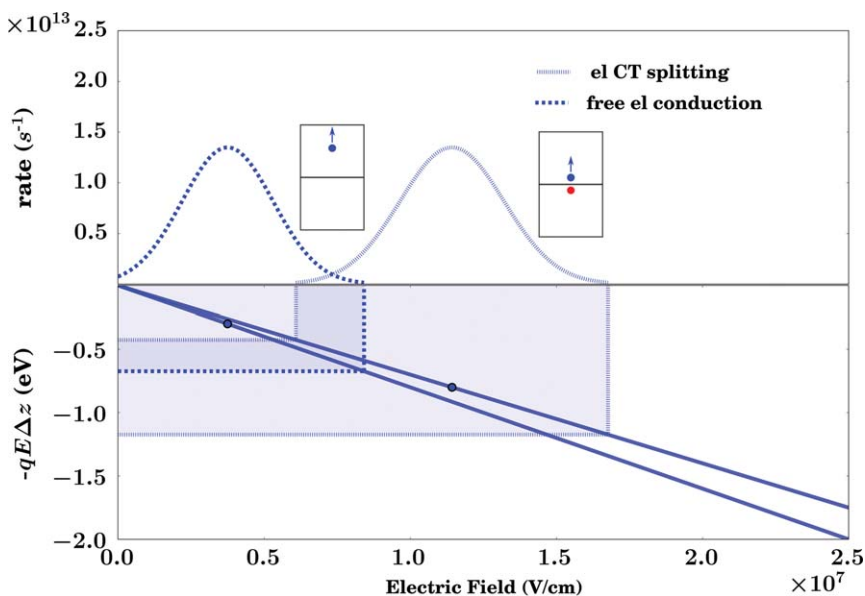

Fig. 8   Efficiency diagram for the electron (for conduction close to and away from the interface). [colors online]

where O and T represents the first and second layer for the hole (Fig. 7). Due to the fact that h is present in the superscript, now the subscript 0 in O_0 and T_0 refers to the ground state of the hole instead. The conditions imposed by eqn (33–34) can be illustrated in the efficiency diagram for the hole, Fig. 9. To have both an efficient CT state splitting and free charge carriers collection at the electrodes the field should respect eqn (31–34). This results in a combination of Figs. 8 and 9 in what will be called a CT state splitting diagram, Fig. 10. The CT state splitting diagram plots together the electric field strength, the energy of the field gained in jumping to the next layer and the efficiencies of the 4 processes of interest in eqn (31–34). From the CT state splitting diagram, it is possible to tell, before performing the KMC simulations, if two materials are suitable to be used together in an organic solar cell and which electric field strengths are required for the simulations to work. Also we can spot immediately from the hypothetical CT state splitting digram of Fig. 10, that the hole is more mobile than the electron. Ideally we would like to have the 4 Gaussians of a CT state splitting diagram to overlap as much as possible. This means that there is an electric field at which the 4 processes are efficient. Suboptimal situations are also possible, leading sometimes to the same efficient CT state splitting. When reading a CT state splitting diagram indeed it is mandatory to keep in mind that the two Gaussians corresponding to the splitting of the CT state (first to second layer) somehow represent the state of the system in the initial situation. We can think that the state of the system at every instant of time is represented by two Gaussians (since only two charges are present). As soon as the system starts to evolve, if the CT state splits and one of the two charge carriers starts moving, the two Gaussians representing
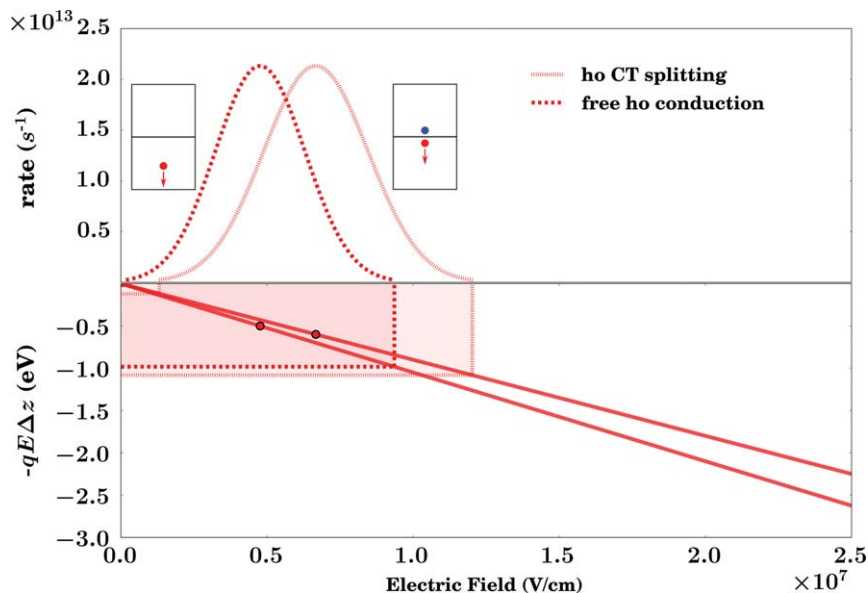

Fig. 9 Efficiency diagram for the hole (for conduction close to and away from the interface). [colors online]
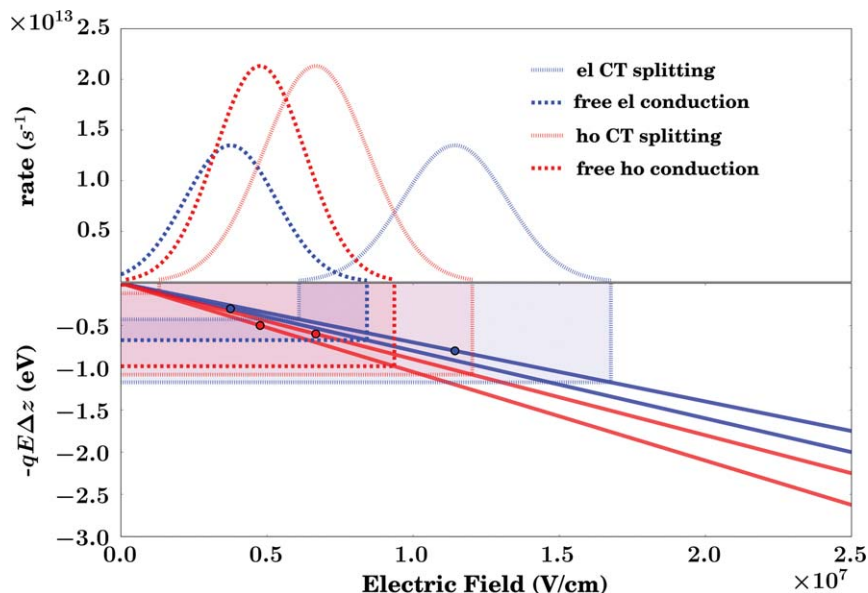
Fig. 10 CT state splitting diagram. [colors online]

the state of the system will be in an intermediate situation, between charge at the interface and charge in the bulk. In the example of Fig. 10 then, the field that is expected to work best for both CT state splitting and free charge carrier conduction is around 0.5×10^7 V cm^{-1}. At this field indeed, the CT state splitting will always be initiated by the hole, but as soon as the hole moves away from the interface, the Gaussian of the electron will approach the Gaussian of the free conduction and also the electron will start moving efficiently.

The main question that can arise now is which are the factors that keep the Gaussians separated. Let us first focus on the two Gaussians of the same charge carrier, for this task we have to look at eqn (31) and (32). Usually the difference between the Δz near to and far from the interface is not predominant. Also, if considering donor and acceptor materials with similar polarizabilities, we can argue that the external reorganization energies do not change so much at the interface or in the bulk. Under these conditions the main factor shifting the Gaussian of the free conduction with respect to the Gaussian of the CT state splitting is the energy difference in passing from first to second layer without field contribution (noF), *i.e.* $\Delta E_{F_0 S_0}^{\text{e-noF}}$ and $\Delta E_{F_0 S_0}^{\text{h-noF}}$ for electron and hole, respectively

$$\Delta E_{F_0 S_0}^{\text{e-noF}} = -(\Delta E_{F_0 S_0}^{\text{e-int}} + \Delta E_{F_0 S_0}^{\text{e-h}}) \tag{35}$$

$$\Delta E_{O_0 T_0}^{\text{h-noF}} = -(\Delta E_{O_0 T_0}^{\text{h-int}} + \Delta E_{O_0 T_0}^{\text{h-e}}). \tag{36}$$

In the initial CT state, the charge carriers will feel two contrasting forces, for example the electron will feel the attraction of the hole $\Delta E_{F_0 S_0}^{\text{e-h}}$ and the contribution of the interface $\Delta E_{F_0 S_0}^{\text{e-int}}$. To have the two Gaussians as much

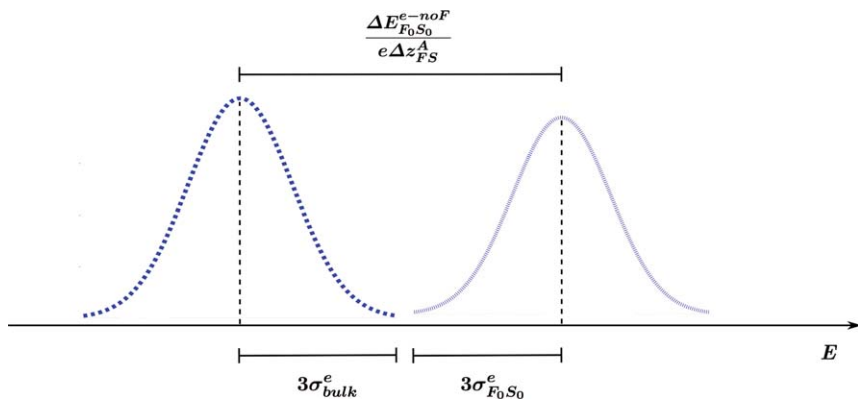**Fig. 11** Illustration of the two gaussians for electron in the CT state and free electron conduction. The factors promoting or impeding the overlap are noted on the figure. [colors online]

overlapped as possible the interface should have a destabilization contribution balancing the hole attraction (Fig. 11)

$$\frac{|\Delta E_{F_0 S_0}^{\mathrm{e-noF}}|}{\mathrm{e}\Delta z_{FS}^{\mathrm{A}}} \lesssim 3(\sigma_{F_0 S_0}^{\mathrm{e}} + \sigma_{\mathrm{bulk}}^{\mathrm{e}}). \tag{37}$$

This agrees somehow with the intuition that the interface should destabilize "sufficiently" the CT and thus helps the free charge carriers formation. Eqn (37) clarifies how much is "sufficiently". Polarization helps for two reasons. It introduces a screening of the Coulombic attraction between electron and hole, and increases the destabilization of the interface, thus reducing $\Delta E_{F_0 S_0}^{\mathrm{e-noF}}$. Secondly it increases the reorganization energy, adding the external reorganization contribution and thus increasing the Gaussians standard deviations $\sigma_{F_0 S_0}^{\mathrm{e}}$ and $\sigma_{\mathrm{bulk}}^{\mathrm{e}}$. The reasoning for the hole is analogous, leading to the condition

$$\frac{|\Delta E_{O_0 T_0}^{\mathrm{h-noF}}|}{\mathrm{e}\Delta z_{OT}^{\mathrm{D}}} \lesssim 3(\sigma_{O_0 T_0}^{\mathrm{h}} + \sigma_{\mathrm{bulk}}^{\mathrm{h}}). \tag{38}$$

Another question is if the two materials selected are compatible. Assuming that either one of eqn (37) or (38) are respected, we have established that it is possible to split the CT state and conduct at least one of the free charge carriers. To assure that both the charge carriers can be collected we have to analyze the factors that favor the overlap between the free electron conduction and the free hole conduction Gaussians. For this task we have to compare eqn (32) and (34). It turns out that we would like the ratio between reorganization energy and layer spacing in the direction of the field to be quite similar, meaning

$$\left| \frac{\lambda_{\mathrm{bulk}}^{\mathrm{e}}}{\mathrm{e}\Delta z_{\mathrm{bulk}}^{\mathrm{A}}} - \frac{\lambda_{\mathrm{bulk}}^{\mathrm{h}}}{\mathrm{e}\Delta z_{\mathrm{bulk}}^{\mathrm{D}}} \right| \lesssim 3(\sigma_{\mathrm{bulk}}^{\mathrm{e}} + \sigma_{\mathrm{bulk}}^{\mathrm{h}}). \tag{39}$$

If this condition is respected, it exists a field at which both electrons and holes can move. In combination with the fact that eqn (37) or (38) is respected we have just found two materials for which it is possible to find a field at which the CT state splits and both the free charge carriers can be collected at the electrodes.

The analysis done so far focuses on the ground CT state, in which the electron is always on the LUMO of an acceptor molecule and the hole is always on the HOMO of a donor molecule. There is a quite large consensus in the literature on the importance of the role of hot charge carriers for an efficient functioning of organic solar cells.[62–65] This can also be rationalized with the CT state splitting diagram just presented. Introducing excited levels for the electron and the holes introduces new Gaussians in the diagram and thus creates more possibilities for overlaps and complex path mixing of decay and transport. We can argue that in the bulk the main transport is given by the ground levels, thus we can assume that the free conduction Gaussians will not change (at least not significantly). The main effect of the excited molecular levels is expected at the interface in the hot CT states. The transfer integrals $H_{F_m S_n}$ have an oscillatory behaviour but are expected to increase with increasing m and n due to the increasing delocalization of the molecular orbitals. This will result in Gaussians with higher amplitude. Neglecting the change of the reorganization contribution $(\lambda_{F_m S_n})$, we focus again on the energy difference $(\Delta E_{F_m S_n}^{noF})$, that is the main factor determining the overlap. We had seen in eqn (37) and (38) which are the conditions for an overlap between the Gaussians of the charge carrier in the ground CT state and the free conduction Gaussian. Analogously, the condition for the overlap between the Gaussians of the charge carriers in the hot CT state and the free conduction is

$$\frac{|\Delta E_{F_m S_n}^{e-noF}|}{e \Delta z_{FS}^A} \lesssim 3 \left(\sigma_{F_m S_n}^e + \sigma_{bulk}^e \right) \tag{40}$$

$$\frac{|\Delta E_{O_m T_n}^{h-noF}|}{e \Delta z_{OT}^D} \lesssim 3 \left(\sigma_{O_m T_n}^h + \sigma_{bulk}^h \right). \tag{41}$$

For different values of $\Delta E_{F_m S_n}^{e-noF}$, we will have hot CT Gaussians centered at different positions on the field axis, determining the possibility of the overlap (analogous to the ground CT case of Fig. 11). It will thus be possible to find an energy difference $\Delta E_{F_m S_n}^{e-noF}$ for which the CT splitting Gaussian from F_m to S_n is overlapping with the free conduction Gaussian in the bulk, or with another CT splitting Gaussian that in turn overlap with the free conduction Gaussian in the bulk. In this way it will be possible to establish, at a particular field, which is the most probable chain of events leading to the free charge carriers conduction. If it is possible to find one.

Which chain of events (the system passing from one hot CT state to another) is more favourable depends on the specific case. The definition of $\Delta E_{F_m S_n}^{\text{e-noF}}$ is analogous to the ground state case (eqn (35))

$$\Delta E_{F_m S_n}^{\text{e-noF}} = -(\Delta E_{F_m S_n}^{\text{e-int}} + \Delta E_{F_m S_n}^{\text{e-h}}) \tag{42}$$

i.e. the average energy difference in passing from the m orbital level in the first layer F to the n orbital level in the second layer S. Higher level wavefunctions are more delocalized and can produce a different momentum in the charge distribution on the molecule (dipoles, quadrupoles, *etc.*). These represent higher order corrections to the electrostatic energies and we can argue that in first approximation the electron–hole interactions do not vary significantly from the ground case, *i.e.* $\Delta E_{F_m S_n}^{\text{e-h}} \approx \Delta E_{F_0 S_0}^{\text{e-h}}$. Similar reasoning can be done for the interaction of the charge carrier with the surrounding in term of permanent charges (ΔE^{perm}) and induced dipoles (ΔE^{ind}). Consequently, the contribution of the interface will mainly differ caused by the orbital energy difference $\Delta\epsilon$. In first approximation the differences in energy without the field can be written as

$$\Delta E_{F_m S_n}^{\text{e-noF}} \approx \Delta E_{F_0 S_0}^{\text{e-noF}} - \Delta\epsilon_{F_0 S_0}^{\text{e}} + \Delta\epsilon_{F_m S_n}^{\text{e}} \tag{43}$$

$$\Delta E_{O_m T_n}^{\text{h-noF}} \approx \Delta E_{O_0 T_0}^{\text{h-noF}} - \Delta\epsilon_{O_0 T_0}^{\text{h}} + \Delta\epsilon_{O_m T_n}^{\text{h}}. \tag{44}$$

An analogous argument can justify somehow the neglecting of the change in reorganization energy, leading to $\sigma_{F_m S_n}^{\text{e}} = \sigma_{F_0 S_0}^{\text{e}}$ and to the simplified set of equations for the hot CT states

$$\frac{|\Delta E_{F_0 S_0}^{\text{e-noF}} - \Delta\epsilon_{F_0 S_0}^{\text{e}} + \Delta\epsilon_{F_m S_n}^{\text{e}}|}{e\Delta z_{FS}^{A}} \lesssim 3(\sigma_{F_0 S_0}^{\text{e}} + \sigma_{\text{bulk}}^{\text{e}}) \tag{45}$$

$$\frac{|\Delta E_{O_0 T_0}^{\text{h-noF}} - \Delta\epsilon_{O_0 T_0}^{\text{h}} + \Delta\epsilon_{O_m T_n}^{\text{h}}|}{e\Delta z_{OT}^{D}} \lesssim 3(\sigma_{O_0 T_0}^{\text{h}} + \sigma_{\text{bulk}}^{\text{h}}). \tag{46}$$

Selecting appropriately the initial and final levels of the hop m and n, it is possible to obtain $\Delta\epsilon_{F_m S_n}^{\text{e}}$ and/or a $\Delta\epsilon_{O_m T_n}^{\text{h}}$ respecting the eqn (45–46) for a given electric field. The analysis done so far is focused on the hopping rates of the electron or the hole composing the hot CT. If trying to determine the most probable path for the evolution of the system, also the pure decays of a single charge carrier from higher to lower level on the same layer have to be considered. For example for electron on the first layer, decaying rates $F_m \rightarrow F_p$ with $p < m$ will compete with hopping rates $F_m \rightarrow S_n$. Eqn (45–46) have been obtained neglecting higher order terms that could lead to slightly different results, but the message of the equations is still valid: higher levels help the CT state splitting and the overlap of the CT splitting Gaussians with the free conduction Gaussians. Also other paths can be thought, complicating further the CT state splitting diagram, for example considering intermediate steps in the CT state dissociation, *i.e.* $F_m \rightarrow S_n$ followed by $S_m \rightarrow R_p$, where R represents here a consecutive layer.

5   Conclusions and perspectives

We have presented in this book chapter the steps required to study charge transport in organic solar cells. Starting from a realistic morphology using Ising model and MD simulations, Kinetic Monte Carlo simulations based on Marcus formula can thus be performed on the obtained geometry. The parameters entering in the Marcus formula can be calculated with a combination of semi-empirical, QM and classical methods. A detailed analysis of those parameters allows the construction of the CT state splitting diagram, to determine the qualitative behavior of the materials and the electric fields which allow simultaneously the CT state splitting and the conduction of the free charge carriers in the bulk. This allows to select carefully the electric fields at which the KMC simulations should be performed. With the progress of supercomputing, it is becoming more and more feasible to screen a large variety of donor and acceptor systems and interpenetrating networks. However, the
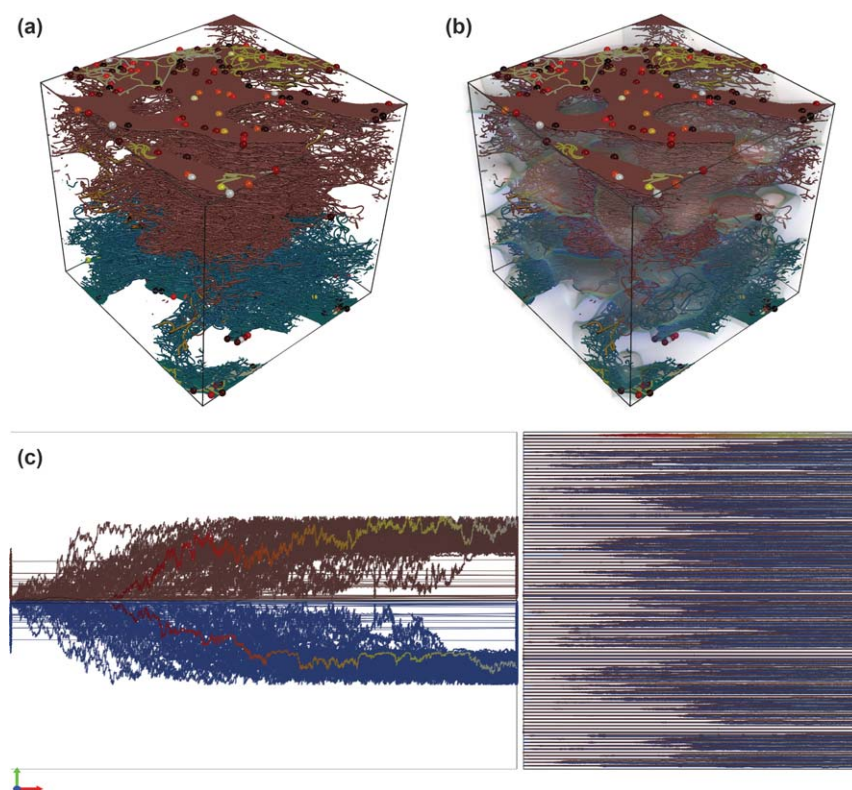
Fig. 12   Visualisation of a set of 100 trajectories of a CT state splitting in a BHJ solar cell. (a) Ensemble of trajectories rendered as flow lines with tube rendering technique using OpenGL shaders. The trajectory lines are sub-sampled and filtered using simple Gaussian smoothing. (b) Same trajectory ensembles are embedded in the volumetric representation of the acceptor and donor material grid. The thin transparent layers in green indicates the boundary between acceptor and donor materials. Direct volume rendering technique is used for this visualization. (c) Distance from the starting point (*y* axis) in function of time (*x* axis). The left panel shows all trajectories imposed over each other, while the right panel has split view of all individual plots. [colors online]

analysis of the KMC simulations is not always trivial, especially for non planar interfaces (*e.g.* BHJ). For these reasons we are currently collaborating with the Visualization center in Norrköping to develop a visualization tool. Figure 12 shows a set of 100 trajectories, where trajectories for the holes in the donor part and for the electrons in the acceptor part are colored in light gray [red] and dark gray [blue], respectively. [colors online] The spheres represent the end points of the trajectories and their colors indicate how long it took for the charge carrier, initially in the CT state, to arrive at this point (light for short time and dark for long time). It is also possible to highlight a specific trajectory in the different visualization panel in order to identify preferential paths in the BHJ morphology paving the way to establishing structure–property relationship.

Acknowledgements

References

1   National Renewable Energy Laboratory: Best Research-Cell Efficiencies, accessed 2016-05-11. URL http://www.nrel.gov/ncpv/images/efficiency_chart.jpg.
2   M. A. Green, K. Emery, Y. Hishikawa, W. Warta and E. D. Dunlop, *Progr. Photovoltaics*, 2016, **24**, 3.
3   M. Jørgensen, K. Norrman, F. C. Krebs, M. Jorgensen, K. Norrman and F. C. Krebs, *Sol. Energy Mater. Sol. Cells*, 2008, **92**, 686.
4   S. Savagatrup, A. D. Printz, T. F. O'Connor, A. V. Zaretski, D. Rodriquez, E. J. Sawyer, K. M. Rajan, R. I. Acosta, S. E. Root and D. J. Lipomi, *Energy Environ. Sci.*, 2014, **8**, 55.
5   P. Cheng and X. Zhan, *Chem. Soc. Rev.*, 2016, **45**, 2544.
6   K. Vandewal, K. Tvingestedt and O. Inganäs, *Semicond. Semimetals*, 2011, **85**, 261.
7   J.-L. Brédas, J. E. Norton, J. Cornil and V. Coropceanu, *Acc. Chem. Res.*, 2009, **42**, 1691.
8   C. Deibel, T. Strobel and V. Dyakonov, *Adv. Mater.*, 2010, **22**, 4097.
9   E. von Hauff, *Semicond. Semimetals*, 2011, **85**, 231.
10 D. Beljonne, J. Cornil, L. Muccioli, C. Zannoni, J. L. Brédas and F. Castet, *Chem. Mater.*, 2011, **23**, 591.
11 J. Cornil, S. Verlaak, N. Martinelli, A. Mityashin, Y. Olivier, T. Van Regemorter, G. D'Avino, L. Muccioli, C. Zannoni, F. Castet, D. Beljonne and P. Heremans, *Acc. Chem. Res.*, 2013, **46**, 434.
12 O. V. Mikhnenko, P. W. M. Blom and T.-Q. Nguyen, *Energy Environ. Sci.*, 2015, **8**, 1867.
13 L. J. A. Koster, E. C. P. Smits, V. D. Mihailetchi and P. W. M. Blom, *Phys. Rev. B*, 2005, **72**, 085205.
14 G. a. Buxton and N. Clarke, *Modell. Simul. Mater. Sci. Eng.*, 2007, **15**, 13.
15 M. Gruber, B. Stickler, G. Trimmel, F. Schürrer and K. Zojer, *Org. Electron.*, 2010, **11**, 1999.

16 C. de Falco, M. Porro, R. Sacco and M. Verri, *Comput. Methods Appl. Mech. Eng.*, 2012, **245–246**, 102, , arXiv:1206.1440v3.

17 H. Bässler, *Phys. Status Solidi B*, 1993, **175**, 15.

18 P. K. Watkins, A. B. Walker and G. L. B. Verschoor, *Nano lett.*, 2005, **5**, 1814.

19 J. Kirkpatrick, V. Marcon, J. Nelson, K. Kremer and D. Andrienko, *Phys. Rev. Lett.*, 2007, **98**, 227402.

20 Y. Olivier, L. Muccioli, V. Lemaur, Y. H. Geerts, C. Zannoni and J. Cornil, *J. Phys. Chem. B*, 2009, **113**, 14102.

21 J. Idé, R. Méreau, L. Ducasse and F. Castet, *J. Phys. Chem. B*, 2011, **115**, 5593.

22 C. Poelking, E. Cho, A. Malafeev, V. Ivanov, K. Kremer, C. Risko, J. L. Brédas and D. Andrienko, *J. Phys. Chem. C*, 2013, **117**, 1633.

23 C. Deibel, T. Strobel and V. Dyakonov, *Phys. Rev. Lett.*, 2009, **103**, 036402.

24 T. Vehoff, B. Baumeier, A. Troisi and D. Andrienko, *J. Am. Chem. Soc.*, 2010, **132**, 11702.

25 V. Rühle, J. Kirkpatrick and D. Andrienko, *J. Chem. Phys.*, 2010, **132**, 134103.

26 V. Rühle, A. Lukyanov, F. May, M. Schrader, T. Vehoff, J. Kirkpatrick, B. Bauemeier and D. Andrienko, *J. Chem. Theory Comput.*, 2011, **7**, 3335.

27 M. Jakobsson, M. Linares and S. Stafström, *J. Chem. Phys.*, 2012, **137**, 114901.

28 R. Volpi, S. Kottravel, M. S. Nørby, S. Stafström and M. Linares, *J. Chem. Theory Comput.*, 2016, 812.

29 C. W. Tang, *Appl. Phys. Lett.*, 1986, **48**, 183.

30 N. S. Sariciftci, L. Smilowitz, A. J. Heeger and F. Wudi, *Science*, 1992, **258**, 1474.

31 A. Miller and E. Abrahams, *Phys. Rev.*, 1960, **120**, 745.

32 R. A. Marcus, *J. Chem. Phys.*, 1956, **24**, 966.

33 Y. Yi, V. Coropceanu and J.-l. Brédas, *J. Am. Chem. Soc.*, 2009, **131**, 15777.

34 Y. Yi, V. Coropceanu and J.-L. Brédas, *J. Mater. Chem.*, 2011, **21**, 1479.

35 Y. T. Fu, D. A. Da Silva Filho, G. Sini, A. M. Asiri, S. G. Aziz, C. Risko and J. L. Brédas, *Adv. Funct. Mater.*, 2014, **24**, 3790.

36 Y. Olivier, V. Lemaur, J. L. Brédas and J. Cornil, *J. Phys. Chem. A*, 2006, **110**, 6356.

37 N. G. Martinelli, M. Savini, L. Muccioli, Y. Olivier, F. Castet, C. Zannoni, D. Beljonne and J. Cornil, *Adv. Funct. Mater.*, 2009, **19**, 3254.

38 S. Verlaak, D. Beljonne, D. Cheyns, C. Rolin, M. Linares, F. Castet, J. Cornil and P. Heremans, *Adv. Funct. Mater.*, 2009, **19**, 3809.

39 M. Linares, D. Beljonne, J. Cornil, K. Lancaster, J.-L. Brédas, S. Verlaak, A. Mityashin, P. Heremans, A. Fuchs, C. Lennartz, J. Idé, R. Méreau, P. Aurel, L. Ducasse and F. Castet, *J. Phys. Chem. C*, 2010, **114**, 3215.

40 L. Muccioli, G. D'Avino and C. Zannoni, *Adv. Mater.*, 2011, **23**, 4532.

41 M. C. Heiber and A. Dhinojwala, *Phys. Rev. Appl.*, 2014, **2**, 014008.

42 S. Novikov and A. Vannikov, *J. Exp. Theor. Phys.*, 1994, **79**, 482.

43 S. Novikov and A. Vannikov, *J. Phys. Chem.*, 1995, **99**, 14573.

44 R. Volpi, S. Stafström and M. Linares, *J. Chem. Phys.*, 2015, **142**, 094503.

45 P. T. van Duijnen and M. Swart, *J. Phys. Chem. A*, 1998, **102**, 2399.

46 J. Wang, P. Cieplak, J. Li, T. Hou, R. Luo and Y. Duan, *J. Phys. Chem. B*, 2011, **115**, 3091.

47 B. Thole, *Chem. Phys.*, 1981, **59**, 341.

48 L. Gagliardi, R. Lindh and G. Karlstrom, *J. Chem. Phys.*, 2004, **121**, 4494.

49 R. Mulliken, C. Rieke, D. Orloff and H. Orloff, *J. Chem. Phys.*, 1949, **17**, 1248.

50 A. Hansson and S. Stafström, *Phys. Rev. B*, 2003, **67**, 075406.

51 J.-L. Brédas, D. Beljonne, V. Coropceanu and J. Cornil, *Chem. Rev.*, 2004, **104**, 4971.

52   J. Kirkpatrick, *Int. J. Quantum Chem.*, 2008, **108**, 51.

53   B. Baumeier, J. Kirkpatrick and D. Andrienko, *Phys. Chem. Chem. Phys.*, 2010, **12**, 11103.

54   P. Petelenz, *Chem. Phys. Lett.*, 1984, **103**, 369.

55   M. N. Bussac, J. D. Picon and L. Zuppiroli, *Europhys. Lett.*, 2004, **66**(392), 0306353.

56   K. Hannewald, V. Stojanović, J. Schellekens, P. Bobbert, G. Kresse and J. Hafner, *Phys. Rev. B*, 2004, **69**, 075211.

57   K. Hannewald and P. a. Bobbert, *Appl. Phys. Lett.*, 2004, **85**, 1535.

58   F. Ortmann, F. Bechstedt and K. Hannewald, *Phys. Status Solidi B*, 2011, **248**, 511.

59   F. Castet, P. Aurel, A. Fritsch, L. Ducasse, D. Liotard, M. Linares, J. Cornil and D. Beljonne, *Phys. Rev. B*, 2008, **77**, 115210.

60   J. C. Sancho-García, G. Horowitz, J. L. Brédas and J. Cornil, *J. Chem. Phys.*, 2003, **119**, 12563.

61   D. P. McMahon, D. L. Cheung and A. Troisi, *J. Phys. Chem. Lett.*, 2011, **2**, 2737.

62   X. Zhu, Q. Yang and M. Muntwiler, *Acc. Chem. Res.*, 2009, **42**, 1779.

63   J. Lee, K. Vandewal, S. R. Yost, M. E. Bahlke, L. Goris, M. A. Baldo, J. V. Manca and T. Van Voorhis, *J. Am. Chem. Soc.*, 2010, **132**, 11878.

64   S. L. Smith and A. W. Chin, *Phys. Chem. Chem. Phys.*, 2014, **16**, 20305.

65   A. Melianas, F. Etzold, T. J. Savenije, F. Laquai, O. Inganäs and M. Kemerink, *Nat. Commun.*, 2015, **6**, 8778.

Exploring the electronic structure of nanohybrid materials for their application in solar cell

Sunandan Sarkar,[a] Supriya Saha,[b] Sougata Pal[c†] and Pranab Sarkar*[a]

1 Introduction

The rapid decrease of the fossil fuel and the increasing energy demand forces the researchers to think of alternative energy resources.[1–3] The photovoltaic cell that employs solar energy is found to be a viable alternative for the purpose of meeting energy demand. The solar cells made of Si-based photovoltaic materials are of extensive use for the conversion of solar energy to electricity. However, the Si-based devices suffer from high manufacturing costs and also the photon to current efficiency for such devices reaches only about 16%.[4] The need of the hour is the invention of solar cell devices which are of relatively low cost and at the same time having higher efficiencies. The first successful attempt in this direction was made by O'Regan and Grätzel in the year 1991.[1] Inspired by the principle of natural photosynthesis, they developed a new photovoltaic cell popularly known as dye-sensitized solar cell (DSSC).[5] In DSSC, the light absorption and charge transport are performed by a dye and a semiconductor, respectively, in contrast to traditional Si-based solar cells where the function of both the light absorption and charge transport are accomplished by the single semiconductor. The interest in DSSCs also stems from the fact that, fabrication costs of these devices are much lower as compared to Si-based solar cells. Among the different studied sensitizers, ruthenium based sensitizers are found to show photovoltaic efficiencies more than 11%.[4] However, the problem with the ruthenium is its low natural abundance and more importantly it is environmentally hazardous. This promotes the use of organic dyes as sensitizer. The use of organic dyes is advantageous not only for their low cost but also one can tune their absorption characteristics through suitable functionalization.

During the last two decades, the research in nanochemistry has made a significant progress and nanomaterials have emerged as the new building blocks in solar cell devices. As mentioned by Kamat,[6–18] the

[a]Department of Chemistry, Visva-Bharati, Santiniketan – 731235, India.
E-mail: pranab.sarkar@visva-bharati.ac.in
[b]Newcastle Institute for Energy and Resources, The University of Newcastle, Callaghan 2308, Australia
[c]Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA
[†]Permanent Address: Department of Chemistry, University of Gour Banga, Malda – 732103, India.

nanostructures can be utilized for solar energy conversion in three different ways: the first one is the donor–acceptor assemblies that mimic photosynthesis;[8–11,13] the second one is to make use of semiconductor nanoparticles (NPs) as photocatalysts to produce fuels.[15,16] Finally, the third and the most appealing one is the semiconductor based solar cells where the nanostructure plays the role of an effective charge carrier.[6,14,18]

The utility of semiconductor quantum dots (QDs) for their use in photovoltaic cells primarily depends on the ability of charge transfer from the photoexcited QDs before the recombination of charge carriers. The studies on the individual QDs reveal that the process of the charge transport and charge recombination takes place on an ultrafast time scale so that the charge recombination is a competing event along with the charge transfer and this puts severe limitations in their application in solar cell.[7,10,19] In the recent past, hybrid nanostructures[6,8,13,20–25] with suitable donor–acceptor systems is in the limelight since these can lower the recombination rate and are thus suitable in designing solar cell. The donor–acceptor systems composed of inorganic–organic hybrid nanostructures[23,26–28] are of particular importance to design next generation solar cells where zinc Zn- and Cd-chalcogenide (X = O, S, Se, Te) semiconductor nanocrystals (NCs) are the best choice as inorganic component.[11,17,29–41] This is so, because these possess unique physical properties such as large extinction coefficients, ability of multiple exciton generation and more importantly the size-dependent tunability of the optical and electronic properties. Different carbon nanostructures such as graphene,[42–53] fullerenes[6,54–58] or carbon nanotubes (CNTs)[17,20,39,40,59–67] have been of great use as organic components in the hybrid nanostructures. During the last few years several experimental groups have synthesized different inorganic–organic hybrid nanostructures made of group II–VI (CdS, CdSe, CdTe, ZnO, ZnS) QDs[6,8,38,52,68] or other nanostructures such as nanotubes (NTs) or nanowires (NWs)[29,31,35] and carbon nanostructures such as graphene, CNTs and fullerenes. The ultra-fast charge transfer dynamics in those systems have been studied in detail.[49,69–73] The suggestion that is offered by the experimentalists is that the charge separation, a prerequisite of efficient photovoltaic devices is greatly improved in hybrid nanostructures as compared to that of an individual component.

Numerous efforts from theory are also being devoted to study the excited state electron transfer dynamics.[49,69–78] For successful implementation of the hybrid nanostructures the exploration and understanding of the electronic structure of this class of materials is of primary importance. This is simply because, the electron injection efficiencies and the rate of electron–hole recombination are the most important factors that control the efficiency of a solar cell, and those two factors in turn depend on the position of frontier energy levels of the donor–acceptor hybrid nanostructures.[79–83] For example, the energy gap between the lowest unoccupied molecular orbital (LUMO) of the donor to the conduction band minimum (CBM) of the acceptor dictates the rate of electron transfer (ET) from the donor to the acceptor. On the other hand, it is the spatial charge separation that controls the rate of electron–hole recombination. In this

chapter, we are going to present results of our theoretical studies on different hybrid nanostructures of contemporary interest. But, before going to the details of our theoretical results, let us put a brief review on the experimental studies of this class of materials which motivates us to perform the theoretical studies.

2 Experimental studies, a brief review

In experiment, time resolved spectroscopy is the most viable tool for obtaining the rate of an individual process and determining the relationship between the carrier transport dynamics that govern the solar cell performance.[25] Many experimentalists have investigated the electron transfer dynamics in dye-TiO_2 composite systems.[84,85] Time-resolved laser spectroscopic techniques can monitor the electron injection in real time.[69,72,73,75,76,86,87] The first pioneering work was performed by group of Willig[72,75,76] and gave some insights regarding the mechanism of electron injection. They studied the ET between modified perylene chromophores and TiO_2 and reported that the coherent wave packet created in the donor molecule by the pump pulse survived the ET and it was detected in the molecular cation by the probe pulse. This observation rectified the usual concept of ET where it is assumed that the reaction starts from a thermally equilibrated occupation of the vibrational modes in the donor. Thus, it motivated the real-time theoretical studies of the injection dynamics. The experimental observation of Willig and coworkers demonstrated that on leaving the dye, the electron is not trapped on the surface; rather it enters into the bulk TiO_2.[72] Lian and coworkers[69–71,74] performed numerous studies on the ET dynamics of different ruthenium (Ru) dyes-TiO_2 composites by monitoring infrared (IR) absorption of electrons in the semiconductor conduction band (CB). Their studies conferred that the injection dynamics is biphasic; the ultrafast component arises as a result of injection from a non-thermalized excited state and the slower component corresponds to injection from a relaxed excited state. They found that the slower component depends on excitation wavelength, dye energetics, the pH of the solution, and the solvent. Wachtveitl, Grätzel, and coworkers[77,88] reported the electron injection and cation radical formation of alizarin on a TiO_2 surface with an unprecedented life time of 6 fs which is among the fastest charge-separation reactions reported and studied in real time. Many other systems also exhibit timescales on the order of tens or hundreds of femtoseconds.[71,74,76] By measuring the line width using resonant photoemission spectroscopy, Schnadt and coworkers[89] reported even faster, 3 fs of injection time for bi-isonicotinic acid on TiO_2.

The group of P. V. Kamat has designed various assemblies based on QDs and other mesoscopic substances.[9,10,12] By using bifunctional surface modifiers (SH–R–COOH), they have assembled CdSe QDs onto mesoscopic TiO_2 films and performed femtosecond transient absorption and emission quenching experiments.[12] The experimental results indicate that upon visible light excitation, CdSe QDs inject electrons into TiO_2 nanocrystallites. Electron transfer from the thermally relaxed s-state

occurs over a wide range of rate constant values between 7.3×10^9 and 1.95×10^{11} s^{-1}. The TiO$_2$–CdSe composite, when employed as a photo-anode in a photoelectrochemical cell, exhibits a high (12%) charge carrier generation efficiency. However, significant loss of electrons has occurred due to scattering as well as charge recombination at TiO$_2$/CdSe hybrid interfaces and internal TiO$_2$ grain boundaries.

The factors that limit the overall power conversion efficiency of a QD sensitized solar cell (QDSC) include limited absorption of incident light, slow hole transfer rate, back electron transfer (BET), and low fill factors arising from poor counter electrode performance.[8] CdS NPs anchored on mesoscopic TiO$_2$ films exhibit limited absorption below 500 nm with a poor power conversion efficiency when employed as a photoanode in QDSC. QDSCs supersensitized with a squariand dye (JK-216) shows good stability during illumination with standard global AM 1.5 solar conditions, delivering an improved overall power conversion efficiency of 3.14%. The synergy of combining semiconductor QDs and the dye capable of absorbing in near infra-red (NIR) provides new opportunities to harvest photons from different regions of the solar spectrum.[8]

Very recently, Kaniyankandy et al.[49] have synthesized CdTe QD decorated graphene and studied the ultrafast electron transfer from a photoexcited QD to the graphene surface by femtosecond visible spectroscopy. Their study reveals that graphene induces a drastic charge separation in the composite system as compared to the pure QD. The photoconductivity properties of graphene oxide (GO)–CdTe QD composites with visible illumination have been studied by Kundu et al.[90] using time resolved fluorescence spectrophotometry to understand transport properties of the GO–CdTe QD composite. Cyclic voltammetry has been used to investigate the interaction between reduced GO and CdTe QDs to estimate the band structure parameters.[52] A blue shift in the quasi-particle gap was observed indicating strong electronic coupling between GO and the QDs.

Alivisatos' group reported a hybrid solar cell composed of CdSe nanorods (NRs) and a conjugated polymer, poly(3-hexylthiophene) (P3HT).[19,23,91] Efficient free carrier generation via exciton dissociation at the nanoscale interface depends upon the nanoscale morphology, extrinsic and intrinsic properties of the inorganic NCs, their size, geometry and also on the nature of the passivating ligands.[92] Various ligand modification procedures have been developed for NCs either using organic or inorganic ligands which not only change the QD–QD and QD-polymer distances, but also affects the conductivity, energy levels, stability, and the morphology of casted films.[22,93] Moreover, some functional groups like thiols (–SR) are known to create charge traps in the NCs and quench their photoluminescence.[94,95] CdSe and CdS NRs have been reported to show favourable performance as compared to that of the colloidal QDs because of their large surface area and high aspect ratio, which considerably increase the overall free carrier generation efficiency at the organic–inorganic interfaces.[96] High-aspect-ratio organic materials such as conjugated polymer nanowires (NWs) have been shown to enhance the power conversion efficiencies through creating a nanoscale morphology

favourable for charge separation and transport.[21] Hybrid composites combining CdSe colloidal nanoplatelets (NPLs) and conjugated polymers, bound by different anchoring groups, have been reported to show strong photoluminescence quenching in both the co-integrated polymers and the NPLs. The emission quenching is occurred due to exciton dissociation at the polymer/NPL interfaces having a staggered band alignment. Among different anchors, sulphides lead to the strongest quenching as attributed to their stronger attachment to the NP surfaces. Also, a shorter ligand size of the same anchor group is shown to yield a larger photoluminescence quenching as compared to the larger ligand size. The proposed hybrid materials with exciton dissociating interfaces and homogeneous NP distribution are highly promising for making hybrid solar cells.

In order to mimic the natural photosynthetic process Guldi *et al.* designed a number of artificial synthetic assemblies (photosynthetic reaction centre, PRC) which can efficiently convert solar energy into useful chemical energy.[55,56,61] An important approach to PRC modelling is the covalent linking of a photoexcitable chromophore with an electron acceptor or donor moiety. It should be noted that in these artificial systems the organizing property is the covalent linkage between the redox active centres. Over the past decades, a number of factors have been systematically altered to overcome the difficulties encountered in the early artificial, covalently linked donor–acceptor dyads.[97] The energies of the donor and acceptor molecules have been adjusted by suitable functionalization to increase the rates of the forward electron transfer and to slow down the back electron transfer (BET). Efforts have been given to alter ET rate constants in favourable directions through the tuning of the electronic coupling between donor and acceptor moieties. The most important strategy, however, focuses on the incorporation of secondary electron donor or acceptor moieties into multicomponent arrays. The C_{60} fullerene has been effectively used as a three-dimensional electron acceptor unit in artificial reaction centres. Different strategies are adopted with the aim towards improving charge separation in covalently linked (supermolecular) or non-covalently linked (supramolecular) fullerene containing systems. The key feature in these dyads relies on additional stabilisation of the radical pair such as a gain in aromaticity and planarity of the electron donor, or a dissociation of the charge-separated state. These effects efficiently diminish the BET thereby increasing the device performance.[55]

Design, synthesis, characterization, and potential applications of hybrid materials based on carbon nanotubes (CNTs) and different types of electron donors have been the subject of numerous research.[61] CNTs, in general, and single-wall CNTs (SWNT), in particular, stand out as unique charge acceptors in composite materials.[98,99] The extraordinary electronic, mechanical, and adsorption properties of CNTs have triggered their many possible applications, especially in solar cells and thus, nanohybrids based on CNTs lead to important breakthroughs.[40,62,63,66] SWNTs are one-dimensional NWs that can be metallic or semiconducting depending upon the chiral symmetry. They readily accept electrons, which can then be transported under nearly ideal conditions along the

chiral axis.[100] CNTs have the further advantage of their shear size. Upon receiving the charge, the transport along the axis of the nanometers-long carbon structure reduces the probability of back transfer to the oxidized donor. Metallic nanotubes provide such a phenomenon in great extent.[99] Thus, selection of a proper set of donors can lead to the achievement a very efficient charge separation and very long-lived radical pairs ultimately providing an efficient photovoltaic material. Generally, two strategies are being adopted for composite formation, (i) covalent functionalization[101,102] and (ii) noncovalent functionalization.[103] Both of them can, however, be formed out with molecules or with polymers. Although the covalent approach is very versatile, the noncovalent supramolecular approach is practically simpler where van der Waals and Coulombic interactions are the leading factors for assembly formation. In many cases, charged pyrene derivatives are used to solubilize the SWNT through the formation of π–π interactions. Fast charge separation and slow charge recombination are consistently observed in a variety of CNT-based hybrid nanocomposites that contain porphyrin derivatives as donors that ensure their good photovoltaic performances.[65] CNTs have been used as a support material for the dispersion and stabilization of metal and semiconductor NPs also.[60] CNTs decorated by QDs could act as NWs that promote direct charge transport and efficient charge transfer to QDs, thus making efficient solar cells.[39,59] Ozkan and collaborators reported the formation of heterojunctions between MWNTs and thiol-stabilized ZnS capped CdSe QDs.[64] An assemble of CdS NP and CNT was made by Willner and co-workers to study the photoinduced charge transport property in the presence of triethanolamine (as a sacrificial electron donor) which shows an impressively high photocurrent.[40] The quantum efficiency for the photon-to-electron conversion was estimated to be 25% that could be increased up to 70% by increasing the digestion time of the CNTs whereby simply shortening the CNT length. The results suggest that the length of the CNT plays a major role in the formation of photocurrent and hence quantum efficiency.[60]

3   Theoretical studies, tuning of energy band alignment

Having discussed the recent experimental studies on various aspects of hybrid nanostructures we now turn our attention to the major focus of this chapter, the electronic structure of the hybrid nanostructures. As the system sizes are very large we employ self-consistent charge-density-functional tight-binding (SCC-DFTB) method, an approximate quantum chemistry method which can treat systems with large number of atoms with reasonable accuracy.[104–112] We present the results of our static energy calculation on various hybrid nanostructures *viz.*, CdTeQD–CNT, CdSeQD–C$_{60}$, CdTeQD–graphene, Porphyrin–grapheneQD, GrapheneQD–C$_{60}$, CdTeNT–C$_{60}$, ZnO/ZnSNW–dye, *etc.* The emphasis of our discussion will be on the qualitative understanding of the electron injection efficiencies and electron–hole recombination dynamics through the analysis of electronic energy levels of the nanohybrid systems to identify the suitable materials for solar cell application.

3.1 CdTe quantum dot–CNT hybrid nanostructures

There are many experimental groups who are working on CdTe quantum dot–CNT hybrid systems and their results reveal that the electrostatic binding of CNT and CdTe QDs facilitates the rapid photo-induced ET and thus serves as a potent alternative in fabricating photovoltaic devices.[20,56,67] Drost et al.[20] have reported a novel donor–acceptor inorganic–organic nanohybrid CdTeQD–pyrene/SWCNT and explored the possibilities of using this nanohybrid system for the construction of photovoltaic devices. The interaction between the components plays an important role in determining the efficacy of these particular nanohybrid systems in photovoltaic applications. The basic understanding of ET in QD–CNT hybrid nanostructures, which in turn depends on the position of the frontier energy levels of its component, therefore is of crucial for its effective utilization in designing solar cells. In this section we would like to discuss the electronic structure calculation of CdTeQD–CNT nanohybrid structures. To model the hybrid nanostructure, all the surface Cd atoms except one of the bare CdTe QDs are passivated by thiol groups (–SH). The remaining Cd atom is functionalized with the long chain of thiol molecule, 3-aminopropane-1-thiol (–S–CH$_2$–CH$_2$–CH$_2$–NH$_2$) to form an amide linkage with the CNT, one carbon atom of which has been functionalized by a –COOH group. The support of this kind of –COOH functionalization of the SWCNT came from the experimental work of Zebli and coworkers.[67] However, the charge of the SWCNT is balanced by attaching one –H atom to the neighbouring carbon atom where the –COOH group is attached. The whole system is then treated in a periodic manner to obtain the ground state geometry and related electronic properties.[113] The optimized structure of one representative CNT–CdTe hyrbid nanostructure is shown in Fig. 1.

The dynamics of the electron injection and electron–hole recombination that dictate the photovoltaic efficiencies are controlled by the energy states close to band gap region. To understand the contribution of the two components of hybrid nanostructures to the frontier energy
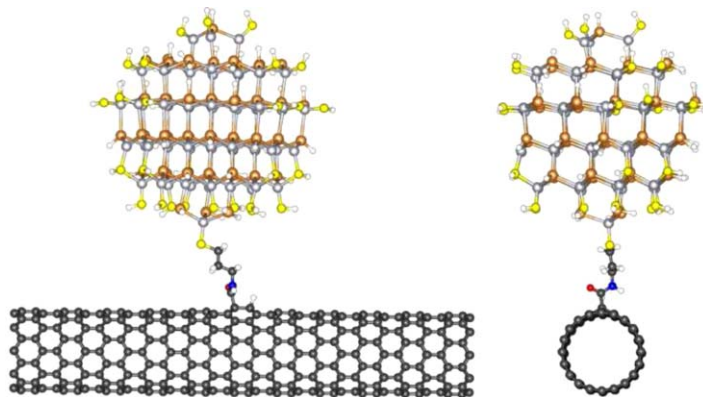

Fig. 1 Optimized structure (both top (right) and side (left) view) of one representative Cd$_{65}$Te$_{65}$ QD–CNT (10, 0) nanohybrid system. (Reprinted with permission from S. Sarkar et al., J. Phys. Chem. C, 2012, **116**, 21601. © 2012, American Chemical Society.)
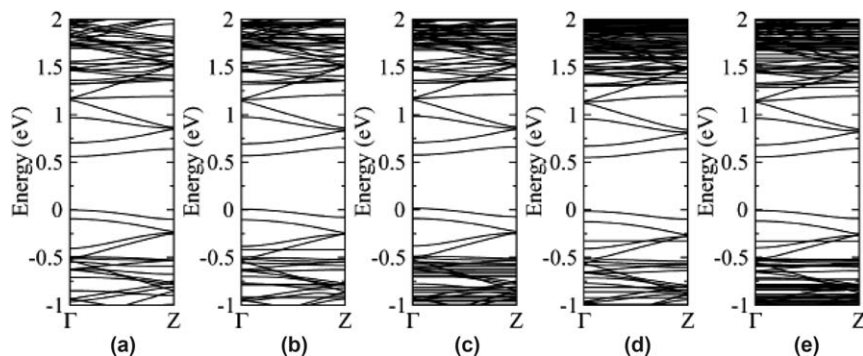
Fig. 2 The band structure of (a) CNT [zigzag (10, 0)] and CdTeQD–CNT [zigzag (10, 0)] nanohybrids with CdTeQDs of different sizes: (b) $Cd_{34}Te_{34}$, (c) $Cd_{55}Te_{55}$, (d) $Cd_{65}Te_{65}$, and (e) $Cd_{98}Te_{98}$, The zero energy is set at the top of the valence band. (Reprinted with permission from S. Sarkar et al., J. Phys. Chem. C, 2012, **116**, 21601. © 2012, American Chemical Society.)

states, we have analyzed both the band structure and the density of states (DOSs) of the CdTeQD–CNT nanohybrids (QDs of four different sizes). From the band structure as shown in Fig. 2, it is clear that the band gap region of the nanohybrids is dominated by the CNT, while the contributions of CdTe QDs in DOSs are in the deep of both valence and conduction band region. So, the band gap absorption of the CdTeQD–CNT nanohybrids will be controlled by the CNT and this result is in good agreement with the experimental observation of Drost et al.[20] Thus we conclude that the onset of photocurrent generation in these particular nanohybrids occurs at the wavelength corresponding to band gap of the CNT.

To have some qualitative idea about the role of the size of the CdTe QD in the electron injection from the QD to the CNT we have shown the DOSs of four CdTeQD–CNT nanocomposite systems with CdTe QDs of different size in Fig. 3. The PDOSs of individual QDs and CNT are also shown in the same figure. One may note an interesting feature from the DOSs figure that the difference between the LUMO of the QD and CB of the CNT decreases with increasing the size of the QDs. Following Marcus theory of ET[114] where the energy difference between the LUMO of donor and the CB of the acceptor is the driving force for electron injection, we could expect faster ET rate for smaller CdTeQD as compared to larger QDs when they are coupled with the CNT. However, the faster ET rates with smaller sized QDs do not necessarily mean that those systems will show better photovoltaic efficiencies since these depend on other factors such as the rate of hole transfer and rate of recombination of charge carriers, etc.[6]

3.2 CdSe quantum dot–fullerene nanocomposites

The excellent exciton mobility, relatively large exciton diffusion length and high electron affinity make fullerenes a good electron-acceptor and electron-transporting material and thus a suitable candidate to be used in nanohybrids for solar cell application.[6,7,57,58,115] The first successful report of the synthesis of a CdSe QD–fullerene nanohybrid is made by
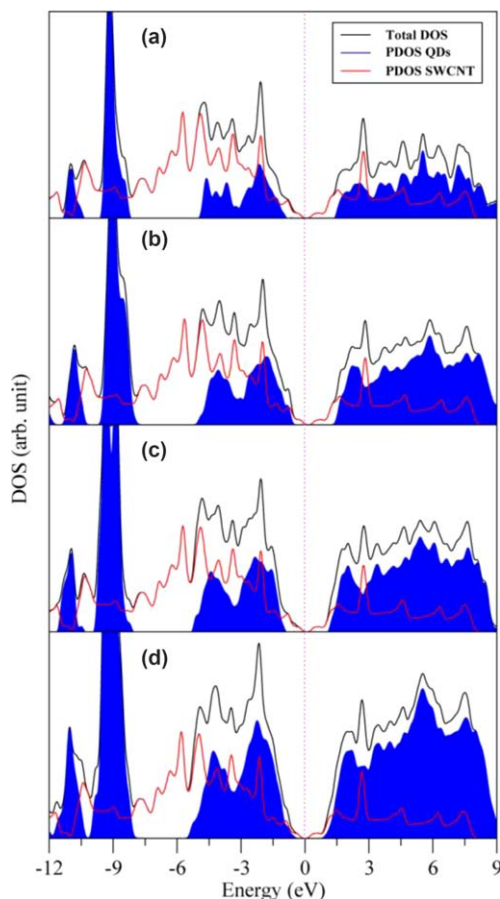
Fig. 3 The density of states (DOS) of CdTeQD–CNT [zigzag (10, 0)] nanohybrids with CdTeQDs of different sizes: (a) $Cd_{34}Te_{34}$, (b) $Cd_{55}Te_{55}$, (c) $Cd_{65}Te_{65}$, and (d) $Cd_{98}Te_{98}$. The zero energy is set at the Fermi level. (Reprinted with permission from S. Sarkar et al., J. Phys. Chem. C, 2012, **116**, 21601. © 2012, American Chemical Society.)

Brown et al.[7] Their study reveals that the photocurrent generation efficiency of CdSeQD–C_{60} films is two orders of magnitude greater than CdSe films alone. Gigli and co-workers have designed a novel method for the fabrication of hybrid NC–C_{60} cells with a potentially long lifetime and thus has the potential for the development of optimized organic–inorganic hybrid photovoltaic cells.[115] Shibu et al.[57] have synthesised a novel fullerene-shelled CdSe QD hybrid nanostructures by the covalent tethering of a fullerene-thiol monolayer to the QD followed by the photochemical reactions of free fullerene-thiol to the tethered monolayer. These authors have studied the ET from the highly stable QD to the protecting fullerene-shell and argued that this nanoarchitecture is among the most promising antenna systems for solar energy harvesting. Bang et al.[6] have recently synthesized and studied the light harvesting efficiencies of CdSeQD–C_{60} nanohybrid systems. They have shown that the rate constant for ET between excited CdSeQD and C_{60} was governed by the energy difference between the LUMO of CdSeQD and LUMO of C_{60},

and followed the trend predicted by the Marcus theory.[13,114] Their study suggests that nanohybrids with larger CdSe QDs exhibit higher energy conversion efficiency than smaller NPs despite being relatively slower electron injection rate in larger CdSe QDs. This study also suggests that it is not the forward electron injection process, rather the slower hole transfer and faster recombination of charge carriers are more important in dictating the performance of these nanocomposites as solar cells. This particular insightful experimental study has promoted many theoretical investigations on the electronic structure of these novel CdSe QDs and fullerene nanocomposites.[54,116] In the following, we will discuss some important electronic features of CdSeQD–C_{60} nanohybrids that we have obtained from our SCC-DFTB calculations.

The DOSs of five CdSeQD–C_{60} nanocomposite systems with CdSe QDs of different sizes are shown in Fig. 4. The PDOS of individual QDs (represented by shaded areas) and C_{60}-thiol molecule (represented by blue line) are also shown in the same figure. From the figure it is seen that the bands that arise from the QDs (specified as shaded region) are
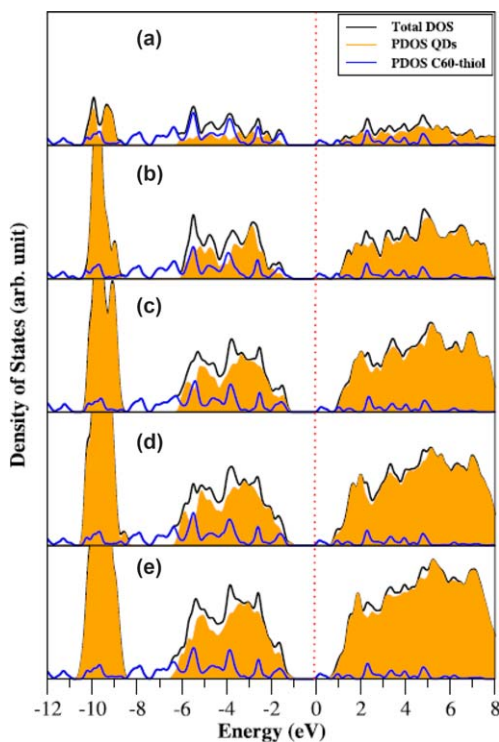
Fig. 4 The density of states (DOS) of CdSeQD–C_{60} nanohybrids constituted with one C_{60}-thiol molecule and CdSe QDs of different sizes (a) $Cd_{10}Se_{10}$ (1.24 nm), (b) $Cd_{34}Se_{34}$ (1.85 nm), (c) $Cd_{55}Se_{55}$ (2.10 nm), (d) $Cd_{65}Se_{65}$ (2.15 nm), and (e) $Cd_{80}Se_{80}$ (2.44 nm). The shaded region represents the partial density of states of the corresponding CdSe QDs and the blue line represents the partial density of states of C_{60}-thiol (PDOS C60-thiol: for colored figures, please see the online version). The black line represents the total density of states of the corresponding nanohybrids (Total DOS). The zero energy is set at the Fermi level. (Reprinted with permission from S. Sarkar *et al.*, *J. Appl. Phys.*, 2014, **116**, 114303. © 2014, American Institute of Physics.)

broadened with increasing the size of the QDs and also the HOMO–LUMO gap of an individual QD decreases because of the quantum confinement effect. The interesting feature of the Fig. 4 is the appearance of a low intensity peak just above the Fermi level. This peak is essentially attributed to the LUMO of the C_{60}-thiol molecule. Interestingly, for nanohybrids with smaller CdSe QDs the HOMO of the C_{60} represents the valence band edge of the nanohybrids (Fig. 4a and b) whereas for larger QDs, HOMO of the QDs represents the valence band edge of the nanohybrids (Fig. 4c–e).

One may have the qualitative understanding on the rate of recombination of charge carriers from the spatial distribution of the frontier orbitals. So, we have shown the spatial distribution of LUMO, HOMO and HOMO-1 charge densities of the CdSe QD–C_{60} nanohybrids in Fig. 5. The figure shows that the LUMO (light blue) charge densities of all the nanohybrids are localised only on the fullerene moiety of the CdSeQD–C_{60} nanohybrids. However, the distributions of HOMO (red) charge densities of the hybrid systems are different for QDs of different size. It is very interesting to see that HOMO of the CdSeQD–C_{60} nanohybrid with 1.24 nm diameter of CdSeQD is located on the fullerene and the thiol chain while in case of a QD of diameter 1.85 nm, it is located on the thiol chain and the adjacent surface of the CdSe QD. But for the larger sized CdSe QDs (e.g. 2.05, 2.15 and 2.44 nm of diameter), the HOMO is localised only on the QDs. So, the spatial distribution of the frontier orbitals of CdSeQD–C_{60} nanohybrids reveals that the conduction band edge of the hybrid systems has the characteristic of the LUMO of the thiol-functionalized fullerene.
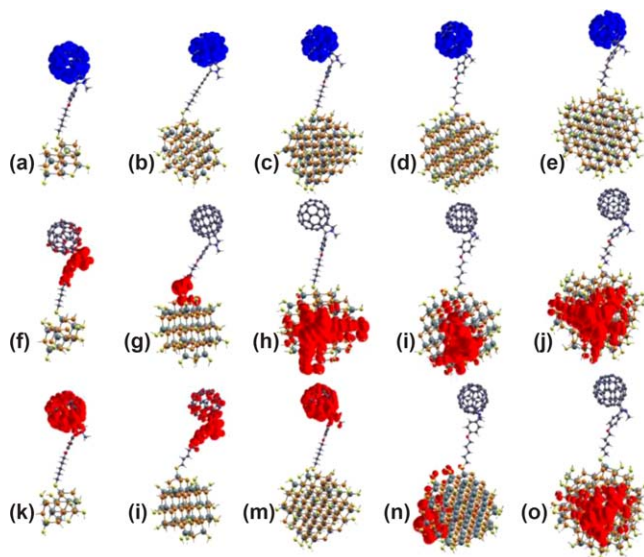
Fig. 5 (a)–(e) LUMO (blue), (f)–(j) HOMO (red), and (k)–(o) HOMO-1 (red) charge densities for the CdSeQD–C_{60} nanohybrids with CdSe QDs of different sizes; (a), (f), and (k) for $Cd_{10}Se_{10}$ (1.24 nm); (b), (g), and (l) for $Cd_{34}Se_{34}$ (1.85 nm); (c), (h), and (m) for $Cd_{55}Se_{55}$ (2.10 nm); (d), (i), and (n) for $Cd_{65}Se_{65}$ (2.15 nm); (e), (j), and (o) for $Cd_{80}Se_{80}$ (2.44 nm). The isosurface value used was 0.0001 e$Å^{-3}$. (Reprinted with permission from S. Sarkar et al., J. Appl. Phys., 2014, **116**, 114303. © 2014, American Institute of Physics.)

However, contribution to the valence band edge of the CdSeQD–C$_{60}$ nanohybrids very much depends on the size of the CdSe QDs. Thus, for larger sized QDs the valence band edge of the nanohybrids has the characteristic of the HOMO of the CdSe QDs but for smaller size of QDs, it has the characteristic of the HOMO of the thiol-functionalized fullerene. The HOMO-1 charge densities of nanocomposites of smaller QDs are localized in the fullerene whereas the same for larger QDs are localized in the QDs. So, for the nanohybrids with larger CdSe QDs, the electrons and holes are spatially separated and accordingly the life time of the charge carriers will be long for these systems and thus they are considered to be good for application in solar cell. From the study of the spatial distribution of charge densities of CdSeQD–C$_{60}$ nanohybrids one may conclude that, nanocomposites, for which the HOMO and LUMO densities are on different components, facilitates the charge separation and consequently the charge recombination dynamics will be delayed in those systems making them suitable for use in solar cells.

To get some qualitative idea about the electron injection dynamics from CdSe QDs to fullerenes, we have analyzed the band energy alignment, *i.e.*, the position of the HOMO and LUMO of the composite systems along with the individual components (C$_{60}$ and CdSeQDs) for different sized CdSe QDs and these are shown in Fig. 6. From the figure it is seen that the energies of HOMO and LUMO of the composite systems are not exactly superposed on their constituents. The energies are either blue or red shifted (although small) as compared to the individual components. In this context it should be mentioned that, Chaban *et al.*,[54] through an time-domain *ab initio* study, have shown that the covalent
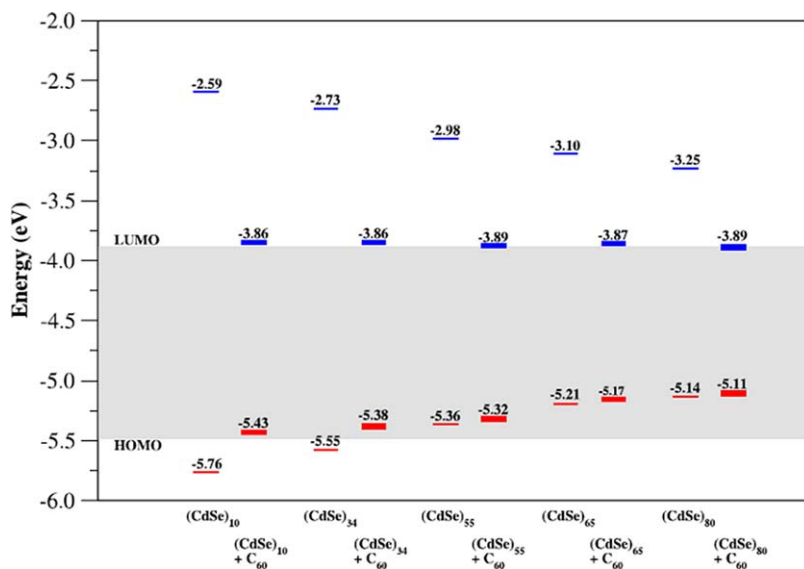

Fig. 6 HOMO (red) and LUMO (blue) energy alignment of the pure CdSe QDs with different sizes and corresponding CdSeQD–C$_{60}$ hybrid systems (bold) in the gas phase. The shaded region identifies the HOMO–LUMO gap of the C$_{60}$-thiol molecule. (Reprinted with permission from S. Sarkar *et al.*, *J. Appl. Phys.*, 2014, **116**, 114303. © 2014, American Institute of Physics.)

QD–C_{60} binding enhances the charge separation efficiencies and thus improves the photo-conversion efficiencies in the covalently linked systems as compared to mechanical mixtures. The figure shows that the valence band-edge of the QDs undergoes an upward shift with increasing size of the QD and it crosses the HOMO level of the C_{60}-thiol molecule thereby constructing the nanohybrid with a type-II band energy alignment with respect to the C_{60}-thiol molecule and may be useful material to harvest the solar energy. It is worthy to mention here that the band energy alignment of a hybrid system is termed as type-I or type-II depending whether both the VBT (valence band top) and CBM lie on a same species (donor or acceptor) or they are on two different constituents. The Fig. 6 clearly demonstrates that the CdSeQD–C_{60} nanohybrid systems with smaller CdSeQD (of diameters 1.24 and 1.85 nm) exhibit type-I band energy alignment while that for QDs having larger diameter (2.10, 2.15 and 2.44 nm) is of type-II. So, the CdSeQD–C_{60} hybrid systems show a transition from type-I to type-II with increasing the size of the CdSeQDs of the nanohybrids. Another important point is to note that the energy difference between the conduction band edge of the CdSeQD and LUMO of the C_{60}-thiol molecule $[\Delta G_{\text{electronic}} = E_{C_{60}}(\text{LUMO}) - E_{\text{CdSe}}(\text{CB})]$ decreases with increasing the size of the CdSeQDs. The calculated $\Delta G_{\text{electronic}}$ values are shown in Table 1. On the basis of Marcus theory of ET,[13,114] as discussed earlier, we could expect faster ET rates for smaller QDs as compared to larger QDs when these are coupled with the C_{60}-thiol molecule. In fact, the rate of ET increases, reaching a maximum when the driving force is equal to the reorganization energy. The experimental study of Bang et al.[6] on this particular nanohybrid system infact found a faster ET rate for smaller QDs as compared to larger QDs. The idea of electronic energy levels therefore rationalizes the experimental results. Thus, these size-dependent inter-particle ET rates will have to be taken into account in band gap engineering of QDSCs.

From the above discussion on band energy alignments and the spatial distribution of charge densities of frontier orbitals of CdSeQD–C_{60} nanohybrids, we have come across two opposing events: nanohybrids with smaller QDs have better electron injection efficiencies while the electron and hole are not spatially separated. On the other hand, for the nanohybrids with larger QDs, the electron and holes are spatially separated thus having delayed electron–hole recombination but the rate of electron injection is slower. So, the best performance of solar cell would be

Table 1 Variation of band gap and free energy of electron transfer $[\Delta G_{\text{electronic}} = E_{\text{CBM}}$ (CdSeQD) $- E_{\text{LUMO}}$ (C_{60})] in CdSe–C_{60} nanohybrids with size of CdSeQD. (Reproduced with permission from S. Sarkar et al., J. Appl. Phys., 2014, **116**, 114303. © 2014, American Institute of Physics.)

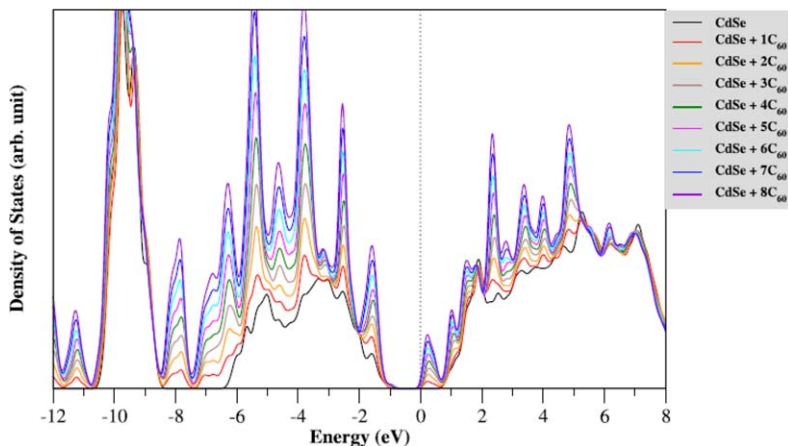| Identity	Diameter (nm)	Band gap (eV)	$\Delta G_{\text{electronic}}$ (eV)
$Cd_{10}Se_{10}$	1.24	3.17	−1.30
$Cd_{34}Se_{34}$	1.85	2.82	−1.16
$Cd_{55}Se_{55}$	2.10	2.38	−0.91
$Cd_{65}Se_{65}$	2.15	2.11	−0.79
$Cd_{80}Se_{80}$	2.44	1.89	−0.64

Fig. 7 The total density of states (DOS) of CdSeQD of a particular diameter (2.44 nm) and the corresponding CdSeQD–C_{60} nanohybrids with different concentrations of C_{60}. The zero energy is set at Fermi level. (Reprinted with permission from S. Sarkar et al., J. Appl. Phys., 2014, **116**, 114303. © 2014, American Institute of Physics.)

achieved by the optimal choice of the size of the CdSeQD (for which it is the best compromise between the electron injection rate and the electron–hole recombination rate) in the nanohybrids.

Bang et al.[6] in their experimental study found that the concentration of fullerene molecules in the nanohybrids plays a great role in dictating the light harvesting efficiencies. They have shown that the photovoltaic efficiency increases with increasing fullerene concentration in the nanohybrids. By measuring the ensemble-averaged transient absorption and fluorescence decay as well as single QD-fluorescence decay, Song et al.[117] have shown that the ET rate from CdSeQD to C_{60} molecules increases with increasing the C_{60}-to-QD ratio. Now to explain this experimental observation we have shown the density of states CdSeQD–C_{60} nanohybrid systems for various concentrations of C_{60}-thiol moieties in Fig. 7. The figure clearly demonstrates that the density of states (DOS) of the composite system increases at the position of the LUMO of C_{60} with increasing concentration of C_{60}-thiol moieties. The increase in DOSs at the position of the LUMO of C_{60} results the shortening of the lifetime of the excited electrons of CdSeQDs and thus there is faster deactivation of charge-separated states of CdSeQDs *via* ET from CdSeQD to the thiolated C_{60}. So, one can increase the light harvesting efficiency of these nanohybrids by increasing the concentration of C_{60}-thiol moieties in the nanohybrids.

3.3 CdTe quantum dot–graphene hybrid nanostructures

Graphene, the latest discovery of carbon nanostructure, has extraordinary properties such as high thermal conductivity, high elasticity and higher drift velocities and thus may be a good alternative material to CNTs and fullerenes as organic component to be integrated with inorganic QDs for designing novel nanohybrids.[118,119] The different experimental groups have explored the QD–graphene nanohybrids for their potential

application in photovoltaic cells and these studies established that graphene is a good acceptor of photoexcited electrons from QDs and can also delay the process of recombination of charge carriers.[44,47,49,52,68,90,120,121] The ultra-fast transient absorption spectroscopic and time-resolved emission spectroscopic studies have indicated the efficient ultra-fast charge separation on graphene–CdTe nanohybrids.[49] This novel nanohybrid can provide a new and promising direction toward developing high-performance light-harvesting devices for the next generation solar cells.

To study the electronic structure of CdTeQD–graphene nanohybrids, we have modelled the nanohybrids as follows.[122] We have considered a 16×16 supercell of a graphene sheet in the xy plane which contains 512 carbon atoms. One carbon atom of the sheet is functionalized with –COOH group and one –H atom is atatched to the neighbouring carbon atom to maintain the charge neutrality. The modelling of CdTe QD and its integration with the –COOH functionalized graphene is done similar to that of CdTeQD–CNT as discussed in Section 3.1. There is experimental report that the covalent linking of the two components in nanohybrids increases the photo-conversion efficiency by 2–3 orders as compared to their mechanical mixture.[54] The optimized structures (both top and side views) of one representative CdTeQD–graphene hybrid nano-composites [GR–(CdTe)$_{65}$] are shown in Fig. 8.

To understand the contribution of electronic energy levels of graphene and the QDs in the nanohybrids, we have shown in Fig. 9 the DOSs of GR–CdTeQD nanohybrids with CdTe QDs of four different sizes. The projected density of states (PDOSs) of individual QDs and graphene are also shown in the same figure. From the figure it is clear that the band gap region of the nanohybrid is dominated by the graphene sheet, while
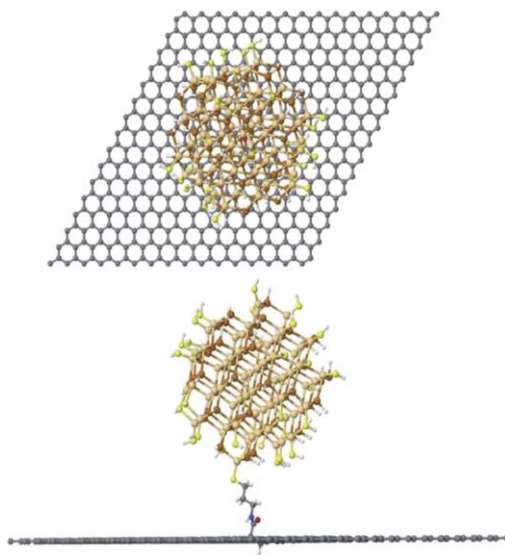

Fig. 8 Optimized structures [both top (upper) and side (lower) views] of one representative GR–CdTeQD [GR–Cd$_{65}$Te$_{65}$ QD] nanohybrid system. (Reprinted from B. Rajbanshi *et al.*, *J. Mater. Chem. C*, 2014, **2**, 8967 with permission from Royal Society of Chemistry.)
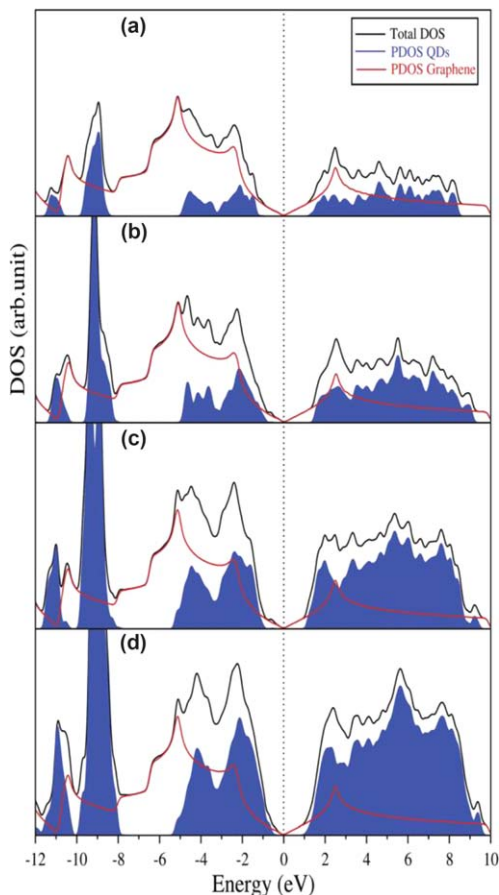
Fig. 9 The density of states (DOS) of GR−CdTeQD nanohybrids with QDs of four different sizes: (a) $Cd_{17}Te_{17}$, (b) $Cd_{34}Te_{34}$, (c) $Cd_{65}Te_{65}$ and (d) $Cd_{98}Te_{98}$. The zero energy is set at the Fermi level. (Reprinted from B. Rajbanshi *et al.*, *J. Mater. Chem. C*, 2014, **2**, 8967 with permission from Royal Society of Chemistry.)

the contributions of CdTe QDs are in the deep of both the valence and conduction band regions. However, with increasing the size of the CdTe QDs, both the HOMO and LUMO of QDs approach toward the Fermi level. Following the similar argument for that of CdTeQD–CNT and CdSeQD–C$_{60}$ composites, here also we see that the driving force of ET from CdTeQD to graphene $[E_{LUMO}(QD) - E_{CB}(graphene)]$ increases with decreasing the size of QD. However, this fast transfer of electrons from the QD to the graphene results the fluorescence quenching of the composite system as compared to the isolated CdTeQD that has already been found in experimental studies on this type of system.[44,49]

The band gap values of CdTeQD–graphene nanohybrids are small as compared to isolated CdTeQDs. The DOSs figures of the hybrid systems show that the graphene states are introduced in between the HOMO and LUMO states of the QD. So, our theoretical result suggests that the absorption spectra of GR–CdTeQD nanohybrid will show a clear red shift as compared to the CdTeQD. Markad *et al.*[52] in their experimental studies

on these particular nanohybrids observed a red shift in the absorption spectra, the origin of which is now clear from our theoretical studies. As the ET from QD to the graphene are faster for smaller QDs as compared to larger ones, the fluorescence quenching rate will decrease with increasing size of the QDs. This qualitative picture of fluorescence quenching rate has recently been observed by Kundu *et al.*[90] in their experimental studies on CdTeQD–GO composite systems. In this regard it should be pointed out that, surface states of QDs also have a great role on fluorescence quenching. It has been observed that the luminescence decay kinetics may be governed by both intrinsic and surface states. In our case, the passivation removes the surface states and hence only the intrinsic valence and conduction bands of the CdTe QD are involved during charge transfer processes.

The zero band gap of the graphene is one of the major obstacles for its application in electronics and photonics. Thus the researchers have devised several strategies *e.g.*, through chemical functionalization to have a tunable band gap in the intermediate energy range (*e.g.*, 1–3 eV). Fully hydrogenated graphene, *i.e.*, graphane has a large band gap (5.4 eV). This suggests that, it is possible to tailor the band gap in a wide range by controlling the hydrogen coverage. So, one can have the additional parameter, *i.e.*, the percentage of hydrogen covering in addition to the size of the CdTe QD to have proper band energy alignment in CdTeQD–graphene hybrid nanostructures for solar cell application. In the next section we would like to present few interesting results on the electronic structure of hydrogenated graphene–CdTeQDs (HG–CdTeQD). Our objective is to find the appropriate combination of size of the CdTeQD and the percentage hydrogenation of graphene for which there will be faster ET from the QD to the graphene and slower electron–hole recombination. We have chosen four partially hydrogenated graphene surfaces (HG) with different percentage of hydrogen as 12.5%, 25%, 50% and 62%. The trend of band gap variation of these partially hydrogenated graphene agrees well with the results as reported by Gao *et al.*[45] However, we would like to discuss the effect of both the extent of hydrogen passivation on graphene and also the size of the CdTe QD on the electronic structure of the HG–CdTe nanohybrids.

The DOSs and PDOSs of few HG–CdTe nanohybrids along with the individual QDs and hydrogenated graphenes are shown in Fig. 10. The interesting point to note from the DOSs figure is that when the percentage of hydrogenation in HG–CdTeQDs is low, (say, 12.5% or 25%) the band gap region is dominated by graphene for nanohybrids with smaller QDs (of radius 0.91 nm or 1.08 nm) as in case of GR–CdTeQD nanohybrid systems. Thus, both the VBM and CBM of the composite system are localized on graphene. So, we can conclude that the HG–CdTeQDs with lower percentage of hydrogen and smaller QDs show type I band alignment. But for nanohybrids with fixed percentage of hydrogenation on graphene, as the size of the CdTeQD increases the VBM of HG–CdTeQD hybrid system is dominated by the valance band maximum (VBM) of HG, while the CBM of HG–CdTeQDs hybrid system is controlled by the LUMO of CdTeQD. Thus, HG–CdTeQDs hybrid systems with lower percentage of
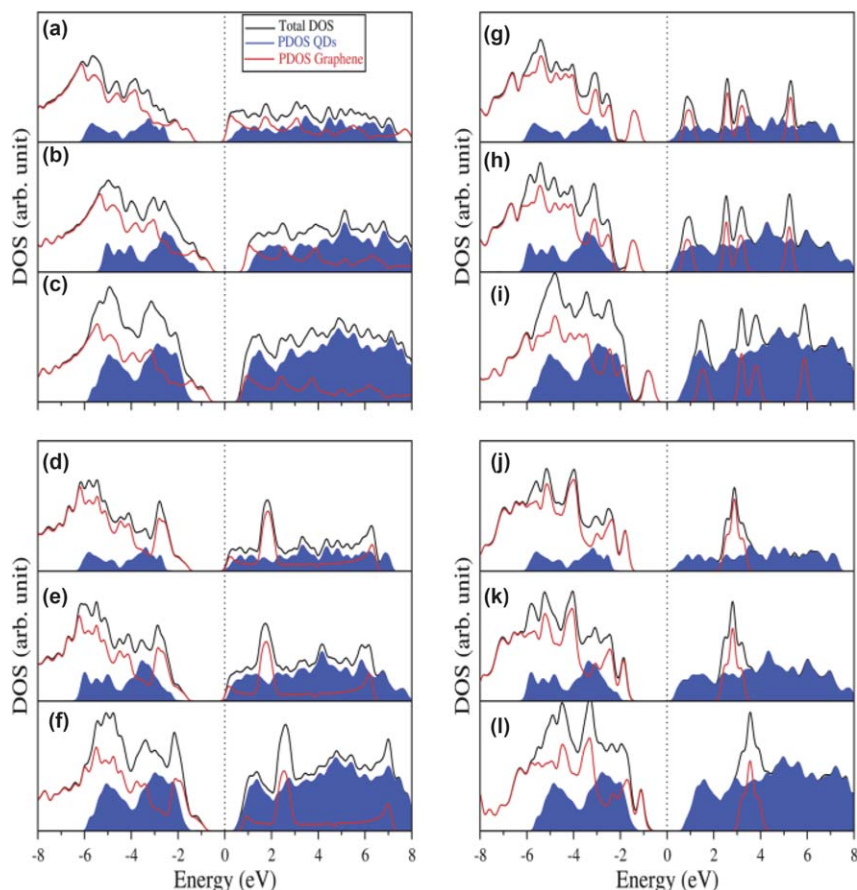
Fig. 10 The density of states (DOS) of HG−CdTeQD nanohybrids with different percentages of hydrogen, each coupled with three QDs of different sizes (radius): 12.5% HG with 0.91, 1.08 and 1.28 nm CdTeQDs (a−c); 25% HG with 0.91, 1.08 and 1.28 nm CdTeQDs (d−f); 50% HG with 0.91, 1.08 and 1.28 nm CdTeQDs (g−i) and 62.5% HG with 0.91, 1.08 and 1.28 nm CdTeQDs (j−l). The zero energy is set at the Fermi level. (Reprinted from B. Rajbanshi *et al.*, *J. Mater. Chem. C*, 2014, **2**, 8967 with permission from Royal Society of Chemistry.)

hydrogenated graphene and CdTeQDs of larger size exhibit type-II band alignment. Hence, increasing the size of QDs in the QD−graphene nanohybrid is one possible pathway for making a transition from type-I to type-II band alignment. More importantly, HG−CdTeQDs with higher percentage of hydrogen (50% or more), always show type II alignment for QDs of all sizes. The LUMO of QDs appears in lower energy than CBM of HG, whereas VBM of these HG−CdTeQDs is dominated by VBM of HG as before. Alternatively, one can change the percentage of hydrogenation on graphene for a fixed size of the QD in the nanohybrids to have type-II band alignment.

To demonstrate such situation we have plotted VBM and CBM densities of three HG−CdTeQD nanohybrids such as (a) 25% hydrogenated graphene with QD of radius 0.91 nm, (b) 25% hydrogenated graphene

Fig. 11 VBM and CBM densities of (a) 25% HG with 0.91 nm CdTeQD, (b) 25% HG with 1.28 nm CdTeQD and (c) 50% HG with 0.91 nm CdTeQD. The red surfaces represent the densities of VBM whe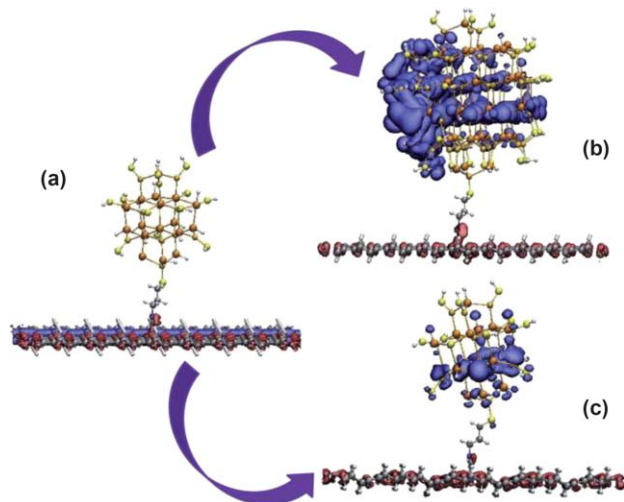reas the blue surfaces represent the CBM. (Reprinted from B. Rajbanshi *et al.*, *J. Mater. Chem. C*, 2014, **2**, 8967 with permission from Royal Society of Chemistry.)

with QD of radius 1.28 nm and (c) 50% hydrogenated graphene with QD of radius 0.91 nm respectively, in Fig. 11. From the Fig. 11(a) it can be noticed that, both the VBM and CBM are residing on the hydrogenated graphene sheet, indicating that it is a type-I system. But, as we increase the size of the QD from radius 0.91 nm to 1.28 nm keeping the percentage of hydrogenation on graphene fixed (*i.e.* the same HG), the VBM density is delocalized over the HG whereas CBM density is now localized on the QD [Fig. 11(b)] thus exhibiting the type-II band alignment. So, one can achieve the type-II band alignment by changing the size of the CdTeQD in the HG–CdTeQD nanohybrid systems. Alternatively, if we look at Fig. 11(a) and (c) [differs only in the percentage of hydrogenation, the size of the QDs are same], we notice that with increasing the percentage of hydrogenation on graphene (from 25% to 50%), the VBM density remains delocalized over the HG and CBM is delocalized over the CdTeQD. So, one may have type-II band alignment by changing the percentage of hydrogenation on graphene sheet.

Therefore, we conclude that the CdTeQD–graphene nanohybrids with CdTeQD of suitable size and appropriate percentage of hydrogenation form type-II band alignment and may be a suitable donor–acceptor systems to be used in solar cells. The formation of type-II band alignment and resulting delayed charge recombination increases the photovoltaic performance for these nanocomposites. Our qualitative understanding of delayed charge recombination is in good agreement with the experimental results of Kaniyankandy *et al.*[49] The bleach recovery kinetic study of these authors indicates that the recombination dynamics of CdTeQD–graphene nanocomposites is much slower as compared to that of the isolated CdTeQD.

3.4 Porphyrin-sensitized graphene quantum dot solar cell

Among the different dyes used in dye-sensitized solar cells, the porphyrin molecule has a very different position because of its excellent light-harvesting property mimicking natural photosynthesis.[79,123,124] So, the porphyrin molecule is considered to be one of the useful components for integration with semiconductor QDs to form hybrid nanostructures. Although among the different semiconductor QDs, Cd-chalcogenides are mostly used for forming nanohybrids, yet, because of the environmental issues the use Cd-chalcogenides in designing solar cells will be rather limited. It is therefore highly desirable to replace Cd-chalcogenides with environmentally safe QDs but having the similar properties as those of Cd-chalcogenides. In this context, graphene quantum dots (GQDs), which have recently been synthesized by Li and co-workers,[125–129] may be a good alternative to integrate with the organic dyes for making solar cell devices. The most important advantage of the GQDs is their large Bohr exciton radius. The quantum confinement effect, because of large Bohr exciton radius, is found to operate at any finite size of the QD. Thus, one has the opportunity to tune the HOMO and the LUMO energy levels over a wide range of wavelengths spanning from the ultraviolet to the visible to the infrared. Numerous reports are now available in the literature demonstrating that the photo-luminescence properties of the GQDs can be tuned through controlling their size, shape, edge configurations and chemical functionalization.[130–132]

In the very recent past, there are many experimental studies devoted to the synthesis of porphyrin functionalized graphene hybrid materials.[43,46,48,50,53,133–135] Chen and co-workers[53] have synthesized porphyrin–graphene nanohybrid and showed that this nanohybrid material has superior optical limiting properties and the authors argued that this nanohybrid may be a new entry in the realm of light harvesting devices. Because of the many exciting properties of GQDs as mentioned in the preceding section, one may expect that nanohybrids consisting of GQDs and a porphyrin molecule would show better photovoltaic properties. In view of this, we will discuss the electronic structure of these particular nanohybrids in the next section. Our emphasis in the discussion will be on how the functionalization of the GQDs with a porphyrin molecule modifies the electronic structure of the GQDs and whether the electronic energy levels of the resulting nanohybrids have proper band energy alignment for the efficient transfer of photoexcited electrons or have favourable spatial charge separation for lower recombination of the charge carriers.[136]

To model the GQDs of different sizes we have first built up a large graphene sheet and then chosen GQDs of different shapes and sizes.[136] In this way, we have generated GQDs up to 10 nm in size retaining their symmetries. Finally, to model the GQD–porphyrin nanohybrids we have attached the porphyrin molecules with the GQD through amide linkage. Porphyrin molecules contain four phenyl rings, one of them involves in the formation of the amide linkage and the different functional groups are placed at the para position of the remaining phenyl rings by removing their para hydrogens. The optimized structures of two representative
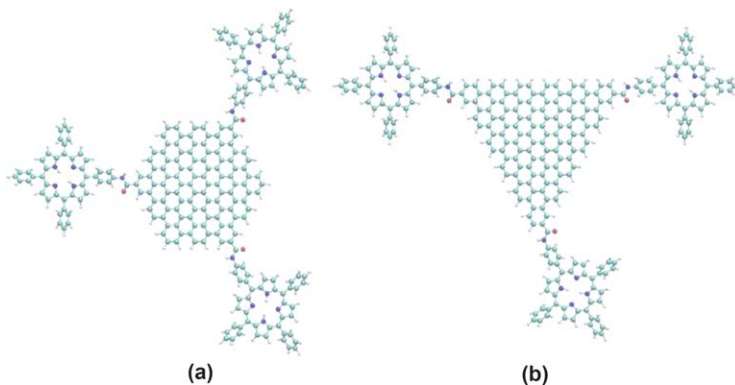
Fig. 12 Optimized geometries of GQD–porphyrin nanocomposites [hexagonal (a) and trigonal (b)] having the number of carbon atoms 114 and 126, respectively, in hexagonal and trigonal GQD. (Reprinted with permission from B. Mandal *et al.*, *J. Phys. Chem. C*, 2015, **119**, 3400. © 2015, American Chemical Society.)

porphyrin functionalized GQDs (hexagonal and triangular) are shown in Fig. 12.

To start our discussion let us first examine the effect of functionalization of the GQDs with a porphyrin molecule. It is now well established that functionalization of nanostructures with organic molecules in addition to size and shape plays an important role in tuning the electronic energy levels of the nanostructures. However, several studies on GQDs showed that the functionalization with small organic molecules has very little effect on the HOMO–LUMO gap of the GQDs. Now, to see whether the functionalization of GQDs with porphyrin molecules can bring a change in the HOMO–LUMO gap, we have shown the variation of HOMO–LUMO gap as a function of the size of the GQD in Fig. 13. The figure clearly establishes that the HOMO–LUMO gap of smaller sized GQDs gets reduced because of the functionalization with the porphyrin molecules. So, the GQD–porphyrin nanocomposite will show a red shift in its absorption spectra as compared to GQD and this is in good agreement with experimental study on porphyrin–graphene nanohybrids of Xu *et al.*[53] However, the HOMO–LUMO gap of the GQD–porphyrin composite system for larger sized GQD is same as that of GQD itself. This tempted us to conclude that the HOMO–LUMO gap of the nanohybrids with larger GQDs is governed by the electronic energy levels of GQD. One can have a strong foundation to this conclusion from the analysis of the DOSs and PDOSs of GQD–porphyrin nanohybrids. In Fig. 14 we have shown the total density of states (DOS) of GQD–porphyrin nanohybrids along with the projected density of states (PDOS) of GQDs and porphyrin molecules. The most striking finding of the DOSs figure is that, the nanohybrids with smaller sized GQDs [Fig. 14(a) and (b)] exhibit type-II band energy alignment. The HOMOs of the hybrid systems are controlled by the porphyrin molecules while the LUMOs are controlled by the GQDs. The nanohybrids with larger GQDs (Fig. 14(e)) show type-I band alignment where both the HOMOs and LUMOs of the hybrid systems are controlled by the GQDs. So there is a transition from type-II to type-I band
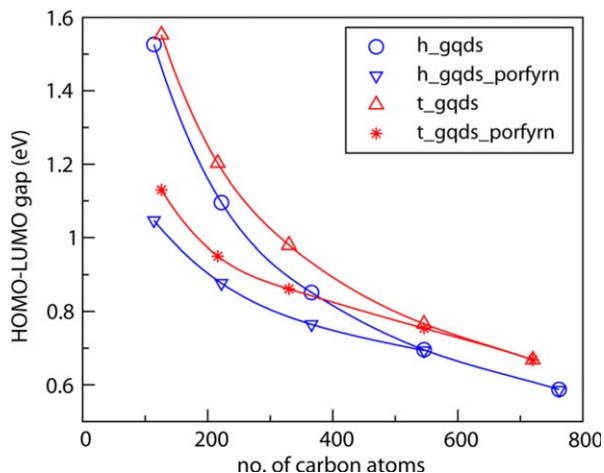
Fig. 13 Variation of energy gap against number of conjugated carbon atoms (N) of GQD for hybrid systems as well as for GQDs. Blue line with circle and triangle (down) used for hexagonal GQD and hexagonal composite, respectively, while red line with triangle (up) and star used for triangular GQD and triangular nanocomposites, respectively. (Reprinted with permission from B. Mandal *et al.*, *J. Phys. Chem. C*, 2015, **119**, 3400. © 2015, American Chemical Society.)

alignment with the increase in size of the GQD. Thus, for nanohybrids with smaller QDs, there is spatial charge separation and as a consequence of that the chance of recombination of charge carriers is expected to be low. The figure also clearly reveals that the energy gap between the LUMO of the donor porphyrin molecule and the LUMO of the acceptor GQD increases with increasing the size of the GQD. Thus, the rate of ET from porphyrin to GQD will increase with increasing size of the GQD. So, we conclude that this nanohybrid has proper band energy alignment and may be suitable to be used as building block of third generation solar cell. To get support of the conclusion of the above statement, the HOMO and LUMO densities of three different GQD–porphyrin nanohybrids are shown in Fig. 15. From the figure it is clear that, for nanohybrids of smaller GQD, the LUMO is localized on porphyrin and HOMO density is localized on the GQD. So, there is a clear charge separation in GQD–porphyrin nanohybrids with smaller GQD. However, in nanohybrids with larger GQDs both HOMO and LUMO densities are on GQD ((e) and (f) of Fig. 15).

Now, let us see the optimum size of GQD for offering better photovoltaic performance. To elucidate, let us compare the possibility of efficient ET from dye to the QD and the spatial separation of the charge carriers that lowers the electron–hole recombination rate. Similar to other donor–acceptor nanohybrids discussed earlier, here also we noticed two contrasting events. The electron injection efficiency which depends on the difference in energies between the LUMOs of the porphyrin and GQD increases with increase in size of the GQD in the nanohybrids. But at the same time for bigger GQDs, both the HOMO and LUMO charge densities are localized in the same material, namely GQD. So, the recombination of charge carriers will be much faster, which
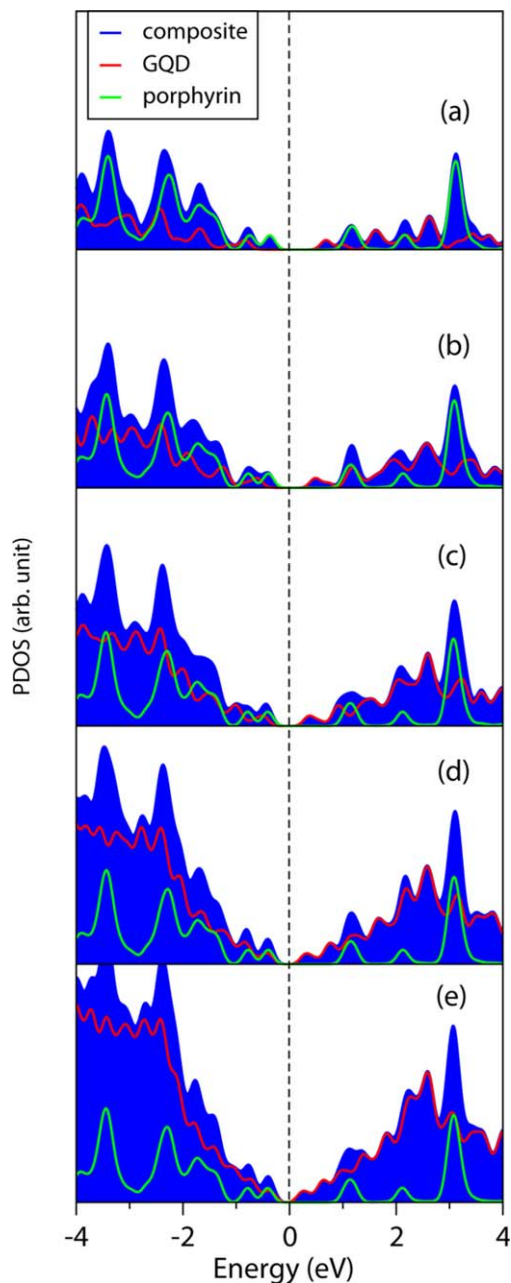
Fig. 14 The projected density of states (PDOS) of five different hexagonal nanohybrids having number of conjugated carbon atoms 114 (a), 222 (b), 366 (c), 546 (d), and 762 (e) in GQDs. Shaded regions represent the total density of states of nanocomposites, while the regions under red and green line represent the contribution of GQD and porphyrin, respectively. The dotted line represents the Fermi level. (Reprinted with permission from B. Mandal *et al.*, *J. Phys. Chem. C*, 2015, **119**, 3400. © 2015, American Chemical Society.)

essentially reduces the photovoltaic efficiency. Now, it would be interesting if it is possible to achieve type-II band energy alignment in larger nanohybrids through some chemical functionalization retaining its
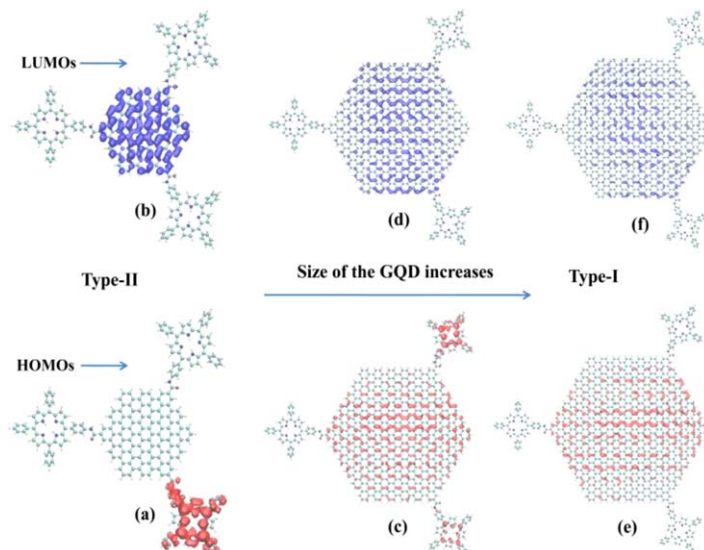
Fig. 15 HOMO (red) and LUMO (blue) charge density of hexagonal composite systems with carbon atoms in GQD 114 (a and b), 546 (c and d), and 762 (e and f), respectively. Isovalue = 0.0001 eÅ^{-3} is used for plotting. (Reprinted with permission from B. Mandal et al., J. Phys. Chem. C, 2015, **119**, 3400. © 2015, American Chemical Society.)

faster electron injection rate. The functionalization with organic functional groups is one of the many ways through which one can tune the electronic structure of NPs. Now, with a goal to achieve type-II band alignment in GQD–porphyrin with larger GQDs we have functionalized the porphyrin molecule by electron donating groups –OCH$_3$ and –NH$_2$.

In Fig. 16, we have shown the wave functions corresponding to the HOMO and LUMO of one representative system (GQD–porphyrin nanocomposite functionalized with –OCH$_3$). From the figure it is evident that HOMO is localized on one porphyrin molecule, while the LUMO is localized on the GQD. So, one achieves the desired distinct charge separation in these nanocomposites. We also found the same feature for –NH$_2$ functionalized porphyrin–GQD nanocomposites. Thus, the functionalization of porphyrin with an electron donating group such as –OCH$_3$ and –NH$_2$ results in forming type-II band energy alignment. The functionalization with electron donating groups raises the HOMO energy of the porphyrin molecule above that of the GQD. Thus, we found that, nanocomposites with larger size of GQDs which were type-I, now become type-II with functionalization of the porphyrin molecule with electron donating groups such as –OCH$_3$ and –NH$_2$. Another interesting feature of these nanocomposites as evident from the density of states (not shown here) is that the energy gap between the LUMO of porphyrin and that of the GQD increases when an electron donating group is attached to the porphyrin molecule as compared to nonfunctionalized nanocomposites rendering better electron injection efficiency for the former. So, the GQD–porphyrin nanocomposite with either smaller or chemically modified larger GQD can offer type-II band energy alignment. From Fig. 16(c) and (d) it is very clear that the HOMO, LUMO wave functions are on the
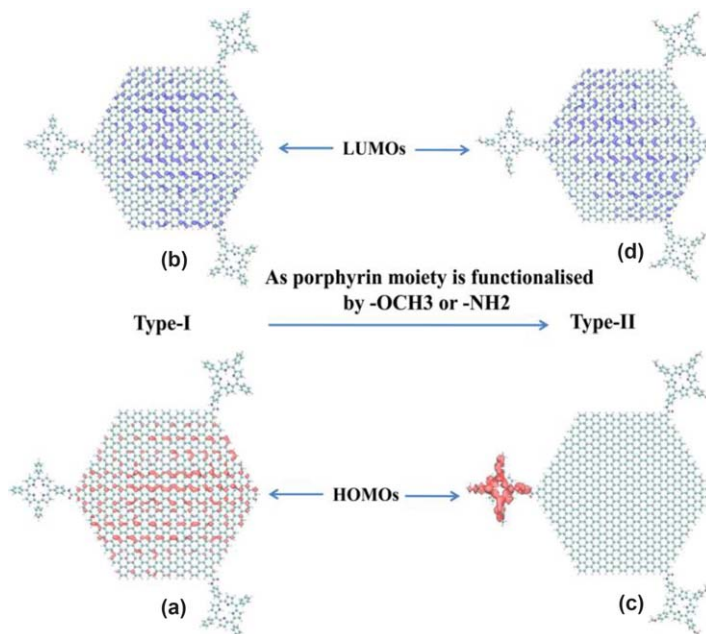
Fig. 16 HOMO (red) and LUMO (blue) charge density of largest size hexagonal nanohybrid with unsubstituted (a and b) and $-OCH_3$ substituted (c and d) porphyrin moiety. (Reprinted with permission from B. Mandal *et al.*, *J. Phys. Chem. C*, 2015, **119**, 3400. © 2015, American Chemical Society.)

different components of the nanohybrids. This results a negligible overlap between them and hence the rate of charge recombination for these systems will be low. Thus, GQD–porphyrin nanohybrid systems have the promise to be good building blocks for realization of dye-sensitized solar cells.

3.5 Graphene quantum dot–fullerene nanohybrids

In this section, we have reviewed the results of the electronic structure calculations of GQD–fullerene nanohybrids to understand the suitability of using these materials for solar cell applications.[42] The test of suitability for applications in solar cells is judged as also done for previous systems by spatial charge separation and efficient ET from donor to the acceptor.

The electron injection rate depends on the energy difference between the LUMO of the donor and the LUMO of the acceptor (ΔG). Now, one can tune the energy gap of the GQD donor over a broad range by changing the size of the GQD. Thus, one may have control over the charge injection rate through controlling the size of the GQD. To understand the role of size in controlling the electron injection rate the positions of the frontier energy levels, *i.e.*, the positions of the HOMO and the LUMO of the hybrid systems along with those of the individual components, GQDs and the fullerene are shown in Fig. 17. The point to note from the figure is that the HOMO, LUMO energies of the hybrid systems are not exactly the superposition of the energies of the isolated components. The figure
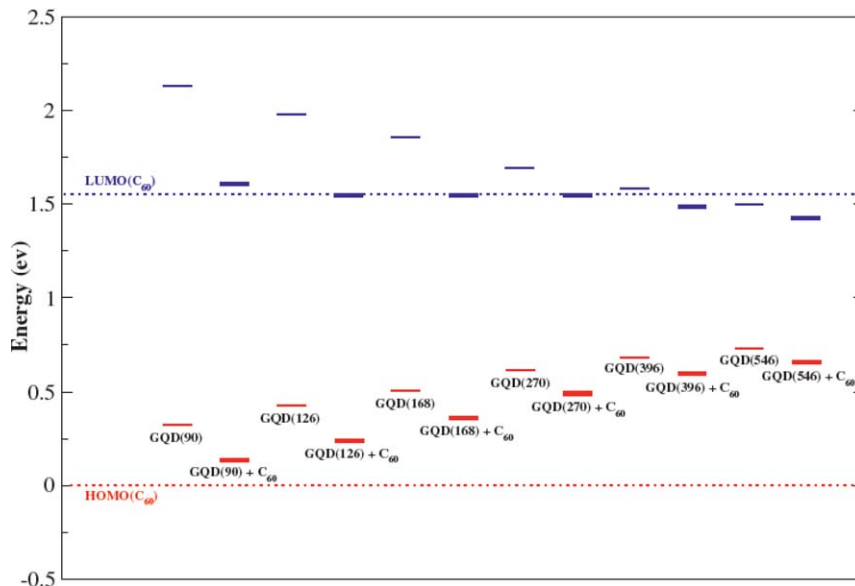
Fig. 17 HOMO (red) and LUMO (blue) energy alignment of the pure GQDs with different size and corresponding GQD–fullerene hybrid systems (bold). (Reprinted with permission from C. Chakravarty et al., Z. Phys. Chem., 2016, **230**, 777.)

clearly reveals that the energy difference between the LUMO of the GQD and the LUMO of the acceptor fullerene decreases with increasing the size of the GQD. So, we may conclude that the electron injection rate will be higher for the nanohybrids with smaller GQDs and it decreases with increasing the size of the GQD. The position of energy levels of the hybrid systems and the individual components show that for nanohybrids with smaller GQDs, both the HOMO and LUMO energies of the GQD lie above the HOMO and LUMO of the fullerene, respectively and thus form type-II energy alignment. But, with increasing the size of the GQD, the LUMO energy shifts to lower energy and that of the HOMO shifts to higher energy and after a certain size of the GQD, the LUMO of the GQD passes over the LUMO energy of the fullerene forming type-I energy alignment. So, with increasing the size of the GQD we notice a transition from type-II system to type-I.

To get support of the type-II band alignment as discussed above, we have shown the HOMO, LUMO charge densities of six nanohybrids with GQD of different sizes in Fig. 18. The figure clearly shows that for GQD–fullerene nanohybrids with smaller GQDs (e.g. GQDs with 90, 126, 168, 270 and 346 carbon atoms), the HOMO is distributed over only GQDs and the LUMO is localized only on fullerene molecules. For these systems, the charge carriers are localized on different components of the hybrid systems thus there is distinct charge separation for these systems. So the life time of the charge carriers will be long and accordingly the rate of recombination of the charge carriers will be low. Among the different nanohybrids, only the nanohybrids with the largest GQDs (GQDs with 546 number of carbon atoms) show type-I band alignment, both the
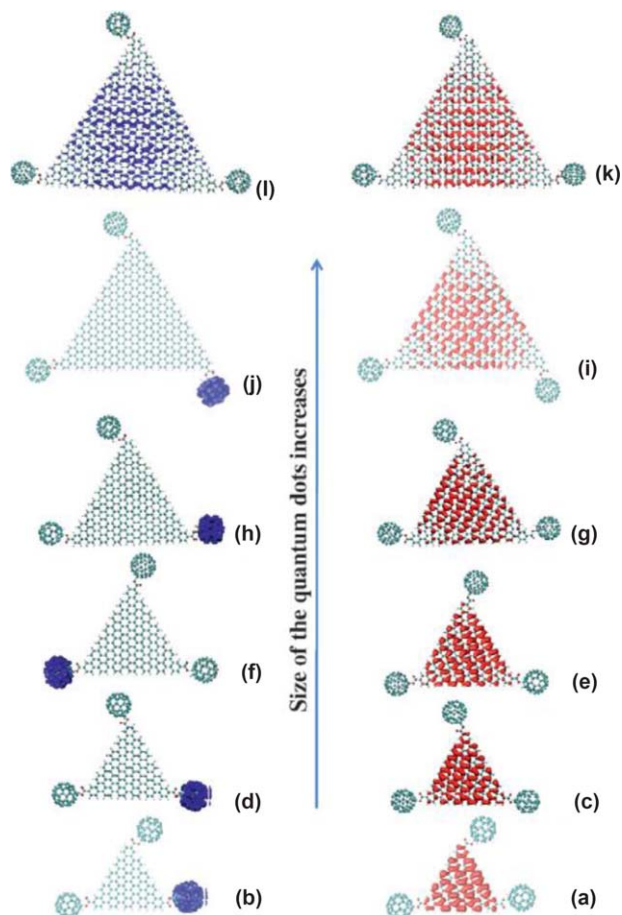
Fig. 18 The HOMO (red) and the LUMO (blue) charge densities of GQD–fullerene nanohybrid systems with number of carbon atoms 90 (a) and (b), 126 (c) and (d), 168 (e) and (f), 270 (g) and (h), 396 (i) and (j), 546 (k) and (l) in GQDs, respectively. Isosurface value $= 0.0001$ e$Å^{-3}$ is used for plotting. (Reprinted with permission from C. Chakravarty *et al.*, *Z. Phys. Chem.*, 2016, **230**, 777.)

HOMO and the LUMO charge densities are distributed only on GQDs. So, our theoretical study suggests that GQD–fullerene nanohybrids with smaller sized GQD may be a good candidate for designing solar cells.

3.6 CdTe nanotube–fullerene hybrid nanostructures

We have already discussed that the research on fullerene based nanohybrids has a phenomenal growth during the last few years. This is primarily because of the fact that, fullerenes have excellent exciton mobility, the relatively large exciton diffusion length and the high electron affinity.[6,54–58,137–140] These properties make fullerenes a good electron acceptor and electron-transporting material and thus very effective for constructing hybrid nanostructures. There are a number of experimental studies establishing that the electrostatic binding of C_{60} and CdTe NPs facilitates the rapid photo-induced ET and thus serves as a

potent alternative in fabricating photovoltaic devices.[6,7,56,117] In this section we will discuss the results of electronic structure calculations of CdTe nanotube (CdTeNT)–fullerene hybrid nanostructures to design a nanotube sensitized solar cell. The emphasis of our discussion will be on the relative positions of the frontier energy levels of the nanohybrid systems and how the variation of the various components of the hybrid systems influence the ET from the nanotube to the fullerene.[141]

The CdTeNT–fullerene hybrid nanostructures are modelled as follows. A thiol derivative of the fullerene enables covalent linking by the S atom of the –SH group to the Cd site of the 1010 outer surface of the CdTeNTs. This kind of covalent linking of the thiol derivative of C_{60} and the CdTeNT has already been demonstrated by the recent experimental work of Bang et al.[6] on a CdSeQD–fullerene hybrid nanocomposite. We have used 3 units of the CdTeNTs periodically repeated along the z-axis containing 144, 270 and 432 CdTe pairs per unit cell for 2WNT, 3WNT and 4WNT, respectively.

To understand how the density of states (DOSs) of the clean nanotube (2WNT) is being modified in the CdTeNT–fullerene (2WNT–C_{60}) hybrid nanostructures we have shown the DOSs of 2WNT–C_{60} hybrid nano-structure along with the isolated 2WCdTeNT in Fig. 19. The figure clearly reveals that a new peak appears above the Fermi level in the nanohybrid which is attributed to the LUMO of the C_{60}–thiol derivative. To have a detailed understanding of the different orbital contributions of the Cd, Te and C atoms we have shown the projected density of states (PDOS) of the 2WNT–C_{60} hybrid systems in the same figure. The figure clearly indicates that the major contribution to the VBT in the 2WNT–C_{60} system
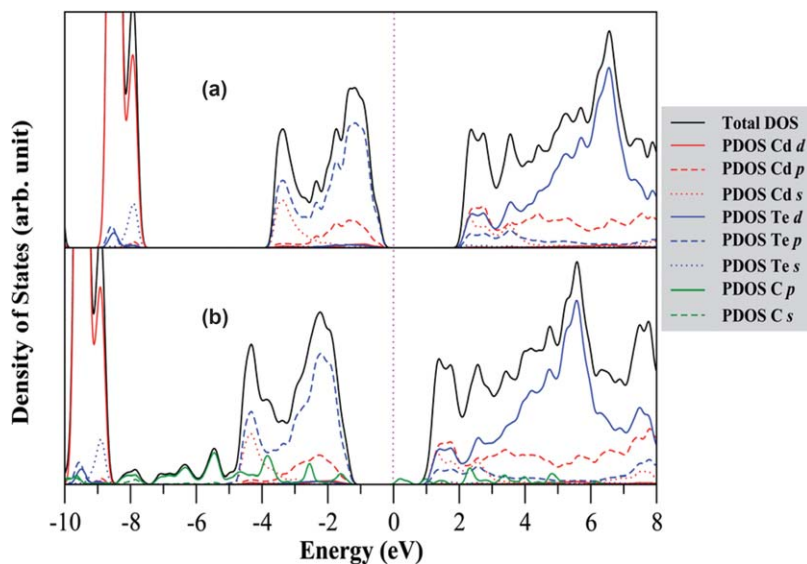

Fig. 19 The projected local density of states of the (a) 2WNT and (b) 2WNT–C_{60} showing the contributions from different orbitals. The value of Gaussian smearing used to plot the DOS is 0.1. The zero of the energy is set at the Fermi energy in each case. (Reprinted from S. Sarkar et al., RSC Adv., 2014, 4, 14673 with permission from Royal Society of Chemistry.)
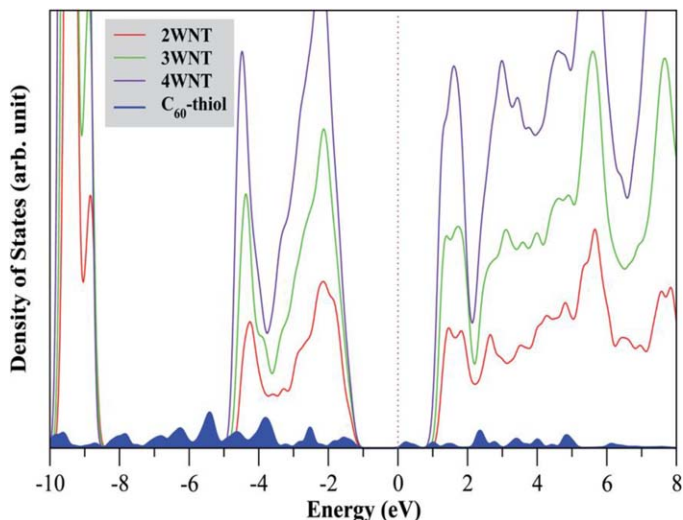
Fig. 20 The projected local density of states of 2WNT, 3WNT, 4WNT and C_{60} in the NT–C_{60} hybrid systems. The value of Gaussian smearing used to plot the DOS is 0.1. The zero of the energy is set at the Fermi energy. (Reprinted from S. Sarkar *et al.*, *RSC Adv.*, 2014, **4**, 14673 with permission from Royal Society of Chemistry.)

comes from the p orbitals of the Te atoms of the nanotube while the p orbitals of the C atoms of the fullerene have major contributions to the CBM.

To understand the effect of nanotube thickness on the electronic energy levels we have shown in Fig. 20, the projected density of states of 2WNT (red), 3WNT (green), 4WNT (violet) and C_{60} (blue shaded) in the NT–C_{60} hybrid systems. The figure shows that the position of the valence band top of the hybrid nanostructures is fixed; however, the position of the conduction band minimum of the hybrid nanostructures is shifted towards lower energies with the increase in the wall thickness of the CdTeNT and accordingly the band gap of these nanohybrids decreases with the increase in the wall thickness. In this nanohybrids like others discussed earlier the difference between the CBM of the NT and the LUMO of the C_{60} decreases with the increasing thickness of the NTs. So, following Marcus theory[114] we could expect a faster ET rate for thin walled CdTe NTs as compared to thick-walled NTs when they are coupled with the C_{60}.

The spatial distribution of the charge densities for the VBT and CBM of CdTe 2WNT–C_{60} nanohybrid at the Γ point are shown in Fig. 21. The figure shows that the VBT charge density (Fig. 21a) of the hybrid systems is localized on the CdTeNT while the CBM charge density (Fig. 21b) of the hybrid systems is localized on the C_{60} only. From the detailed analysis of orbital contributions to the charge densities of VBT and CBM we have seen that the major contribution to VBT comes from the Te p orbitals and little contribution from the Cd p orbitals while the CBM has all contributions from the C p orbitals. The VBT of the 2WNT–C_{60} hybrid systems has therefore the characteristic feature of the CdTeNT and the CBM of the 2WNT–C_{60} hybrid systems has the characteristic feature of C_{60}.
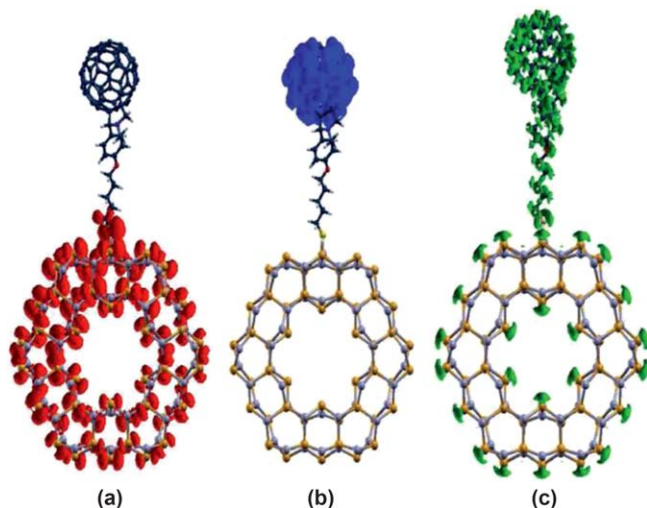
Fig. 21 Isosurface plots of the charge density (squared wave functions) of the (a) valence band top (VBT) and (b) conduction band minimum (CBM) for the 2WNT$-C_{60}$ hybrid system at the Γ point. (c) Isosurface plots of the charge density difference distribution for the 2WNT$-C_{60}$ hybrid system at the G point. The green indicates an increase in the charge density. The isosurface value used for (a) and (b) was 0.0002 e$Å^{-3}$ and for (c) 0.01 e$Å^{-3}$. (Reprinted from S. Sarkar *et al.*, *RSC Adv.*, 2014, **4**, 14673 with permission from Royal Society of Chemistry.)

So, this study suggests the type-II hetero-junctions for CdTeNT$-C_{60}$ nanohybrids. The consequent charge separation between the electrons and holes at the interface of the nanohybrids therefore leads to the localization of the electrons on the C_{60} and holes on the CdTeNT. This spatial separation of the charge carriers is highly desirable to slow down the rate of electron–hole recombination in nanotube sensitized solar cells. In Fig. 21(c) we have shown the charge density difference which is defined as the difference between the total charge density and the atomic charge densities of the 2WNT$-C_{60}$ hybrid system. The figure clearly shows that the thiolated C_{60} part of the hybrid system gains more electrons as compared to the CdTeNT in the nanohybrid. So, the figure suggests that there is an efficient ET from CdTe to C_{60} in 2WNT$-C_{60}$ donor–acceptor material. Several experimental groups have shown that there is a strong emission quenching in fullerene–CdTe nanoparticle hybrid system as compared to isolated CdTe NPs.[6,44,49,90,115] This quenching of emission is evidently due to the charge transfer from the NPs to the fullerenes. So, the theoretical prediction is in good agreement with the experimental observation.

Now, we would like to discuss the electronic structure of the CdTeNT–fullerene nanohybrids with a series of fullerenes *e.g.*, C_{60}, C_{68}, C_{70}, C_{76}, C_{80} and C_{84}. The band structures and density of states of these nanohybrids are shown in Fig. 22. The flat bands appearing in the band gap region of the hybrid systems 2WNT$-C_{68}$ (b), 2WNT$-C_{70}$ (c), 2WNT$-C_{76}$ (d) and 2WNT$-C_{80}$ (e) are due to the molecular orbitals of the fullerenes. The hybrid systems 2WNT$-C_{60}$ (a) and 2WNT$-C_{84}$ (f) have no flat bands
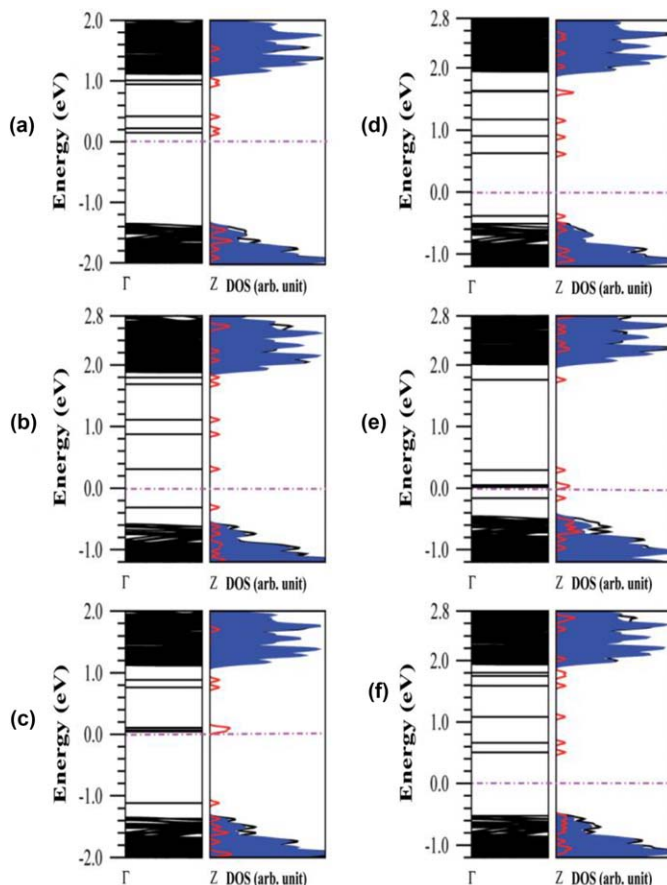
Fig. 22 The left panel of (a), (b), (c), (d), (e) and (f) represent the electronic band structures of the 2WNT–C_{60}, 2WNT–C_{68}, 2WNT–C_{70}, 2WNT–C_{76}, 2WNT–C_{80} and 2WNT–C_{84} nanohybrids, respectively. The right panel of (a), (b), (c), (d), (e) and (f) represent the density of states (DOS) of the 2WNT–C_{60}, 2WNT–C_{68}, 2WNT–C_{70}, 2WNT–C_{76}, 2WNT–C_{80} and 2WNT–C_{84} nanohybrids, respectively. The shaded area in the DOS reflects the projected local density of states of the CdTe NT of the 2WNT–C_n ($n = 60$, 68, 70, 76, 80 and 84) hybrid systems and the red peaks are the projected local density of states of C_n–thiol ($n = 60$, 68, 70, 76, 80 and 84) of the 2WNT–C_n ($n = 60$, 68, 70, 76, 80 and 84) hybrid systems. The value of Gaussian smearing used to plot the DOS is 0.02. The zero of the energy is set at the Fermi energy in each case. (Reprinted from S. Sarkar *et al.*, *RSC Adv.*, 2014, **4**, 14673 with permission from Royal Society of Chemistry.)

below the Fermi levels. Thus, the VBT and CBM of the 2WNT–C_{60} and 2WNT–C_{84} hybrid systems have the characteristic feature of the CdTeNT and fullerenes, respectively. But, both the VBT and CBM of the 2WNT–C_{68} (b), 2WNT–C_{70} (c), 2WNT–C_{76} (d) and 2WNT–C_{80} (e) hybrid systems have the characteristic feature of the fullerenes. So, two hybrid systems 2WNT–C_{60} and 2WNT–C_{84} would exhibit type-II band energy alignment. The presence of new states induced by the fullerenes in the hybrid NT–Fullerene nanostructures has a great impact on the characteristic properties of the final device, such as the polarization offset of the photodiodes or the open-circuit voltage in excitonic solar cells.

3.7 ZnO/ZnS nanowire-based dye-sensitized solar cell

Law et $al.$[31,142] have pointed out that nanoparticle-based DSSCs rely on trap-limited diffusion for electron transport, a slow mechanism that can limit device efficiencies. These authors have introduced a new type of the dye-sensitized cell in which the traditional nanoparticle film is replaced by crystalline ZnO nanowires (ZnONWs). These authors showed that NWs have excellent electron transport capabilities because of internal electric fields in the direction of the c-axis of wurtzite ZnO crystals and this causes the suppression of recombination for injected electrons from surrounding electrolytes. Although, ZnONW enables rapid electron transport with a large electron diffusion coefficient of 1.8×10^{-3} cm^2 s^{-1}, the cell efficiency of the ZnONW-based DSSCs is still lower than that of the other nanoparticle-based ones.[36] The low conversion efficiency is because of the recombination of the injected electrons into the ZnO with either the dye or the redox electrolyte.[37,143] One of the most promising approaches to improve the efficiency is to use a core/shell NW. The formation of core/shell NWs results an energy barrier between the core and the shell thereby reducing the recombination rates. The presence of a shell also removes the surface states thereby decreasing the dark current and increasing the V_{oc} (open-circuit voltage). In the recent literature, there are extensive reports that a ZnS layer is a viable option as a shell layer on ZnONW for the application of solar cells.[11,33–35,37,144] The oxygen vacancies of ZnONWs, which were responsible for the high recombination rate of electrons in DSSCs, have been filled up by S atoms of the ZnS layer. The formation of a shell of ZnS over the core ZnONW results a type-II band alignment of heterostructures through the induction of charge separation at the interface of the two different materials. Thus the electrons and holes are localizing in different regions of the nanostructures and this reduces the chance of recombination of the charge carriers. The type-II band alignment also forms appropriate conduction band edge lines to facilitate ET. There are extensive reports on the synthesis and photovoltaic properties of ZnO/ZnS core/shell NWs.[29,32–35,37,144–146] The recent experimental work of Yu et $al.$[37] on ZnO/ZnS heteronanostructures showed that the variation of shell thickness and roughness of the ZnS shell has remarkable influence on the photovoltaic properties of ZnO-based DSSCs. Zhu et $al.$[41] in their study on CdSe/ZnS core/shell nanosystems have demonstrated that both the charge separation and recombination rates can be controlled by altering the shell thickness and the nature of the shell. By using band-corrected pseudopotential density functional theory, Schrier et $al.$[30] studied optical absorption and carrier localization in NW-based ZnO/ZnS heterostructures. In this section we would like to explore the electronic structure of ZnO/ZnS core/shell NWs and ZnO/ZnS core/shell NW–dye nanohybrids to understand the suitability of their application in solar cells.

We have studied two different types of ZnO/ZnS core/shell NW heterostructures, one is with fixed core radii and increasing shell thickness (FCIS) and the other type of heterostructures is with increasing shell thickness and decreasing core radii (ISDC) gradually, keeping a fixed core plus shell diameter. We denote these as (ZnO)$_3$/(ZnS)$_1$, (ZnO)$_3$/(ZnS)$_2$,

$(ZnO)_3/(ZnS)_3$, $(ZnO)_3/(ZnS)_4$ and $(ZnO)_5/ZnS)_1$, $(ZnO)_4/(ZnS)_2$, $(ZnO)_2/$ $(ZnS)_4$, and $(ZnO)_1/(ZnS)_5$, where the subscripts indicate the number of atomic layers in the cross section in the ZnO core and ZnS shell regions, respectively.

We have shown the band alignment of both FCIS and ISDC systems in Figs. 23 and 24, respectively. In the insets of both the figures we have shown the variation of the band gap with varying core/shell thickness for both FCIS and ISDC systems. The band gap for FCIS system decreases with increasing shell thickness. The band alignment of the FCIS system clearly shows that the VBM of the hybrid system remains almost at the same position or close to the VBM of the ZnO NW. The CBM of the hybrid system becomes more negative with increasing ZnS shell thickness. So, the band gap for FCIS system decreases with increasing shell thickness. Thus, the absorption and emission of the exciton band are red shifted with increasing the thickness of the shell. This is due to the increased leakage of the exciton wave function from the ZnO core to the ZnS shell. The red shift in the band gap for ZnO/ZnS core/shell NW with increasing shell thickness is in good agreement with experimental studies of this particular system.[29,33,144,147] For ISDC systems, the VBM of the hybrid systems slightly becomes more positive as compared to ZnONWs while the energies of the CBM shift to lower energies as compared to CBM of ZnO NWs. So, here also a red shift occurs in the band gap as we go from $(ZnO)_5/(ZnS)_1$ to $(ZnO)_2/(ZnS)_4$. As the size of ISDC NW systems are same, the variation of the band gap for these systems is not due to the quantum
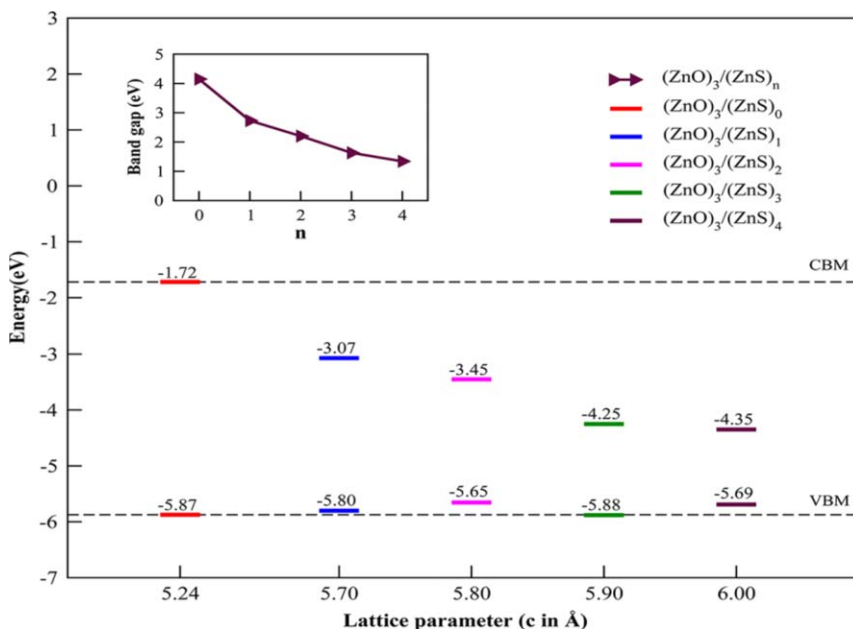
Fig. 23 Band alignment of FCIS ZnO/ZnS heterostructure NW systems. The dotted lines indicate the VBM and CBM energy levels of ZnO NW without ZnS shell. In the inset, the variation of band gap is shown as a function of ZnS layer. (Reprinted with permission from S. Saha *et al.*, *J. Phys. Chem. C*, 2013, **117**, 15890. © 2013, American Chemical Society.)
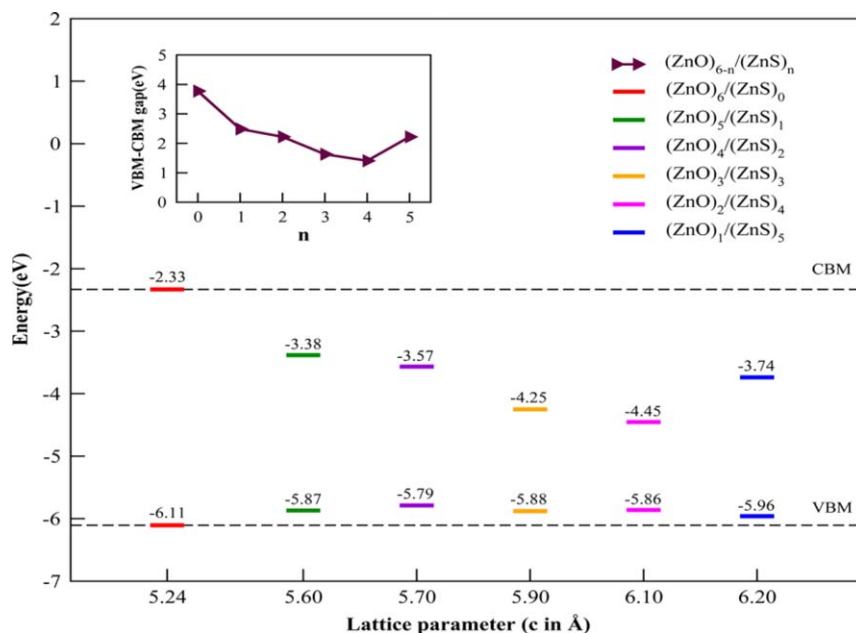
Fig. 24 Band alignment of ISDC ZnO/ZnS heterostructure NW systems. The dotted line indicates the VBM and CBM energy levels of ZnO NW without ZnS shell. Variation of band gap is shown in the inset. (Reprinted with permission from S. Saha *et al.*, *J. Phys. Chem. C*, 2013, **117**, 15890. © 2013, American Chemical Society.)

confinement effect, rather it is the stress of the NW that controls the variation of the band gap. The increased heterogeneity in the ISDC systems causes the increase in stress as the value of "n" increases from 0 to 3 for $(ZnO)_{6-n}/(ZnS)_n$ core/shell NWs. The increased stress causes the decrease in CB energies resulting a decrease in the band gap. So, one can tailor the band alignment of ZnO/ZnS NW heterostructures by either varying the ZnS shell thickness or by varying both ZnO core diameter and ZnS shell thickness. The understanding of this energy level alignment plays a crucial role in choosing a perfect dye molecule as the difference between the conduction band energy levels of heterostructure NWs and the dye molecule serves as a driving force for the interparticle ET. Hence, for a dye molecule whose LUMO energy remains above the CBM of the hybrid system, the charge injection rate increases with the increasing ZnS shell thickness of the core/shell NW.

To assess whether there is a charge separation at the interface of the heterostructures, the wave functions corresponding to the VBM (upper row) and CBM (lower row) of the FCIS core/shell NWs for three different shell thickness at the Γ point are shown in Fig. 25 (the first, second, and third columns are for $(ZnO)_3/(ZnS)_1$, $(ZnO)_3/(ZnS)_2$, and $(ZnO)_3/(ZnS)_3$ systems, respectively). The same for three different ISDC core/shell NW systems are shown in Fig. 26 (the first, second, and third columns are for $(ZnO)_5/(ZnS)_1$, $(ZnO)_4/(ZnS)_2$, and $(ZnO)_2/(ZnS)_4$ systems, respectively). Both Figs. 25 and 26 clearly show that the VBM and CBM are localized on different components namely ZnS shell and ZnO core, respectively.
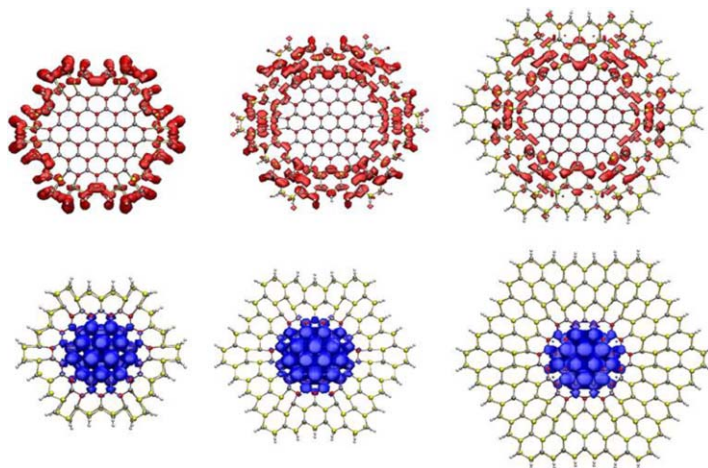
Fig. 25 Band-edge wave functions of the ZnO/ZnS FCIS heterostructure NWs. First row represent the VBM (red) and second row for CBM (blue). Gray, yellow, red, and white spheres represent Zn, S, O, and H atoms, respectively. (Reprinted with permission from S. Saha *et al.*, *J. Phys. Chem. C*, 2013, **117**, 15890. © 2013, American Chemical Society.)
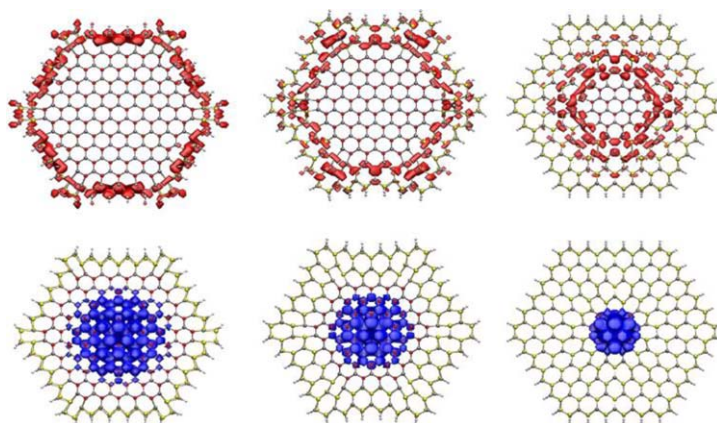
Fig. 26 Band-edge wave functions of the ZnO/ZnS ISDC heterostructure NWs. First row represent the VBM (red) and second row for CBM (blue). Gray, yellow, red, and white spheres represent Zn, S, O, and H atoms, respectively. (Reprinted with permission from S. Saha *et al.*, *J. Phys. Chem. C*, 2013, **117**, 15890. © 2013, American Chemical Society.)

Thus we see that this particular hetero NW system form type-II band alignment irrespective of the shell thickness and also of the variation both the core radius and shell thickness of ZnO/ZnS core/shell NW. The localization of the electron and the hole in core and the shell material suggests the possibility of charge separation of an exciton into an electron and hole in this particular nanoheterostructure that would result in an envelope-function-induced dark exciton. This dark exciton can suppress electron–hole recombination, resulting in an enhancement of the carrier collection efficiency, making these NW systems very useful for solar cell applications.[30,41,148]

To test the applicability of these systems in more realistic environment, we construct some ZnO/ZnS NW–dye hybrid nanostructures and discuss their electronic properties. We choose the anthraquinone-2,3-dicarboxylic acid (AQ) dye molecule and it is adsorbed on the surface of the ZnO/ZnS NW. The performance of DSSC depends very much on the relative position of the energy bands of semiconductors and dye molecules. The injection efficiency of photogenerated electrons from the dye molecule to the semiconductor heterostructures largely depends on the relative position of the VBM and CBM of the semiconductors and the HOMO and LUMO of dye molecules.[148–152] We herein present theoretical results of the effect of the ZnS shell thickness on the electron injection efficiency, charge separation, and recombination dynamics of ZnO/ZnS core/shell NW–AQ composite systems. We would like to examine how the energy levels (CBM and VBM) of the dye adsorbed ZnO/ZnS nanoheterosystems are modulated by either changing the thickness of the ZnS shell of the ZnO/ZnS core/shell NW and also with AQ dyes with different functional groups.[153]

The adsorption of AQ on the ZnS shell of a CdSe/ZnS core/shell QD was investigated earlier by Zhu *et al.*[41] These authors have shown that the recombination rates decreases exponentially with the shell thickness. Now to understand what will be the situation for ZnO/ZnS NW systems, the band alignment of two representative ZnO/ZnS core/shell hetero systems with fixed core diameter and increasing shell thickness along with a few AQ dyes with different functional groups are shown in Fig. 27(a) and (b), respectively. The band alignment of few hybrid systems (hetero structure NWs and AQ–molecules) are also shown in the same figure. To understand the effect of different functional groups we have considered hybrid systems: AQ, 7-amino-AQ (NH_2-AQ) and 7-nitro-AQ (NO_2–AQ) dyes adsorbed on both $(ZnO)_2/(ZnS)_1$ and $(ZnO)_2/(ZnS)_2$ core/shell hetero structure NWs. The analyses of the Fig. 27(a) and (b) reveal that the ZnS shell has a pronounced effect on the conduction band of the nanohybrids. The figures show that with increasing shell thickness, the CB states, primarily localized in the ZnO core, extend to the ZnS shell and are shifted to lower energies. This lowering of the CB energy with increasing shell thickness is in good agreement with the study of Yang *et al.* and Lo *et al.*[148,154] The figure also shows that the position of the conduction band edge very much depends on the shell thickness and also on the nature of the functional group of the dye molecule. So, the electron injection rate which depends on the energy difference between LUMO of the dye and CBM of the NW can be controlled through the variation of either thickness of the shell and the nature of the functional groups. The band gap of the ZnO/ZnS NW–AQ dye is smaller than the ZnO/ZnS NW. So, as it is found for other hybrid systems, the core/shell NW–dye composite system would show a red shift in their absorption spectra with increasing shell thickness of ZnO/ZnS core/shell NW. This pronounced red shift with increasing shell thickness is in good agreement with recent experimental studies on these core/shell NW systems and other related systems.[34,35,144,147] Now, we try to understand which NW–dye hybrid systems will form type-II band energy alignment. As it is evident from Fig. 27(a), for core/shell NWs of smaller shell thickness,
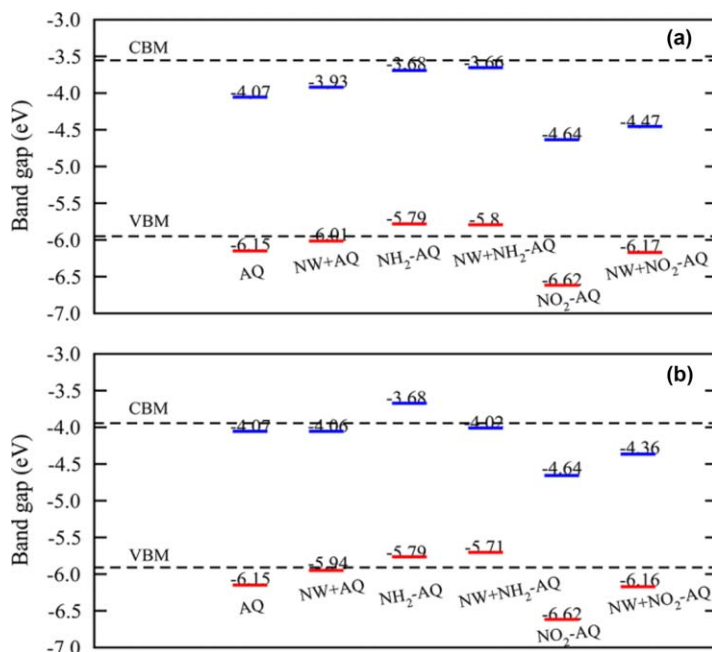
Fig. 27 VBM (red) and CBM (blue) energy alignment of NW–dye composite systems (ZnO/ZnS core/shell heterostructure NWs and AQ-molecules with different functional group). (a) $(ZnO)_2/(ZnS)_1$ + dye molecules and (b) $(ZnO)_2/(ZnS)_2$ + dye molecules. Dotted lines represent the VBM and CBM of corresponding ZnO/ZnS core/shell NWs without dye molecule. (Reprinted with permission from S. Saha *et al.*, *J. Phys. Chem. C*, 2013, **117**, 15890. © 2013, American Chemical Society.)

all systems except NH_2–AQ system show type-II band alignment. But for ZnO/ZnS core/shell NWs with larger ZnS shell thickness, all systems including –NH_2–AQ have type-II band alignment. So, for NH_2–AQ dye–ZnO/ZnS NW composite system a type-I to type-II band alignment crossover occurs with the increase of the ZnS shell thickness. The presence of electron withdrawing groups in AQ dyes, *i.e.*, for NO_2–AQ, the LUMO of the dye molecule goes to more positive (on the NHE scale) and the AQ acid derivative with electron withdrawing group always forms a type-II band alignment.

From the several systems studied here, we found that the hybrid system made of NH_2–AQ-acid dye-$(ZnO)_2/(ZnS)_2$ NW has the proper band energy alignment for use in NW–dye sensitized solar cells. The LUMO of this dye molecule lies above the CBM of the ZnO/ZnS NW, thus facilitates electron injection from the dye to the NW. For this particular system, the electron and hole wave functions are on different parts of the hybrid system that may suppress the electron hole recombination. As the thickness of the ZnS shell in ZnO/ZnS core/shell NWs increases, the CBM of the NW shifts downward so the ΔG for ET from the dye to the NW $[E_{LUMO}(dye) - E_{CBM}(NW)]$ increases and ΔG for recombination $[E_{CBM}(NW) - E_{HOMO}(dye)]$ decreases. The increased ΔG for ET and decreased ΔG for recombination of electrons will be beneficial for charge transfer and suppression of charge recombination, respectively.
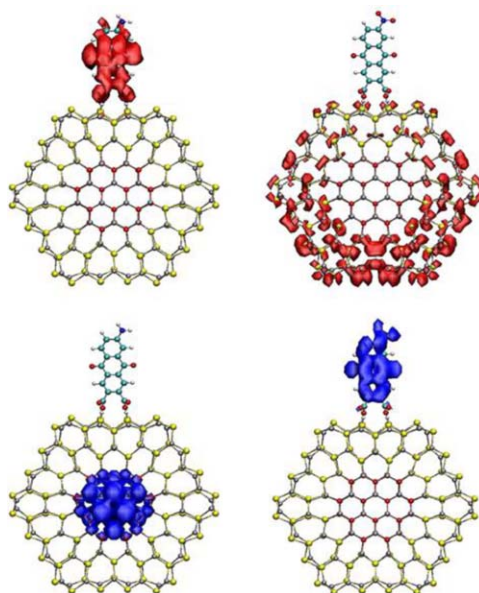
Fig. 28 Band-edge wave functions of the composite nanosystems ((ZnO)$_2$/(ZnS)$_2$ core–shell NW + NH_2–AQ and NO_2–AQ dye molecules). Upper row represents VBM (red) and lower row represents CBM (blue). Left column for (ZnO)$_2$/(ZnS)$_2$ + NH_2–AQ-acid dye and right column for (ZnO)$_2$/(ZnS)$_2$ + NO_2–AQ-acid dye. Gray, yellow, red, blue, and white spheres represent Zn, S, O, N, and H atoms, respectively. (Reprinted with permission from S. Saha *et al.*, *J. Phys. Chem. C*, 2013, **117**, 15890. © 2013, American Chemical Society.)

To understand the type-II behaviour and the localization of charge carriers into different parts, the band-edge wave functions of the hybrid nanosystems ((ZnO)$_2$/(ZnS)$_2$ core/shell NW + NH_2–AQ and NO_2–AQ dyes) are shown in Fig. 28. For the hybrid system with AQ molecules, having an electron donating group such as –NH_2, the VBM is distributed over the dye molecules and the CBM is localized on the ZnO core region. But for hybrid systems having an electron withdrawing group such as –NO_2 in AQ acid molecule, the VBM is distributed throughout the outer ZnS shell while the CBM is localized on the adsorbed AQ molecules. The rate of charge recombination is controlled by the extent of overlap of the VBM and CBM wave functions of the hybrid systems. From the figure it is clearly seen that the electron and hole wave functions are spatially separated from each other so there is negligible overlap for the ZnO/ZnS NW–AQ dye molecule hybrid system and this also depends on the thickness of the ZnS shell. Therefore, we suggest that, the charge recombination rate is low for the above mentioned NW–dye hybrid systems and will be found suitable for application in dye-sensitized solar cells.

4 Conclusions and future perspectives

The research on nanoscale materials to understand the feasibility of their use as building blocks in designing solar cells is an active field of research. In this context, hybrid nanostructures composed of two

different materials have been the focus of many research groups. Several experimental groups have synthesized different hybrid nanostructres and studied the dynamics of ET using the techniques of time-resolved ultra-fast spectroscopy. There are also numerous theoretical studies addressing the electronic structure of this class of materials. In this chapter, we have presented some important results of our electronic structure calculations of different hybrid nanostructures. We analyzed the electronic energy levels of the hybrid systems as a function of the size of the QD, the thickness of the nanotube or thickness of the shell of the core/shell NW, *etc.* and tried to suggest systems with improved photo-voltaic performance that may find application in solar cells. Although such calculations are very useful for device optimization, the major goal of our research remains incomplete without being able to perform dynamical studies of such large realistic systems. As already discussed, the primary mechanism of the solar-to-electric energy conversion is the photo-induced separation of electrons and holes. The subsequent charge separation at an interface of two components of a hybrid material is one of the key factors that define the overall efficiency of energy conversion. This process is intrinsically non-equilibrium and often occurs at the ultrafast timescale, rendering classical reaction rate theories based on the quasi-equilibrium assumption inapplicable. Moreover, the commonly utilized Born–Oppenheimer (adiabatic) approximation is no longer valid. The understanding of non-equilibrium processes involved in the photo-induced charge separation and transport requires explicit time-domain molecular dynamics simulations that account for nonadiabatic effects – the nonadiabatic molecular dynamics (NA-MD). The approach allows one to treat all relevant processes and to take the full complexity of the material into account, including shape, defects, surface passivation, as well as the environment effect at varying electrochemical conditions. Although various theoretical groups now a days perform simulations on different nanoscale materials[155–159] such as inorganic–organic nano-hybrid, semiconductor QDs to capture various photoinduced phenomena, still, the efforts given to investigate from theoretical side are limited by the system size. The efficient approach to study such nonradiative transitions which are responsible for fundamental photochemical processes is based on the mixed quantum-classical dynamics. In this approach, the nuclear motion is described by classical trajectories propagated using quantum chemical ab initio molecular dynamics ''on the fly'', and by combining with the stochastic Tully's surface hopping (TSH) procedure one can give a detailed description of those nonadiabatic electronic transitions.[155,157,158] For the propagation of the classical nuclear trajectories, forces as well as the nonadiabatic couplings are needed which can be calculated by several approximations. Among them, the TDDFT represents an efficient, gener-ally applicable method which allows one to treat all relevant processes and to take the full complexity of the material into account, including shape, defects, surface passivation, as well as the environment effect at varying electrochemical conditions.

The group of Oleg Prezhdo developed a semi-classical treatment[155–158] which has successfully been employed in various nano structure systems

to simulate different nonequlibrium phenomena, such as charge transfer, energy transfer, and thermal relaxation of excited electrons and so on.[78,160,161] It implements a number of basic and more advanced functionalities, including fewest-switches surface hopping (FSSH), decoherence-induced surface hopping (DISH), multi electron adiabatic representation of the time-dependent Kohn-Sham (TD-KS) equations, and the direct simulation of photoexcitation *via* explicit light matter interaction. The classical path approximation (CPA) achieves additional and considerable computational savings; however, the computational demand is still limited by the scaling of the *ab initio* DFT methods for electronic structure calculations. Pal *et al.*[162] have very recently incorporated tight binding approaches for obtaining the nonadiabatic electronic surfaces and the coupling between a pair of these surfaces (DFTB–NAMD) which decreases the computational demand quite in large extent and thus it is expected that it would be applicable for very large systems with larger time scale simulations. As an illustration of the applicability, the authors have studied different ultrafast nonradiative relaxation dynamics of two different systems, CdSe QD and CNT. The computed time scales for the various ET processes, obtained from this method match well with the experimental results as well as other theoretical methods with plane-wave basis justifying the accuracy of the methodology. One can perform real time simulations of hybrid nano-structures of relatively large size by using the DFTB–NAMD method and be able to calculate the rate constants of electron injection and electron–hole recombination. So, it is now possible to make a quantitative prediction of the photovoltaic efficiency of solar cells.

Acknowledgements

References

1   B. O'Regan and M. Grätzel, *Nature*, 1991, **353**, 737.
2   M. K. Nazeeruddin, P. Pechy, T. Renouard, S. M. Zakeeruddin, R. Humphry-Baker, P. Comte, P. Liska, L. Cevey, E. Costa and V. Shklover, *J. Am. Chem. Soc.*, 2001, **123**, 1613.
3   M. K. Nazeeruddin, F. De Angelis, S. Fantacci, A. Selloni, G. Viscardi, P. Liska, S. Ito, B. Takeru and M. Grätzel, *J. Am. Chem. Soc.*, 2005, **127**, 16835.
4   M. A. Green, K. Emery, Y. Hishikawa, W. Warta and E. D. Dunlop, *Prog. Photovoltaics*, 2015, **23**, 1.
5   M. Grätzel, *J. Photochem. Photobiol., C*, 2003, **4**, 145.
6   J. H. Bang and P. V. Kamat, *ACS Nano*, 2011, **5**, 9421.
7   P. Brown and P. V. Kamat, *J. Am. Chem. Soc.*, 2008, **130**, 8890.
8   H. Choi, R. Nicolaescu, S. Paek, J. Ko and P. V. Kamat, *ACS Nano*, 2011, **5**, 9238.

9   P. V. Kamat, *J. Phys. Chem. C*, 2007, **111**, 2834.

10   P. V. Kamat, *J. Phys. Chem. C*, 2008, **112**, 18737.

11   A. Kongkanand, K. Tvrdy, K. Takechi, M. Kuno and P. V. Kamat, *J. Am. Chem. Soc.*, 2008, **130**, 4007.

12   I. Robel, V. Subramanian, M. Kuno and P. V. Kamat, *J. Am. Chem. Soc.*, 2006, **128**, 2385.

13   K. Tvrdy, P. A. Frantsuzov and P. V. Kamat, *Proc. Natl. Acad. Sci.*, 2011, **108**, 29.

14   K. Gopidas, M. Bohorquez and P. V. Kamat, *J. Phys. Chem.*, 1990, **94**, 6435.

15   V. Subramanian, E. Wolf and P. V. Kamat, *J. Phys. Chem. B*, 2001, **105**, 11439.

16   P. V. Kamat, *J. Phys. Chem. B*, 2002, **106**, 7729.

17   I. Robel, B. A. Bunker and P. V. Kamat, *Adv. Mater.*, 2005, **17**, 2458.

18   J. S. Manser and P. V. Kamat, *Nat. Photon.*, 2014, **8**, 737.

19   A. P. Alivisatos, *Science*, 1996, **271**, 933.

20   C. Schulz-Drost, V. Sgobba, C. Gerhards, S. Leubner, R. M. Krick Calderon, A. Ruland and D. M. Guldi, *Angew. Chem., Int. Ed.*, 2010, **49**, 6425.

21   S. Ren, L.-Y. Chang, S.-K. Lim, J. Zhao, M. Smith, N. Zhao, V. Bulovic, M. Bawendi and S. Gradecak, *Nano Lett.*, 2011, **11**, 3998.

22   E. Holder, N. Tessler and A. L. Rogach, *J. Mater. Chem.*, 2008, **18**, 1064.

23   W. U. Huynh, J. J. Dittmer and A. P. Alivisatos, *Science*, 2002, **295**, 2425.

24   M. Wright and A. Uddin, *Sol. Energy Mater. Sol. Cells*, 2012, **107**, 87.

25   R. Plass, S. Pelet, J. Krueger, M. Grätzel and U. Bach, *J. Phys. Chem. B*, 2002, **106**, 7578.

26   Z. Liu, Y. Sun, J. Yuan, H. Wei, X. Huang, L. Han, W. Wang, H. Wang and W. Ma, *Adv. Mater.*, 2013, **25**, 5772.

27   R. Zhou, R. Stalder, D. Xie, W. Cao, Y. Zheng, Y. Yang, M. Plaisant, P. H. Holloway, K. S. Schanze and J. R. Reynolds, *ACS Nano*, 2013, **7**, 4846.

28   B. Guzelturk, F. Menk, K. Philipps, Y. Kelestemur, M. Olutas, R. Zentel and H. V. Demir, *J. Phys. Chem. C*, 2016, **120**, 3573.

29   M.-Y. Lu, J. Song, M.-P. Lu, C.-Y. Lee, L.-J. Chen and Z. L. Wang, *ACS Nano*, 2009, **3**, 357.

30   J. Schrier, D. O. Demchenko and A. P. Alivisatos, *Nano Lett.*, 2007, **7**, 2377.

31   M. Law, L. E. Greene, A. Radenovic, T. Kuykendall, J. Liphardt and P. Yang, *J. Phys. Chem. B*, 2006, **110**, 22652.

32   Y. Hao, J. Pei, Y. Wei, Y. Cao, S. Jiao, F. Zhu, J. Li and D. Xu, *J. Phys. Chem. C*, 2010, **114**, 8622.

33   M. Ahmad, X. Yan and J. Zhu, *J. Phys. Chem. C*, 2011, **115**, 1831.

34   K. Wang, J. Chen, Z. Zeng, J. Tarr, W. Zhou, Y. Zhang, Y. Yan, C. Jiang, J. Pern and A. Mascarenhas, *Appl. Phys. Lett.*, 2010, **96**, 123105.

35   J. Chung, J. Myoung, J. Oh and S. Lim, *J. Phys. Chem. C*, 2010, **114**, 21360.

36   C.-H. Ku and J.-J. Wu, *Appl. Phys. Lett.*, 2007, **91**, 093117.

37   X.-L. Yu, J.-G. Song, Y.-S. Fu, Y. Xie, X. Song, J. Sun and X.-W. Du, *J. Phys. Chem. C*, 2010, **114**, 2380.

38   S. V. Kilina, D. S. Kilin and O. V. Prezhdo, *ACS Nano*, 2008, **3**, 93.

39   J. M. Haremza, M. A. Hahn, T. D. Krauss, S. Chen and J. Calcines, *Nano Lett.*, 2002, **2**, 1253.

40   L. Sheeney-Haj-Ichia, B. Basnar and I. Willner, *Angew. Chem., Int. Ed.*, 2005, **44**, 78.

41   H. Zhu, N. Song and T. Lian, *J. Am. Chem. Soc.*, 2010, **132**, 15038.

42   C. Chakravarty, P. Ghosh, B. Mandal and P. Sarkar, *Z. Phys. Chem.*, 2016, **230**, 777.

43   Y. Chen, Z.-H. Huang, M. Yue and F. Kang, *Nanoscale*, 2014, **6**, 978.

44 J. Debgupta, S. Mandal, H. Kalita, M. Aslam, A. Patra and V. Pillai, *RSC Adv.*, 2014, **4**, 13788.

45 H. Gao, L. Wang, J. Zhao, F. Ding and J. Lu, *J. Phys. Chem. C*, 2011, **115**, 3236.

46 J. Geng and H.-T. Jung, *J. Phys. Chem. C*, 2010, **114**, 8227.

47 C. X. Guo, H. B. Yang, Z. M. Sheng, Z. S. Lu, Q. L. Song and C. M. Li, *Angew. Chem., Int. Ed.*, 2010, **49**, 3014.

48 Y. Guo, L. Deng, J. Li, S. Guo, E. Wang and S. Dong, *ACS Nano*, 2011, **5**, 1282.

49 S. Kaniyankandy, S. Rawalekar and H. N. Ghosh, *J. Phys. Chem. C*, 2012, **116**, 16271.

50 D. Kiessling, R. D. Costa, G. Katsukis, J. Malig, F. Lodermeyer, S. Feihl, A. Roth, L. Wibmer, M. Kehrer and M. Volland, *Chem. Sci.*, 2013, **4**, 3085.

51 R. Long, N. J. English and O. V. Prezhdo, *J. Am. Chem. Soc.*, 2012, **134**, 14238.

52 G. B. Markad, S. Battu, S. Kapoor and S. K. Haram, *J. Phys. Chem. C*, 2013, **117**, 20944.

53 Y. Xu, Z. Liu, X. Zhang, Y. Wang, J. Tian, Y. Huang, Y. Ma, X. Zhang and Y. Chen, *Adv. Mater.*, 2009, **21**, 1275.

54 V. V. Chaban, V. V. Prezhdo and O. V. Prezhdo, *J. Phys. Chem. Lett.*, 2012, **4**, 1.

55 D. M. Guldi, *Chem. Commun.*, 2000, 321.

56 D. M. Guldi, I. Zilbermann, G. Anderson, N. A. Kotov, N. Tagmatarchis and M. Prato, *J. Am. Chem. Soc.*, 2004, **126**, 14340.

57 E. S. Shibu, A. Sonoda, Z. Tao, Q. Feng, A. Furube, S. Masuo, L. Wang, N. Tamai, M. Ishikawa and V. Biju, *ACS Nano*, 2012, **6**, 1601.

58 M. H. Stewart, A. L. Huston, A. M. Scott, E. Oh, W. R. Algar, J. R. Deschamps, K. Susumu, V. Jain, D. E. Prasuhn and J. Blanco-Canosa, *ACS Nano*, 2013, **7**, 9489.

59 S. Banerjee and S. S. Wong, *Nano Lett.*, 2002, **2**, 195.

60 V. Georgakilas, D. Gournis, V. Tzitzios, L. Pasquato, D. M. Guldi and M. Prato, *J. Mater. Chem.*, 2007, **17**, 2679.

61 D. M. Guldi, G. Rahman, F. Zerbetto and M. Prato, *Acc. Chem. Res.*, 2005, **38**, 871.

62 E. Kymakis and G. Amaratunga, *Appl. Phys. Lett.*, 2002, **80**, 112.

63 E. Kymakis and G. Amaratunga, *Sol. Energy Mater. Sol. Cells*, 2003, **80**, 465.

64 S. Ravindran, S. Chaudhary, B. Colburn, M. Ozkan and C. S. Ozkan, *Nano Lett.*, 2003, **3**, 447.

65 V. Sgobba and D. M. Guldi, *Chem. Soc. Rev.*, 2009, **38**, 165.

66 C. Yang, M. Wohlgenannt, Z. Vardeny, W. Blau, A. Dalton, R. Baughman and A. Zakhidov, *Phys. B*, 2003, **338**, 366.

67 B. Zebli, H. Vieyra, I. Carmeli, A. Hartschuh, J. P. Kotthaus and A. W. Holleitner, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **79**, 205402.

68 Z. Lu, C. X. Guo, H. B. Yang, Y. Qiao, J. Guo and C. M. Li, *J. Colloid Interface Sci.*, 2011, **353**, 588.

69 N. A. Anderson and T. Lian, *Annu. Rev. Phys. Chem.*, 2005, **56**, 491.

70 J. B. Asbury, E. Hao, Y. Wang, H. N. Ghosh and T. Lian, *J. Phys. Chem. B*, 2001, **105**, 4545.

71 H. N. Ghosh, J. B. Asbury and T. Lian, *J. Phys. Chem. B*, 1998, **102**, 6482.

72 T. Hannappel, B. Burfeindt, W. Storck and F. Willig, *J. Phys. Chem. B*, 1997, **101**, 6799.

73 G. Ramakrishna, H. N. Ghosh, A. K. Singh, D. K. Palit and J. P. Mittal, *J. Phys. Chem. B*, 2001, **105**, 12786.

74 J. B. Asbury, R. J. Ellingson, H. N. Ghosh, S. Ferrere, A. J. Nozik and T. Lian, *J. Phys. Chem. B*, 1999, **103**, 3110.

75   P. Persson, M. J. Lundqvist, R. Ernstorfer, W. Goddard and F. Willig, *J. Chem. Theory Comput.*, 2006, **2**, 441.

76   S. Ramakrishna, F. Willig, V. May and A. Knorr, *J. Phys. Chem. B*, 2003, **107**, 607.

77   R. Huber, S. Spörlein, J. E. Moser, M. Grätzel and J. Wachtveitl, *J. Phys. Chem. B*, 2000, **104**, 8995.

78   R. Long and O. V. Prezhdo, *J. Am. Chem. Soc.*, 2011, **133**, 19240.

79   S. M. Pratik and A. Datta, *Phys. Chem. Chem. Phys.*, 2013, **15**, 18471.

80   Y. Y. Tan, W. H. Tu and S. Manzhos, *Chem. Phys. Lett.*, 2014, **593**, 14.

81   N. N. Ghosh, A. Chakraborty, S. Pal, A. Pramanik and P. Sarkar, *Phys. Chem. Chem. Phys.*, 2014, **16**, 25280.

82   K. J. Thomas, J. T. Lin, Y.-C. Hsu and K.-C. Ho, *Chem. Commun.*, 2005, 4098.

83   A. Pramanik, S. Sarkar, S. Pal and P. Sarkar, *Phys. Lett. A*, 2015, **379**, 1036.

84   W. R. Duncan and O. V. Prezhdo, *Annu. Rev. Phys. Chem.*, 2007, **58**, 143.

85   A. Nitzan, *Annu. Rev. Phys. Chem.*, 2000, **52**, 681.

86   Y. Tachibana, S. A. Haque, I. P. Mercer, J. E. Moser, D. R. Klug and J. R. Durrant, *J. Phys. Chem. B*, 2001, **105**, 7424.

87   Y. Tachibana, J. E. Moser, M. Grätzel, D. R. Klug and J. R. Durrant, *J. Phys. Chem.*, 1996, **100**, 20056.

88   R. Huber, J.-E. Moser, M. Grätzel and J. Wachtveitl, *J. Phys. Chem. B*, 2002, **106**, 6494.

89   J. Schnadt, P. A. Brühwiler, L. Patthey, J. N. O'Shea, S. Södergren, M. Odelius, R. Ahuja, O. Karis, M. Bässler and P. Persson, *Nature*, 2002, **418**, 620.

90   S. Kundu, S. Sadhu, R. Bera, B. Paramanik and A. Patra, *J. Phys. Chem. C*, 2013, **117**, 23987.

91   Y. Yin and A. P. Alivisatos, *Nature*, 2005, **437**, 664.

92   E. Martínez-Ferrero, J. Albero and E. Palomares, *J. Phys. Chem. Lett.*, 2010, **1**, 3039.

93   M. J. Greaney and R. L. Brutchey, *Mater. Today*, 2015, **18**, 31.

94   K. E. Knowles, D. B. Tice, E. A. McArthur, G. C. Solomon and E. A. Weiss, *J. Am. Chem. Soc.*, 2009, **132**, 1041.

95   A. M. Munro, I. Jen-La Plante, M. S. Ng and D. S. Ginger, *J. Phys. Chem. C*, 2007, **111**, 6220.

96   S. Günes, H. Neugebauer and N. S. Sariciftci, *Chem. Rev.*, 2007, **107**, 1324.

97   M. R. Wasielewski, *Chem. Rev.*, 1992, **92**, 435.

98   R. C. Haddon, *Acc. Chem. Res.*, 2002, **35**, 997.

99   H. Dai, *Top. Appl. Phys.*, 2001, **80**, 29.

100   A. Javey, J. Guo, Q. Wang, M. Lundstrom and H. Dai, *Nature*, 2003, **424**, 654.

101   A. Hirsch, *Angew. Chem., Int. Ed.*, 2002, **41**, 1853.

102   J. L. Bahr and J. M. Tour, *J. Mater. Chem.*, 2002, **12**, 1952.

103   R. J. Chen, Y. Zhang, D. Wang and H. Dai, *J. Am. Chem. Soc.*, 2001, **123**, 3838.

104   M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai and G. Seifert, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 7260.

105   T. Frauenheim, G. Seifert, M. Elstner, T. Niehaus, C. Köhler, M. Amkreutz, M. Sternberg, Z. Hajnal, A. Di Carlo and S. Suhai, *J. Phys.: Condens. Matter*, 2002, **14**, 3015.

106   R. Mitric, U. Werner, M. Wohlgemuth, G. Seifert and V. Bonačić-Koutecký, *J. Phys. Chem. A*, 2009, **113**, 12700.

107   T. A. Niehaus, S. Suhai, F. Della Sala, P. Lugli, M. Elstner, G. Seifert and T. Frauenheim, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2001, **63**, 085108.

108   D. Porezag, T. Frauenheim, T. Köhler, G. Seifert and R. Kaschner, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1995, **51**, 12947.

109   G. Seifert, *J. Phys. Chem. A*, 2007, **111**, 5609.

110   T. A. Niehaus, *J. Mol. Struct.: THEOCHEM*, 2009, **914**, 38.

111   S. Saha, S. Pal, P. Sarkar, A. Rosa and T. Frauenheim, *J. Comput. Chem.*, 2012, **33**, 1165.

112   S. Sarkar, S. Pal, P. Sarkar, A. Rosa and T. Frauenheim, *J. Chem. Theory Comput.*, 2011, **7**, 2262.

113   S. Sarkar, S. Saha, S. Pal and P. Sarkar, *J. Phys. Chem. C*, 2012, **116**, 21601.

114   Y. Q. Gao, Y. Georgievskii and R. Marcus, *J. Chem. Phys.*, 2000, **112**, 3358.

115   Y. Li, R. Mastria, A. Fiore, C. Nobile, L. Yin, M. Biasiucci, G. Cheng, A. M. Cucolo, R. Cingolani, L. Manna and G. Gigli, *Adv. Mater.*, 2009, **21**, 4461.

116   S. Sarkar, B. Rajbanshi and P. Sarkar, *J. Appl. Phys.*, 2014, **116**, 114303.

117   N. Song, H. Zhu, S. Jin, W. Zhan and T. Lian, *ACS Nano*, 2010, **5**, 613.

118   K. Novoselov, *Nat. Mater.*, 2007, **6**, 720.

119   K. S. Novoselov, A. K. Geim, S. Morozov, D. Jiang, Y. Zhang, S. A. Dubonos, I. Grigorieva and A. Firsov, *Science*, 2004, **306**, 666.

120   Z. Chen, S. Berciaud, C. Nuckolls, T. F. Heinz and L. E. Brus, *ACS Nano*, 2010, **4**, 2964.

121   G. Katsukis, J. Malig, C. Schulz-Drost, S. Leubner, N. Jux and D. M. Guldi, *ACS Nano*, 2012, **6**, 1915.

122   B. Rajbanshi, S. Sarkar and P. Sarkar, J. *Mater. Chem. C*, 2014, **2**, 8967.

123   L.-L. Li and E. W.-G. Diau, *Chem. Soc. Rev.*, 2013, **42**, 291.

124   S. Mathew, A. Yella, P. Gao, R. Humphry-Baker, B. F. Curchod, N. Ashari-Astani, I. Tavernelli, U. Rothlisberger, M. K. Nazeeruddin and M. Grätzel, *Nat. Chem.*, 2014, **6**, 242.

125   X. Yan, B. Li, X. Cui, Q. Wei, K. Tajima and L.-S. Li, *J. Phys. Chem. Lett.*, 2011, **2**, 1119.

126   X. Yan, X. Cui and L.-S. Li, *J. Am. Chem. Soc.*, 2010, **132**, 5944.

127   X. Yan, X. Cui, B. Li and L.-S. Li, *Nano Lett.*, 2010, **10**, 1869.

128   X. Yan and L.-S. Li, *J. Mater. Chem.*, 2011, **21**, 3295.

129   L.-S. Li and X. Yan, *J. Phys. Chem. Lett.*, 2010, **1**, 2572.

130   M. Alam Sk, A. Ananthanarayanan, L. Huang, K. H. Lim and P. Chen, *J. Mater. Chem. C*, 2014, **2**, 6954.

131   O. Voznyy, A. D. Güçlü, P. Potasz and P. Hawrylak, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **83**, 1654171.

132   S. Schumacher, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **83**, 081417.

133   J. Malig, C. Romero-Nieto, N. Jux and D. Guldi, *Adv. Mater.*, 2012, **24**, 800.

134   Y. Xu, L. Zhao, H. Bai, W. Hong, C. Li and G. Shi, *J. Am. Chem. Soc.*, 2009, **131**, 13490.

135   H. Zhang, Y. Han, Y. Guo and C. Dong, *J. Mater. Chem.*, 2012, **22**, 23900.

136   B. Mandal, S. Sarkar and P. Sarkar, *J. Phys. Chem. C*, 2015, **119**, 3400.

137   Y. Cui, Q. Wei, H. Park and C. M. Lieber, *Science*, 2001, **293**, 1289.

138   Z. W. Pan, Z. R. Dai and Z. L. Wang, *Science*, 2001, **291**, 1947.

139   K. Zhu, N. R. Neale, A. Miedaner and A. J. Frank, *Nano Lett.*, 2007, **7**, 69.

140   S. P. Albu, A. Ghicov, J. M. Macak, R. Hahn and P. Schmuki, *Nano Lett.*, 2007, **7**, 1286.

141   S. Sarkar, S. Saha, S. Pal and P. Sarkar, *RSC Adv.*, 2014, **4**, 14673.

142   M. Law, L. E. Greene, J. C. Johnson, R. Saykally and P. Yang, *Nat. Mater.*, 2005, **4**, 455.

143   V. Thavasi, V. Renugopalakrishnan, R. Jose and S. Ramakrishna, *Mater. Sci. Eng., R*, 2009, **63**, 81.

144 L. Liu, Y. Chen, T. Guo, Y. Zhu, Y. Su, C. Jia, M. Wei and Y. Cheng, *ACS Appl. Mater. Interfaces*, 2011, **4**, 17.

145 J. A. Goebl, R. W. Black, J. Puthussery, J. Giblin, T. H. Kosel and M. Kuno, *J. Am. Chem. Soc.*, 2008, **130**, 14822.

146 A. Listorti, B. O'Regan and J. R. Durrant, *Chem. Mater.*, 2011, **23**, 3381.

147 X. Meng, H. Peng, Y. Gai and J. Li, *J. Phys. Chem. C*, 2009, **114**, 1467.

148 S. Yang, D. Prendergast and J. B. Neaton, *Nano Lett.*, 2010, **10**, 3156.

149 E. M. Barea, M. Shalom, S. Giménez, I. Hod, I. Mora-Seró, A. Zaban and J. Bisquert, *J. Am. Chem. Soc.*, 2010, **132**, 6834.

150 J.-J. Wu, G.-R. Chen, H.-H. Yang, C.-H. Ku and J.-Y. Lai, *Appl. Phys. Lett.*, 2007, **90**, 213109.

151 J. Bisquert, F. Fabregat-Santiago, I. Mora-Sero, G. Garcia-Belmonte and S. Giménez, *J. Phys. Chem. C*, 2009, **113**, 17278.

152 C. F. Negre, V. C. Fuertes, M. B. N. Oviedo, F. Y. Oliva and C. N. G. Sánchez, *J. Phys. Chem. C*, 2012, **116**, 14748.

153 S. Saha, S. Sarkar, S. Pal and P. Sarkar, *J. Phys. Chem. C*, 2013, **117**, 15890.

154 S. S. Lo, T. Mirkovic, C. H. Chuang, C. Burda and G. D. Scholes, *Adv. Mater.*, 2011, **23**, 180.

155 A. V. Akimov, A. J. Neukirch and O. V. Prezhdo, *Chem. Rev.*, 2013, **113**, 4496.

156 L. Du and Z. Lan, *J. Chem. Theory Comput.*, 2015, **11**, 1360.

157 A. V. Akimov and O. V. Prezhdo, *J. Chem. Theory Comput.*, 2014, **10**, 789.

158 A. V. Akimov and O. V. Prezhdo, *J. Chem. Theory Comput.*, 2013, **9**, 4959.

159 M. Barbatti, M. Ruckenbauer, F. Plasser, J. Pittner, G. Granucci, M. Persico and H. Lischka, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 26.

160 K. Hyeon-Deuk and O. V. Prezhdo, *Nano Lett.*, 2011, **11**, 1845.

161 S. V. Kilina, D. S. Kilin, V. V. Prezhdo and O. V. Prezhdo, *J. Phys. Chem. C*, 2011, **115**, 21641.

162 S. Pal, D. J. Trivedi, A. V. Akimov, B. Aradi, T. Frauenheim and O. V. Prezhdo, *J. Chem. Theory Comput.*, 2016, **12**, 1436.

Chemical modelling of fluids near surfaces

J. K. Singh

1   Introduction

The behaviour of a fluid or fluid mixture near soft or hard interfaces is of immense importance in various industries.[1] This includes the self-assembly phenomena on surfaces commonly found in numerous processes ranging from life sustaining processes like folding of proteins to physical processes such as wetting–dewetting, lubrication, detergency *etc.* Thus, the study of self-assembly has direct relevance in the fields of biomaterials, bio-medical devices,[2] biosensors,[3] micro-nano-electronics[4,5] and many others.

The intricacy of the fluid behaviour on surfaces is enormous which grows with the complex nature of the fluid. For a simple fluid, the adsorption of fluid molecules, under bulk gas condition, on solid surfaces at temperatures below their bulk critical temperature can result in different phases such as percolation, layering, prewetting and wetting. In case of a liquid deposited on surfaces, we may either observe a wetted system or the film may break (dewetting behaviour) leading to the formations of holes or drops, or contract to form a single drop. The adsorbed phase depends on the nature of surfaces, and the competition of fluid–fluid and fluid–solid intermolecular forces leads to different kinds of phase transitions, which is seen in the presence of single sub-strate/surface as well as for confined systems.[6,7]

A typical nature of fluids on hard surfaces is quantified by the contact angle. However, in addition to the wetting transition, layering and prewetting transitions are also observed for some cases. Furthermore, a fluid may be unstable at the interface and may lead to dewetting. Dynamical nature of fluids on a substrate is also found to vary with the effective fluid–fluid and fluid–substrate interactions. In many scientific applications coating is necessary such as self-assembled monolayer which essentially imparts new functionality to the surface.[8] The behaviour of a self-assembled monolayer at substrate depends on the nature of the substrate, whether hydrophilic or hydrophobic, and its crystallographical plane.[9] The structure and dynamics of monolayers determine many properties like adhesion, molecular and biological recognition.[10–12] Furthermore, surface diffusion of the molecules in the monolayer plays an important role in many kinetic processes such as chemical reactions, catalysis, crystal growth, and is responsible for wetting/dewetting of a surface.

In recent years, superhydrophic surfaces have been extensively investigated. The motivation for the interest in such materials originated

Department of Chemical Engineering, Indian Institute of Technology Kanpur,
Kanpur 208016, India. E-mail: jayantks@iitk.ac.in

from the 'lotus' leaf characteristic, where a hierarchical nature of roughness is responsible for the self-cleaning nature of the lotus leaf. Numerous works have been done to mimic the nature to create super-ydrophobic surfaces.[13] While the development of superhydrophobic or superhydrophillic is relevant for many applications, controlling the wettability depending on the environment condition propel workers to develop responsive materials, which change their behaviour due to pH, temperature or presence of electrolytes. In such cases, the fluid undergoes wetting/dewetting behaviour different from what typically is seen for non-responsive surfaces.[14]

The success of superhydrophobic design has been extended to ice-repellant surfaces. Thus, there have recently been efforts to design such ice-repellant surfaces, and many of them have been inspired by the successes in the design of superhydrophobic surfaces.[15–26] The effort in recent years in this area is mainly due to the fact that formation of ice adversely affects industrial processes and applications in daily life such as in power lines, aircraft, ships and buildings.[27,28] However, there are conflicting outcomes based on recent experiments. While some experiments indicate that superhydrophobic surfaces can minimize or eliminate ice formation under some conditions,[29] other experiments, for example of Jung *et al.*,[30] conclude that anti-icing design needs optimization of the competing influence of both freezing delay and liquid-shedding ability *i.e.*, low adhesion. A recent work of Nosonovsky and Hejazi,[31] based on the theoretical analysis of mechanical forces acting on a liquid droplet and ice, suggests that it depends on the size of the roughness whether or not a superhydrophobic surface is at the same time ice-phobic. This is in line with the prediction[32] of a theoretical work that there is an optimal pit size on the surface for unrestrained growth of crystals. Furthermore, the lattice commensurability and incommensurability can affect crystal nucleation dramatically.[33] Moreover, the claim of Chen and co-workers that superhydrophobic surfaces cannot reduce ice adhesion[34] further corroborates the findings of Jung *et al.*[30] Hence, the apparent contradictions of experiments confuse the picture of the role of roughness for the design of anti-icing surfaces. Recently, a molecular simulation work from our group clearly suggests that under certain conditions, a surface phase transition on a rough surface can cause more adhesion of ice on the surface even though wetting states may have similar freezing temperature.[35] Thus, superhydrophic surfaces can have variable ice-repellent properties depending on the wetting states. In other words, anti-icing properties of a superhydrophobic surface are not entirely controlled by the surface roughness. Consequently, coating is a best way to go forward for a suitable ice-repellant surface.

In this chapter, we describe molecular modelling methods to investigate various aspects of wetting behaviour of fluids on substrates. First, we present the classification of fluid behaviour near surfaces followed by detailed molecular simulation methodologies relevant for wetting studies. This is followed by examples of different cases, which is primarily motivated by our own studies.

2 Classification of fluid behaviour near surfaces

A pure fluid may undergo different phase transitions in the presence of surfaces as illustrated in Fig. 1. At the saturation condition, a fluid may wet the surface above a particular temperature called the wetting temperature, T_w. The wetting temperature for a system is defined as a temperature where a thin film discontinuously transforms to a thick film. In other words, below the wetting temperature the thickness of the film adsorbed on a surface remains finite at all sub saturation pressures. At a subsaturation condition, a first order transition may be observed between two phases, where the films deposited on the surface differ by their thickness. This transition is known as the prewetting transition, which stems from the saturation line at the wetting temperature and goes away from the saturation line with a decrease in the density difference of thin and thick films, and becomes indistinguishable at the prewetting critical temperature, T_{pwc}, as shown in Fig. 1. To observe such phase transitions, precise control over pressure or density of the system is necessary, as it is very close to the bulk saturation line. Perhaps due to this fact, though it was theoretical prediction in 1977,[36,37] the experimental evidence came much later in 1992. Such phase transitions are observed for inert gases on alkali metal surfaces,[38–40] and acetone and water on graphite surfaces.[41,42]

Figure 2a shows the adsorption isotherm for cases when the temperature is below the wetting temperature. In such case finite adsorption is seen with increase in the pressure. At an intermediate temperature such that the temperature is between T_w and T_{pwc}, we may observe a prewetting transition that manifests itself through the first order transition (sudden jump in the density) at $P < P_{coex}$, as shown in Fig. 2b. The prewetting line on the phase diagram can be located by recording
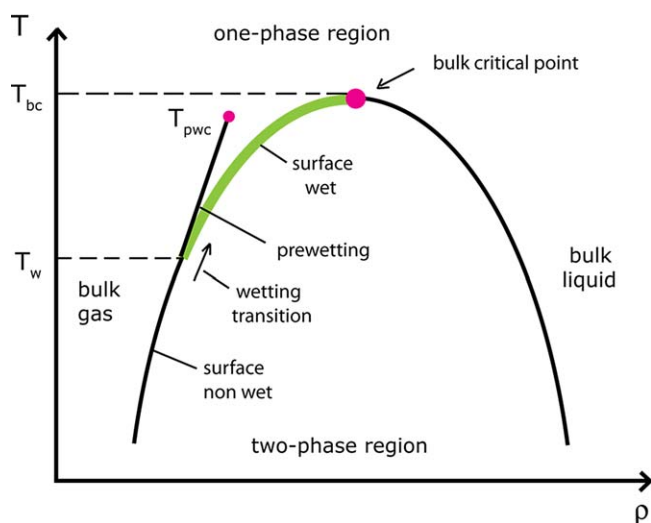

Fig. 1 Typical phase diagram of a fluid near a surface.
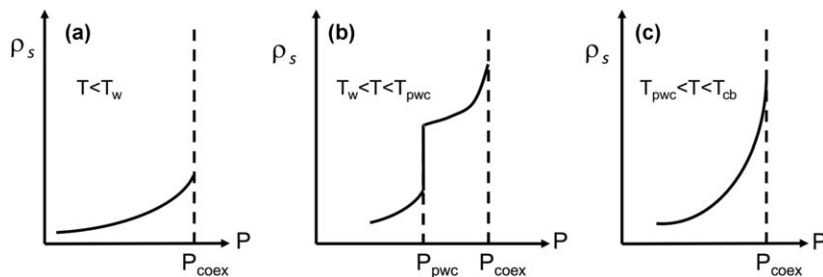
Fig. 2 The adsorption isotherm at different temperatures: (a) $T < T_w$, finite adsorption isotherm is observed, (b) first order transition representing the prewetting transition, (c), continuous adsorption leading to divergence of the film thickness.

prewetting transitions at different temperatures ranging from T_w to T_{pwc}. At T_{pwc} the thin and thick films are indistinguishable *i.e.*, the boundary tension between the two phases vanishes. With decreasing temperature, the prewetting pressure or chemical potential shifts towards the bulk saturation pressure or chemical potential. Moreover, at T_w no first order transition is seen as the prewetting chemical potential is akin to that of the bulk saturation values. Therefore, the difference between the bulk saturation chemical potential and the prewetting chemical potential, $\Delta\mu$, can be used to evaluate the wetting temperature, T_w. It has been found that locating the locus of the prewetting line close to the bulk saturation line is challenging from usual methods as seen from earlier studies.[37,43–48] This is evident from the wide range of the wetting temperature estimates for the argon-solid carbon dioxide model. In a recent work, Errington[48] examined the same system using a more advanced Monte Carlo method, which will be discussed later. The author could accurately capture the coexistence of thin and thick films close to the bulk saturation line and found the difference in the prewetting chemical potential with respect to the bulk saturation chemical potential, $\Delta\mu$, to follow the power law $(T - T_w)^{3/2}$ which is also predicted by a theory.[49–52] In this chapter we will provide some examples of the prewetting transition predicted using molecular simulations, and in general the modelling techniques to examine such phase transitions.

While gas phase deposition on a substrate leading to phase transitions is an intriguing subject, a more commonly studied system is the wettability of a substrate by a liquid. This is primarily motivated by various industrial applications where hydrophobicity/hydrophillicity and permeability of water in the material is of direct relevance. Wettability of any substrate is generally characterized by the contact angle of a liquid drop on the substrate (as shown in Fig. 3a). The liquid drop on a substrate under mechanical equilibrium yields a contact angle due to the interfacial tensions experienced by the drop, at the three-phase contact line. The relation between contact angle and these interfacial tensions at the three-phase contact line, is given by the Young's equation:

$$\gamma_{sv} = \gamma_{sl} + \gamma_{lv} \cos\theta \tag{1}$$
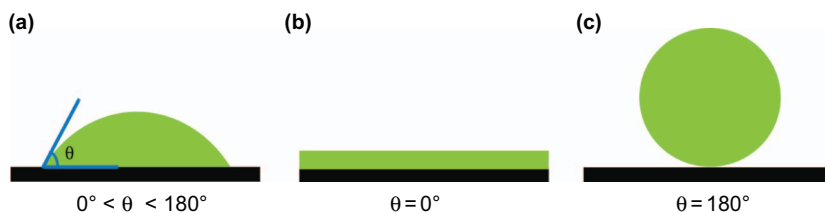
Fig. 3 A liquid drop on a flat substrate: (a) partial wetting state; (b) complete wetting state; (c) complete drying state.

where, γ is the interfacial tension between two phases as indicated by the subscripts. The subscript s, l, v refer to solid, liquid and vapour, respectively, and θ is the contact angle of the droplet on the surface. The contact angle of a droplet can be found in between $0°$ and $180°$ depending on the spreading of the droplet on the surface. The spreading of the droplet or the length of the three-phase contact line is controlled by the interfacial tensions. For example, the length of the three-phase contact line becomes infinite ($\theta = 0°$) when the sum of the solid–liquid and liquid–vapour interfacial tensions is equal to the solid–vapour interfacial tension. As a result, the surface is completely covered by a liquid film as illustrated in Fig. 3b. On the other hand, the three-phase contact line reduces to a point when the sum of solid–vapour and the liquid–vapour interfacial tensions is equal to the solid–liquid interfacial tension ($\theta = 180°$). This situation is known as complete drying state and is shown in Fig. 3c. An intermediate state of the water droplet occurs when there is a finite contact line among the three phases. This state is known as partial wetting, and is recognized by the contact angle being between $0°$ and $180°$ (see Fig. 3a). A surface is called hydrophillic when the contact angle is less than $90°$. In case of hydrophobic surface, the contact angle is greater than $90°$.

The application of Young's equation is not directly suitable for nano size liquid droplets, as typically studied in molecular simulations, due to the absence of three-phase interaction across the three-phase contact line. Hence, to account for the three-phase contact line, the line tension, τ, which is defined as the interfacial energy per unit length across the three-phases contact line, is introduced in the Young's equation:

$$\gamma_{sv} = \gamma_{sl} + \gamma_{lv} \cos \theta + \left(\frac{\tau}{r_B} \right) \qquad (2)$$

where r_B and θ are the base radius of the droplet and the microscopic contact angle of the droplet respectively. Hence, this equation is useful for the analysis of micron (or lesser) size droplet.[53] Macroscopic contact angles can be evaluated from the eqn (2) by extrapolating $1/r_B \rightarrow 0$, which yield the Young's equation:

$$\cos \theta_\infty = \left(\frac{\gamma_{sv} - \gamma_{sl}}{\gamma_{lv}} \right) \qquad (3)$$

where, θ_∞ is the contact angle of an infinitely large droplet on a surface. Therefore, microscopic and macroscopic contact angle on a smooth surface can be correlated through eqn (2) and (3) as given below:

$$\cos\theta = \cos\theta_\infty - \left(\frac{\tau}{\gamma_{lv}}\right)\frac{1}{r_B}. \tag{4}$$

3 Molecular simulation methodologies

3.1 Phase coexistence techniques

Molecular modelling methods to evaluate phase coexistence of pure fluids and mixtures are now well established. Typical slab based techniques where the initial density of the system is chosen such that phase separation occurs, require the stability of the interface. Though the method is straight forward, it is not the most efficient. This technique is employed by few groups[54] to study the surface phase transition such as thin–thick film coexistence on substrates. While the explicit slab based technique is used often, there exist better techniques for evaluating phase transitions in the presence of substrates. One such technique is the grand canonical transition matrix Monte Carlo simulation (GC-TMMC) and which is popular for the study of wetting and prewetting transitions.

GC-TMMC is basically a standard GCMC simulation carried out at a fixed chemical potential μ, volume V, and temperature T in which a transition matrix is used to record the probability of the transition from one microstate to another regardless of whether the moves are accepted or not. The probability of a microstate s, in a GCMC ensemble is given by:

$$\pi(s) = \frac{1}{\Xi}\frac{V^{N(s)}}{\Lambda^{3N(s)}N(s)!}\exp[-\beta E(s)]\exp[\beta\mu N(s)] \tag{5}$$

where β denotes the inverse temperature ($1/k_B T$, k_B is the Boltzmann factor), Λ is the de Broglie wavelength, and Ξ is the grand partition function, which is the summation of probabilities of all microstates. Hence, $\pi(s)$ is the normalized probability. $N(s)$, $E(s)$ are the number of particles and energy of the microstate s respectively. However, in this method, the number of particles and the energy of the system fluctuate according to the thermodynamic state of the system. The macrostate probability, $\Pi(N)$, is calculated by summing probabilities of all microstates at a constant number of particle N, as shown below:

$$\Pi(N) = \sum_{N_s = N}\pi(s). \tag{6}$$

A transition matrix can be used to evaluate the probability of each macrostate. The algorithm[55] was first proposed by Fitzgerald *et al.*, and is described in detail below.

In GCMC simulations, all the possible microstates are generated through different moves such as add, remove, displace, rotation *etc.*

For example, for a move from a microstate s_1 with N_1 particles to a microstate s_2 with N_2 particles, the acceptance probability is defined as

$$a(s_1 \rightarrow s_2) = \min\left[1, \frac{\pi(s_2)}{\pi(s_1)}\right]. \tag{7}$$

The acceptance probability, a, is then used to update the transition matrix, C, as follows:

$$C(N_1 \rightarrow N_2) = C(N_1 \rightarrow N_2) + a(s_1 \rightarrow s_2), \tag{8}$$

and

$$C(N_1 \rightarrow N_1) = C(N_1 \rightarrow N_1) + 1 - a(s_1 \rightarrow s_2). \tag{9}$$

After sufficient sampling over the phase space corresponding to the thermodynamic states of the system, the populated C matrix can be used to evaluate the macrostate transition probability:

$$P(N_1 \rightarrow N_2) = \frac{C(N_1 \rightarrow N_2)}{\sum\limits_{N_k} (N_1 \rightarrow N_k)}. \tag{10}$$

Once the transition probability is obtained, we can use Monte Carlo detailed balance to determine the probability of a macrostate from the macrostate transition probability:

$$\Pi(N_1)P(N_1 \rightarrow N_2) = \Pi(N_2)P(N_2 \rightarrow N_1). \tag{11}$$

In common grand-canonical Monte Carlo simulations, usually addition, deletion, and displacement moves of one molecule at a time are considered. Thus, the following state changes are possible: $N \rightarrow N$, $N \rightarrow N-1$ and $N \rightarrow N+1$. Under such conditions a sequential approach would be an efficient and simple way to obtain the macrostate probabilities:

$$\ln \Pi(N+1) = \ln \Pi(N) - \ln\left[\frac{P(N+1 \rightarrow N)}{P(N \rightarrow N+1)}\right]. \tag{12}$$

The above expression is used for determining macrostate or particle number probability distributions over a specified range of number densities, which is bimodal in nature if two different phases coexist for the given thermodynamic state of the system. However, if the interfacial energy between two phases is sufficiently large then adequate sampling over two phases is very difficult to get from a conventional grand canonical simulation. To overcome this problem, the multi-canonical sampling method[56] is often used to sample all the microstate with equal probability. In this method, the acceptance probability is multiplied with some weighting parameter, $\eta(N)$. An appropriate weighting parameter is the inverse of macrostate probability distribution, $1/\Pi(N)$, obtained from

the transition matrix, which is evaluated at regular intervals. Thus, the modified acceptance probability is given as:

$$a(s_1 \rightarrow s_2) = \min\left[1, \frac{\eta(N_2)\pi(s_2)}{\eta(N_1)\pi(s_1)}\right] \tag{13}$$

However, the transition matrix is updated with the probability without the weighting parameter. Over time all densities of interest are sampled adequately. The final result is an efficient self-adaptive method for determining the probability distribution over a specified range of densities. Once a macrostate probability distribution has been collected at a given value of chemical potential, histogram reweighting[57] is used to shift the probability distribution to other values of the chemical potential using the following equation:

$$\ln \Pi(N; \mu) = \ln \Pi(N; \mu_0) + \beta(\mu - \mu_0)N. \tag{14}$$

The coexistence chemical potential is calculated by recursively applying the histogram reweighting[57] technique until we obtain a probability distribution, Π_C, such that areas under the two phase regions in the probability distribution plot are equal.

Figure 4 presents a typical particle number (or macrostate) probability distribution for the bulk phase at sub-saturation conditions as obtained from a GC-TMMC simulation. The first peak represents the particle number corresponding to coexistence vapour density and the second peak represents the particle number corresponding to coexistence liquid density, and F_L is the interfacial free energy of the coexistence phases. The average properties for vapour and liquid phases can be estimated
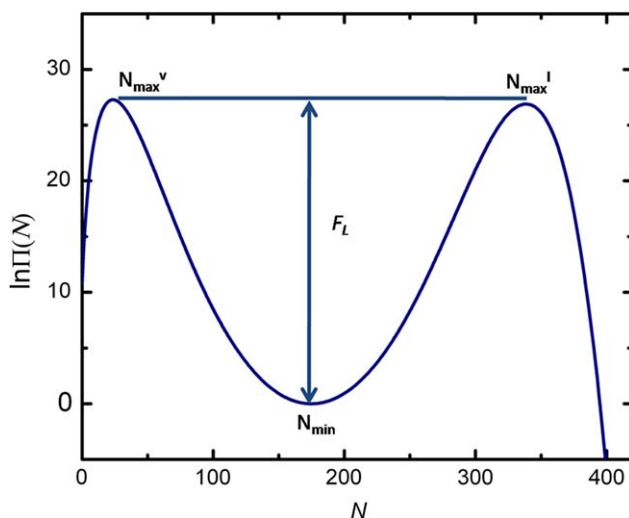

Fig. 4 Particle number probability distribution at a vapour–liquid coexistence condition.

considering the particle range for vapour and liquid phases as shown below:

$$\langle A \rangle = \frac{\sum_N \Pi(N; \mu) A(N)}{\sum_N \Pi(N; \mu)} \tag{15}$$

Saturated densities are related to the first moment of the vapour and liquid peaks of the coexistence probability distribution, Π_C. To calculate the saturation pressure we use the coexistence probability distribution in the following expression:

$$\beta PV = \ln\left(\sum \Pi_C(N)/\Pi_C(0)\right) - \ln 2. \tag{16}$$

In addition to the calculation of saturation pressures, densities, and energies, grand-canonical Monte Carlo simulations can also be combined with finite-size scaling methods to evaluate the surface tension. This method does not require establishing and maintaining an interface, which is necessary in slab-based methods.[57]

The interfacial free energy for a finite–size system with a cell length of L, F_L, is determined from the maximum likelihood in the liquid Π_{max}^l and vapour regions Π_{max}^v, and minimum likelihood in the interface region Π_{min},

$$\beta F_L = \frac{1}{2}\left(\ln \Pi_{max}^l + \ln \Pi_{max}^v\right) - \ln \Pi_{min}. \tag{17}$$

From the formalism of Binder,[58] the interfacial free energy of a two-dimensional surface varies with the system size according to the relation

$$\beta \gamma_L = \frac{\beta F_L}{2A} = C_1 \frac{1}{S} + C_2 \frac{\ln L}{S} + \beta \gamma_\infty \tag{18}$$

where, γ_L is the apparent system size dependent surface tension and γ_∞ is the infinite-system size $(L \to \infty)$ surface tension, A is the interfacial area, and C_1 and C_2 are constants.

The attractive feature of the GC-TMMC method is that it allows discretizing the whole particle range into several sub-domains that can be distributed over different processors separately. The simulations over all processors determine the probability distribution for the particle range independently. However, the final probability distribution can be determined by combining the probability distribution over all processors. The particle range for each processor is typically divided in such a way that each macrostate (*i.e.*, with a number of particle, N) falls within the domain of two processors. At regular intervals, the transition matrix from each processor is interchanged to the transition matrix from an adjacent processor, which helps to sample the entire phase space efficiently.

The GC-TMMC method can be used to obtain phase transitions in bulk and in the presence of a surface. Moreover, the method can be used to determine the vapour–liquid phase transition of pure fluids and mixtures

efficiently. The use of GC-TMMC was extended by Errington to study prewetting transitions.[48] The method has been further extended by Errington and co-workers for obtaining the contact angle on a surface, which is described in the next section.

3.2 Contact angle calculations

Wetting of fluids on surfaces has been in limelight for several decades due to their large number of applications in various fields such as oil recovery, lubrication, coating, painting, printing, spray technology, agriculture *etc.* Wetting/dewetting behaviour is characterized by measuring the contact angle of the liquids. There are several methods to calculate contact angles in experiments, such as sessile bubble and drop, captive drop, Wilhelmy plate *etc.* In molecular simulations, the contact angle can be calculated using typically fitting or center of mass methods. In case of Monte Carlo simulation, in addition the GC-TMMC method can be used to evaluate the contact angle. In the subsequent subsections, we describe these methods briefly.

3.2.1 Fitting method. In molecular dynamics simulations, the *fitting* method, first adopted by Werder *et al.*,[58–60] is one of most popular methods to calculate the contact angle of a liquid droplet on surfaces. In our study, we have implemented this method to calculate the contact angle of a water droplet on a grafted polymeric brush layer. In order to reduce the effect of line tension on the contact angle, we have studied a cylindrical droplet. From the molecular dynamics trajectories, fluid isochore profiles are obtained by introducing a cylindrical binning normal to the surface keeping the center of mass of the droplet as reference axis and horizontal binning parallel to the surface. Averages are taken over all configurations. All bins are assumed to be of equal volume. Generally, bins are considered to be of same height in z direction (see Fig. 5) and
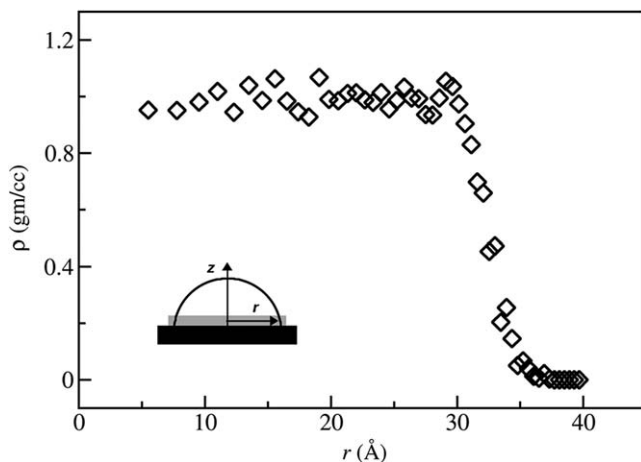
Fig. 5 A typical profile of a water across an axis from the centre of mass of the droplet. The gray shaded region of the cartoon (see the inset) represents the bin, which is used to obtain the profile. The data is being generated using a molecular dynamics simulation of a water model on a graphite surface.
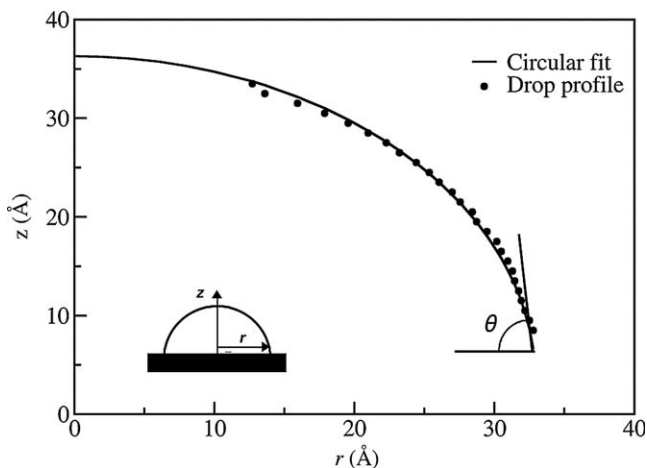
Fig. 6 Schematic for the computation of the contact angle for a water drop on a substrate. The contour of a water drop, represented by filled circles, is fitted to an equation of ellipse, which is shown by the solid curve.

the surface area of horizontal bins is varied to keep the volume fixed. Using the density profiles at different heights, the equimolar dividing surface is identified for each horizontal layer. The equimolar dividing plane can be obtained from the density profile of equilibrated liquid and vapour phases at a specific z using the equation given below:

$$\rho(z) = \frac{1}{2}(\rho_l + \rho_v) - \frac{1}{2}(\rho_l - \rho_v)\tanh\left[\frac{2(z - z_e)}{\delta}\right] \qquad (19)$$

where ρ_l, ρ_v, are the average density of the liquid and vapour phase respectively, z is the distance from the surface, z_e is the height of the equimolar dividing plane from the surface and δ is the thickness of the interface. The position of the interface, as shown in Fig. 6, is fitted with an equation of ellipse and extrapolated to the surface. The contact angle is obtained from the tangent.

3.2.2 Center of mass method. The center of mass method for the calculation of contact angle was developed by Hautman and Klein.[61,62] Hautman and Klein have shown that the contact angle of a water droplet on a substrate can be expressed in terms of the center of mass of the droplet, z_{cm}, as

$$\langle z_{cm} \rangle = 2^{-4/3} R_0 \left[\frac{1 - \cos\theta}{2 + \cos\theta}\right]^{1/3} \frac{3 + \cos\theta}{2 + \cos\theta} \qquad (20)$$

where R_0 is the radius of a free spherical droplet. According to the authors, the two most critical assumptions were considered – *first*, an idealized droplet *i.e.*, uniform bulk water density, which might not always be the case; *second*, the calculation of the center of mass was with respect to the planar surface of the top atomic layer. However, that would be difficult to identify if the layer is soft and responsive. Fan and Cağin[63]

have extended the method and evaluated the contact angle by calculating the volume of the droplet and the interfacial area between droplet and surface.

3.2.3 The GC-TMMC method. The GC-TMMC method can be used to calculate the contact angle on a substrate. However, two independent simulations need to be carried out. First, a separate simulation is conducted to obtain the bulk vapour–liquid surface tension. The other simulation is conducted at the coexistence bulk chemical potential but in the presence of the surface. A bulk simulation is employed to find out the liquid–vapour surface tension γ_{lv}, and Young's equation is used to evaluate the contact angle. This approach was developed mainly by Errington and co-workers[64] to calculate the contact angle using GC-TMMC simulations.

In case of the bulk phase, at a coexistence chemical potential, we observe a bimodal distribution of the probability distribution. In order to obtain the probability distribution in presence of a substrate of interest, a hard wall is included at the end of the simulation box (opposite to the side where the substrate is kept). This is included mainly to enhance the sampling of the configuration space in the simulation, and to reduce the length of the simulation. In such a system, the probability distribution at the bulk vapour–liquid saturation chemical potential μ^{sat} and a temperature of interest is no longer bimodal. Figure 7 shows a typical
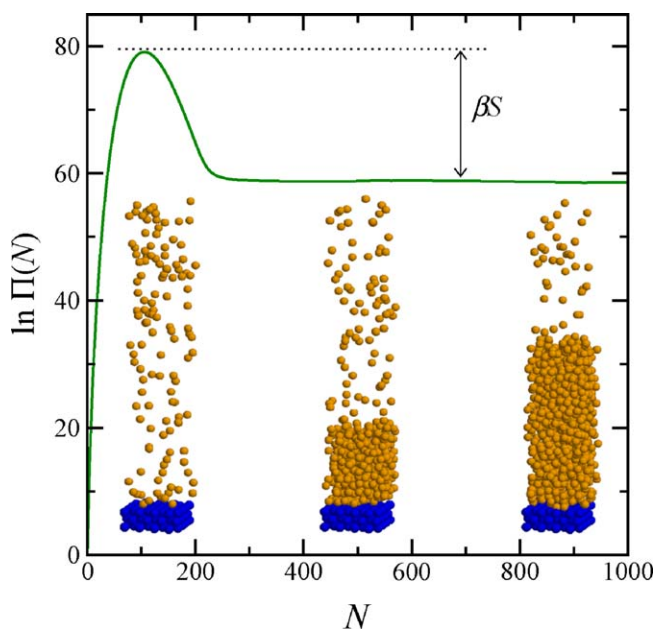
Fig. 7 A typical particle number probability distribution of a Lennard-Jones fluid in the presence of a 9-3 surface generated from a one-wall simulation conducted at the bulk liquid–vapour saturation chemical potential and at a temperature below a wetting point. The snapshots correspond to representative particle configurations for the particle number at which they are located. Reproduced from ref. 64, with the permission of AIP Publishing.

probability distribution for the partial wetting conditions (*i.e.*, for the temperature below the wetting temperature). The peak at the vapour-like number particles, N_v, represents a stable phase. This low-density peak corresponds to a saturated system with a solid–vapour interface. Conversely, at a large enough particle number, $N_{plateau}$, the influence of the substrate diminishes and the probability distribution reaches a limiting value. The system at large N corresponds to a saturated system with a liquid–vapour and solid–liquid interface. The spreading coefficient S can be related to the ratio of the aforementioned probabilities *i.e.*,

$$\beta S = \beta A [\gamma_{sv} - (\gamma_{sl} + \gamma_{lv})]$$
$$= -[\ln \Pi(N_v; \mu^{sat}) - \ln \Pi(N_{plateau}; \mu^{sat})] \tag{21}$$

where, γ_{sv}, γ_{sl} and γ_{lv} are the solid–vapour, solid–liquid and liquid–vapour interfacial tensions, respectively. A is the interfacial area. Combining this interfacial information with Young's equation, once can obtain the contact angle as follows –

$$\cos \theta = 1 + S/A\gamma_{lv} \tag{22}$$

This method, however, requires additional simulations in order to determine the liquid–vapour interfacial tension.

4 Prewetting transitions

The pioneering work of Cahn and Ebner and Saam on prewetting transitions initiated numerous theoretical and computation studies.[48,65–70] Using density functional theory, Ebner and Saam predicted the prewetting transition for Ar on solid carbon-dioxide, which is characterized by a transition between a thin and thick adsorbed film, at a temperature greater than the wetting temperature. Similar to bulk vapour–liquid saturation lines, prewetting saturation lines end at the prewetting critical temperature. The first molecular simulation for Ar–CO_2 was conducted by Finn and Monson[68,69] who observed a prewetting transition, inferred from the discontinuous jump in the excess adsorption isotherm in a Monte Carlo simulation using an isothermal–isobaric ensemble. Later, the authors calculated the particle number probability distribution using grand-canonical Monte Carlo simulations, and found a clear first order phase transition between a thin and a thick film of Ar on the CO_2 surface. Various authors later employed molecular simulations to understand the prewetting transition. However, it was Errington and co-workers who employed transition matrix Monte Carlo technique to precisely obtain the wetting temperature, which was not estimated earlier.

The GC-TMMC method is also an useful technique to study the surface phase transitions as shown by earlier studies.[48,71] The particle number distribution probability for a typical prewetting transition is shown in Fig. 8. The prewetting transition, a first order transition, is associated with the thin and thick films that coexist on the substrate. However, the particle number probability distribution for thin and thick films only
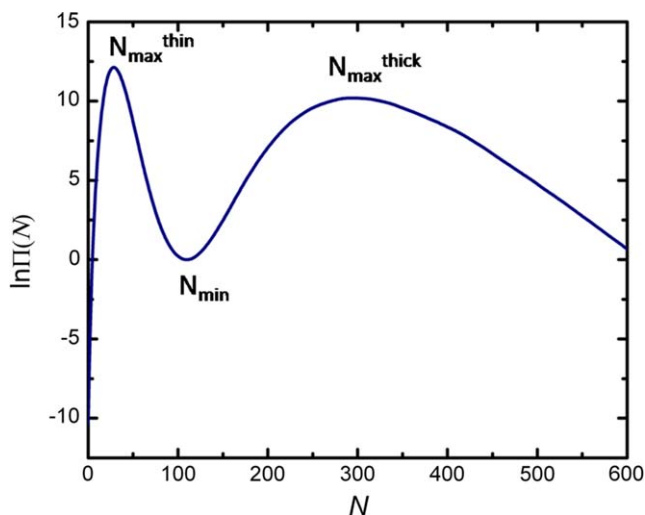
Fig. 8 Particle number probability distribution on a surface at a bulk vapour–liquid coexistence condition.

little different from that for vapour–liquid. The probability distribution for a thick film is significantly wider compared to a liquid phase, which suggests that the thick film density is very sensitive to the thermodynamic condition of the system. A typical snapshot of thin and thick films is given in Fig. 9 for the case of argon on solid CO_2. It has been found from earlier studies that locating the thin and thick film close to the wetting temperature is very difficult and hence prone to an erroneous prediction of the wetting temperature.[37,43–48] The GC-TMMC method, on the other hand, enables to carry out simulations close to the wetting temperature. However, the system size effect is prominent close to wetting temperature.[52]

An adsorption isotherm can be evaluated also from a single simulation with help of the histogram reweighting technique. The average density of the system is determined from the particle number probability distribution at a chemical potential as described by eqn (14). Histogram reweighting is used to determine the average density of the system at other chemical potentials. A typical adsorption as a function of chemical potential is shown in Fig. 10. This method is very efficient to differentiate different types of surface phase transitions such as layering transition and prewetting transition *etc.*

The prewetting transition is a first order transition, which occurs at a temperature in between the wetting temperature and the prewetting critical temperature of a system. Therefore, the length of the prewetting transition is generally measured by the difference between the wetting temperature and the prewetting critical temperature. However, it is very difficult to locate the accurate wetting temperature and prewetting temperature from adsorption isotherms. As the prewetting transition stems from the bulk saturation line at the wetting temperature, the bulk chemical potential and prewetting chemical potential are equal at the wetting temperature. Therefore, the difference between the bulk
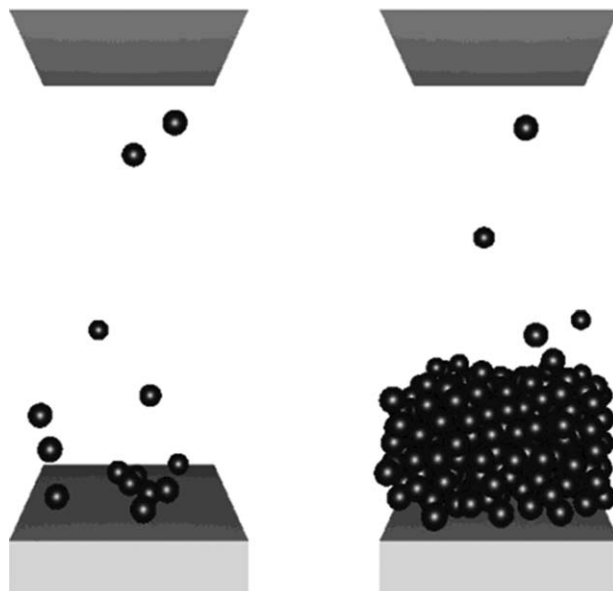
Fig. 9 Typical snapshots of configurations of saturated thin (left) and thick (right) films at a temperature less than the prewetting critical temperature of the Ar–CO$_2$ system. Reprinted (adapted) with permission from ref. 48. Copyright (2004) American Chemical Society.
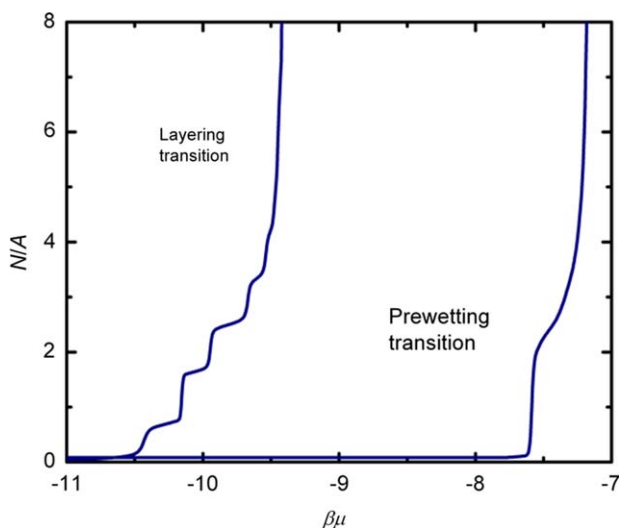
Fig. 10 Typical adsorption isotherms for surface phase transitions for layering and prewetting transitions. Here, N and A are the number of particles and surface area of the substrate, respectively.

saturation chemical potential and the prewetting chemical potential, $\Delta\mu$, can be used to evaluate the wetting temperature.

It has been found that locating the locus of the prewetting line close to the bulk saturation line is very difficult from usual methods.[37,43–45,47,48,72] For example, prediction of the wetting temperature (T_w), in the reduced

unit, for the argon gas on a solid carbon dioxide substrate is found to be quite diverse varying from 0.58 to 0.96, depending on the methods and the temperature range considered for evaluating $T_{\rm w}$.[48,68,69] However, the grand canonical transition matrix Monte Carlo method is quite effective to capture the coexistence of thin and thick films close to the bulk saturation line, and found $\Delta\mu$ to follow the power law $(T - T_{\rm w})^{3/2}$ as also predicted by theory.[73] On the other hand, the prewetting critical temperature can be obtained from extrapolating the series of true boundary tensions as a function of temperature to zero.

4.1 Prewetting transitions of molecular systems

Though the prewetting transition was studied theoretically for different systems with weak adsorption on substrates, the experimental results are scarce. Most of the experimental observations are for metal vapours on weak alkali metal substrates.[74–76] Kruchten and Knorr on the other hand observed a prewetting transition behaviour for acetone on a graphene system,[77] which provided the first experimental evidence of a molecular fluid undergoing the prewetting transition on a weak substrate. Subsequently, using molecular simulations Zhao *et al.* observed a prewetting transition for water on a graphite substrate. Thus, it could be argued that given a suitably weak substrate the associating fluids may also undergo prewetting transitions.

A key feature of associating fluids, such as hydrogen fluoride, acetic acid and water, is the presence of short-ranged, orientational dependent strong attractive forces between molecules. Such fluids can form clusters and ordered structures even at low density. These associating fluids depending on the directional dependent sites can be modelled as one-site, two-site or four-site models. These sites mimic the strong and short-range directional attraction (such as hydrogen bonds) of real associating fluids such as alcohol, hydrogen fluoride and water. On the other hand, isotropic van der Waal interaction can be modelled using Lennard-Jones or square-well models. Figure 11 illustrates a single site model for associating fluids. In our recent works,[78,79] we have investigated the prewetting transition of one-site, two-site and four-site associating fluids on smooth surfaces for various associating strengths. In case of two-site associating fluid, sites are located opposite to each other. On the other hand, sites are arranged in the tetrahedral structure for a four-site associating fluid, similar in spirit of the water models.
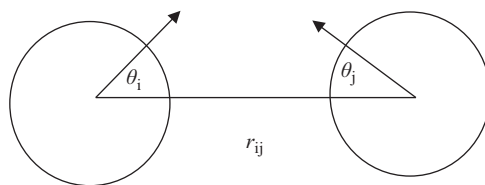
Fig. 11 An illustration of a single–site model of associating fluids. The distance between the molecules, i and j, is given by $r_{\rm ij}$. The arrows represent the vector joining the active sites on the surface with the center of the molecules.

The complete potential model employed for a one-site associating fluid is given below.

$$u_{ff} = u_{LJ\text{-}tr}(r_{ij}) + u_{af}(r_{ij}, \theta_i, \theta_j),$$

$$u_{LJ\text{-}tr}(r_{ij}) = \begin{cases} 4\varepsilon\left[\left(\dfrac{\sigma}{r_{ij}}\right)^{12} - \left(\dfrac{\sigma}{r_{ij}}\right)^6\right] & \text{if } r_{ij} < r_{cut}, \\ 0 & \text{otherwise} \end{cases} \qquad (23)$$

$$u_{af}(r_{ij}, \theta_i, \theta_j) = \begin{cases} -\varepsilon_{af} & \text{if } \sigma < r_{ij} < r_c, \ \theta_i < \theta_c \text{ and } \theta_j < \theta_c, \\ 0 & \text{otherwise} \end{cases}$$

where, θ_i and θ_j are angles between the center–center and the center-site vectors of molecules i and j, respectively. ε_{af} is the association energy and r_c is the range of site–site interaction. σ and ε are the model parameters for the Lennard-Jones (LJ) potential. r_{cut} is the cut-off diameter for the LJ potential. We adopt units such that σ and ε are unity. We kept θ_c, r_c and r_{cut} fixed at $27°$, $1.00 \, \sigma$ and $2.5 \, \sigma$, respectively. Based on the number of sites and the association strength, this model can be used for real associating fluids. This single-site associating model has been used for acetic acid with $\varepsilon_{af} = 20 \, \varepsilon$.[80]

A first order transition between thick and thin films for different temperature ranges depending on the number of sites, and site strength of associating fluids was observed by Khan et al.[67,79] Figure 12 presents the phase diagram of thin and thick films of different associating fluids. In general, the prewetting critical temperature increases with increasing associating strength as well as with number of sites. However, the pre-wetting span decreases with increasing associating strength as well as
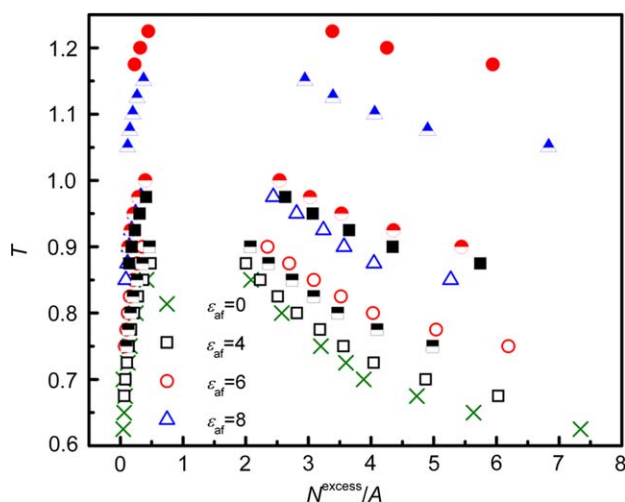
Fig. 12 Surface phase coexistence curves for multiple associating fluids. Crosses, open symbols, half-open symbols, and close symbols represent zero-site, one-site, two-site and four-site associating fluids, respectively. Reprinted (adapted) with permission from ref. 79. Copyright (2012) Taylor & Francis.
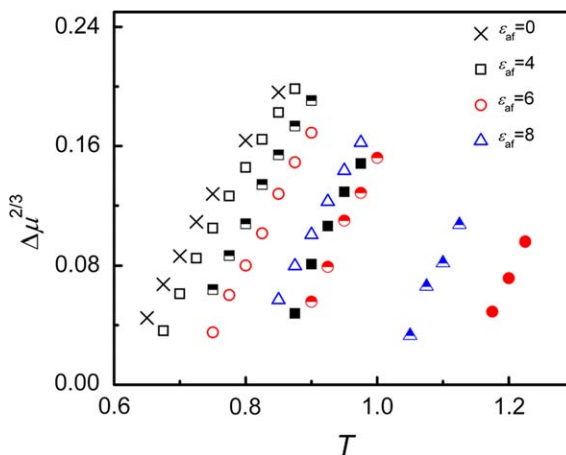
Fig. 13 Difference between the bulk saturation chemical potential and the chemical potential for the surface phase transition *vs.* temperature for multiple-site associating fluids. Reprinted (adapted) with permission from ref. 79. Copyright (2012) Taylor & Francis.

number of sites. Thick film density is also found to increase with increasing associating strength and number of sites. This is due to the dimer formation in case of the one-site associating fluid, chain formation in case of the two-site associating fluid, and three-dimensional complex structures in case of the four-site associating fluid.

Figure 13 shows the span of prewetting transitions for multiple-site associating fluids with various associating strengths. The prewetting regime for all associating fluids is found to fall within a triangular shaped region. The triangular region can provide a complete picture of the region of prewetting transition for all set of associating fluids with variable sites (0 to 4) and associating energies (0 to 8). For example, the prewetting transition for the one-site associating fluid can be observed for large values of associating strength; whereas for two-site and four-site associating fluids the prewetting transition is absent beyond associating strengths of 8 and 6, respectively. Considering the proximity of the prewetting transition near the bulk condition, the prewetting transition of multiple sites associating fluids with high associating strength such as hydrogen fluoride and water may be very difficult to observe, which is evident from the much smaller number of experimental observations for polar molecules.[41]

5 Wetting transition of water on solid surfaces

The behaviour of water on atomistically smooth surfaces has been studied by many. This includes water on mica and graphite surfaces.[81,82] Recently, there has been some debate on the wetting transparency of graphene for water. It was reported that graphene coating on certain materials such as gold, copper or silicon does not affect the intrinsic wetting behaviour of the substrate.[83] The authors used molecular dynamics to provide molecular insight to the behaviour. As mentioned earlier, the wetting behaviour is analysed through the contact angle.

However, in molecular modelling, the system size can have significant effects on the contact angle. In our recent work,[84] we have studied the wetting temperature and the line tension for a water–graphite system using molecular dynamics simulations, which will be described below in detail.

The water–water interaction is described by the extended single point charge (SPC/E)[85] model and the surface–water interaction is represented by the Lennard-Jones potential. We have performed a series of simulations using different system sizes *viz.*, 4000, 5000, 6000 and 7000 water molecules, at different temperatures. The microscopic contact angle for each system size is evaluated using the fitting method described earlier in this chapter.[84] These microscopic contact angles are then extrapolated to get the macroscopic contact angle as shown in Fig. 14.

As described earlier, the wetting temperature is the transition temperature from the partial wetting state to the complete wetting state. Thus the wetting temperature, T_w, is the temperature where the contact angle goes to zero. In order to evaluate T_w a series of contact angles as a function of temperature is used. The data is extrapolated linearly as a function of temperature to obtain the temperature at which contact angle becomes zero as illustrated in Fig. 15. Using the above approach, we have calculated T_w for the graphite–water (470 ± 5 K) and boron–nitride (438 ± 5 K) systems. The wetting temperature of water on the graphite surface is in agreement with that reported by Zhao[86] from a prewetting transition study using GCMC simulations. However, Zhao used the prewetting line of water on a graphite surface to estimate the wetting temperature. Thus, the agreement between the two works strongly supports the existence of a prewetting transition of water on graphite.

Molecular dynamics of drops of variable size on a substrate can be used to calculate the line tension. The dependence of the contact angle with the size of the droplet can be used to obtain the slope $(-\tau/\gamma_{lv})$ as per
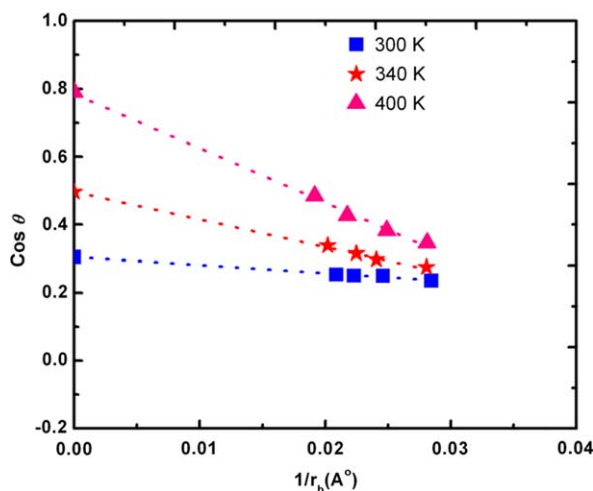
Fig. 14 Dependence of the size of a water droplet on its contact angle on a boron–nitride surface at different temperatures. Symbols, from right to left, represent droplets with 2000, 3000, 4000, 5000 water molecules, respectively. Reprinted (adapted) with permission from ref. 84. Copyright (2011) Elsevier.
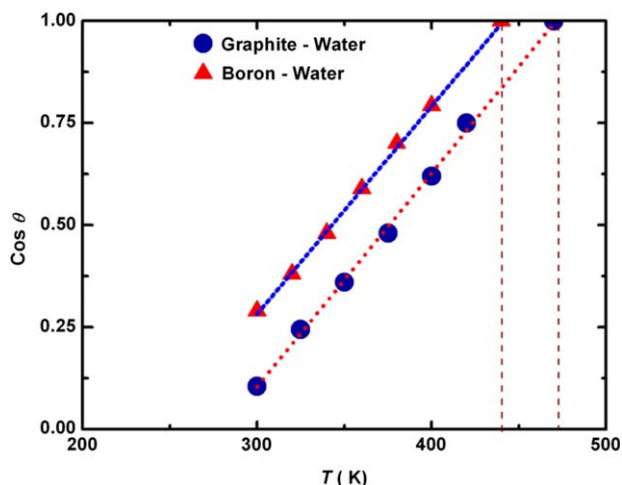
Fig. 15 Dependency of cosine of the macroscopic contact angle θ against temperature for graphite–water and BN-water systems. Dashed lines along the data points are guide to the eyes. Dashed vertical lines represent the estimated wetting temperatures. Reprinted (adapted) with permission from ref. 84. Copyright (2011) Elsevier.

eqn (4). The numerical values for the vapour–liquid surface tension, γ_{lv}, were taken from the study of Vega et al.[87] for the SPC/E water at different temperatures. The calculated values of the line tensions are in the order of 10^{-11} N, and are in good agreement with the line tension values reported in the literature.[88,89] We have obtained 3.06×10^{-11} N as the line tension value at 300 K,[84] which is in good agreement with the previously reported values on graphite–water systems at the same temperature by Werder et al.[90] In a recent work, Zangi et al.[91] assumed a constant line tension for different temperatures for their study of water on a hydrophobic plate. On the contrary, we have found that the line tension substantially increases with increasing temperature. The variation of the line tension for water on graphite and BN surfaces is shown in Fig. 16.

The line tension for water can be modified by adding additives/surfactants. For example, Fig. 17 presents the contact angle of a water–ethanol solution on a graphite surface for different ethanol concentrations at 300 K, for different drop sizes. The variation of the contact angle with the size of the drop is used to evaluate the line tension using eqn (4). Ethanol molecules were modelled using the OPLS force-field.[92] The line tension is found to increase with increasing ethanol concentration. On the other hand, the contact angle of water–ethanol solution is observed to decrease with increasing ethanol concentration. Therefore, it also suggests that the line tension increases as the system approaches complete wetting i.e., with increasing ethanol concentration.

In general, there is a consensus of positive value of the line tension near the first-order wetting transition.[93] At the complete wetting, short-ranged interactions are characterized by a finite τ with finite slope, while the retarded van der Waals interactions exhibit finite τ with diverging slope, and the longer ranged interactions (e.g., non-retarded van der Waals) exhibit diverging τ.
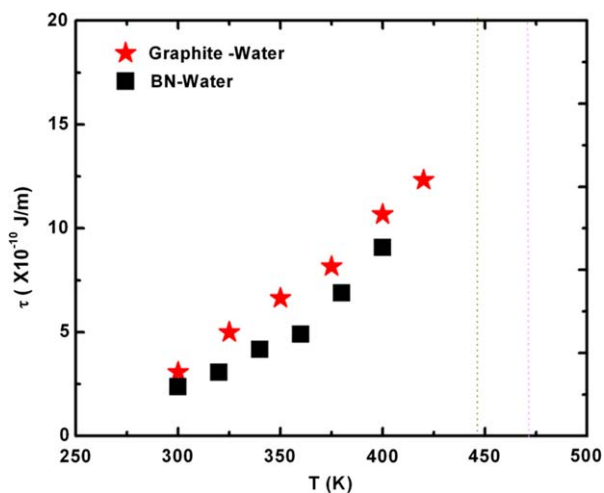
Fig. 16 Line tension *vs.* temperature for graphite–water and BN-water systems. Reprinted (adapted) with permission from ref. 84. Copyright (2011) Elsevier.
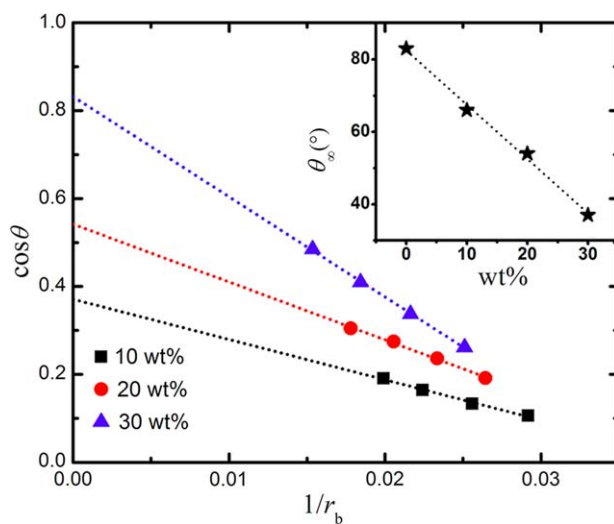
Fig. 17 Dependence of the size of an ethanol–water droplet on its contact angle on a smooth graphite surface, at $T = 300$ K for different ethanol concentrations. Reprinted (adapted) with permission from ref. 92. Copyright (2014) American Chemical Society.

6 Wetting transition of water on soft surfaces

Self-assembly is a fundamental characteristic feature of all living systems, *e.g.* folding of proteins, formation of nucleosomes, endo- and exocytosis *etc*,[94,95] and is one of the most interesting and challenging areas of research. Studies of self-assembly of different kinds of molecules such as alkyl thiols, alcohols, surfactants *etc.*, can help in the further development and implementation of the knowledge in various processes, such as detergency, coating, lubrication,[96] colloidal stabilization, wetting and dewetting,[96] adhesion,[96] corrosion *etc.* Self-assembly of proteins and

other bio-molecules on surfaces has immense importance in the field of biomaterials, biomedical devices,[97] biosensors *etc.*[98] and in the field of micro and nano electronics.[99] Thus, self-assembly is able to generate a vast area of research where a number of questions are still not very clear fundamentally. Several examples of systems and phenomena, related to self-assembly, are well described by Sammalkorpi *et al.*[95] For example, surfactant layers play an important role in protein directed biomineralization in material synthesis. In such cases, inorganic and organic phase interaction and geometric confinement yield many different results. Recent work of Wang *et al.*[100] showed that a well-controlled nanostructure using nanoscale building blocks can yield stable layered nanocomposites using a novel self-assembly process, which can be exploited for various applications.

The self-assembled monolayer (SAM) provides flexibility in the design at molecular and material levels. SAM may also act like a vehicle for investigation of the effect of increasing molecular complexity on the structure (due to specific interactions at interfaces), and stability of two-dimensional assemblies as also shown in a recent work of Wang *et al.*[100] Anchoring and functionalization of surfaces (using SAM) may lead to changes in the properties of the surface. The behaviour of self-assembled monolayers depends on the nature of the substrate, whether hydrophilic or hydrophobic, and its crystallographical plane.[101] The structure and dynamics of monolayers determine many properties like adhesion, molecular and biological recognition.[10] Adsorption of amphiphilic molecules such as surfactants on hydrophilic solids has been studied both experimentally as well as computationally.[101–111] The behaviour of surfactants is strongly governed by the presence or absence of water monolayer on hydrophilic surfaces.[104] The kinetics of adsorption of surfactants is also found to depend strongly on the type of surfactants. For example, the adsorption rate is higher for the cationic surfactants[107] rather than anionic and non-ionic.[103] The general adsorption isotherm of amphiphilic molecules consists of four different regimes:[104]

I. Adsorption increases linearly with concentration of amphiphilic molecules, having a slope of typically one, in a log–log plot.
II. Higher rate of adsorption due to lateral interactions of molecules leading to surface aggregation.
III. Adsorption rate is slower than regime II.
IV. Adsorption isotherm reaches a plateau that depends on many other factors like hydrocarbon chain length, functional group, electrolyte and temperature.

The simplest examples of amphiphilic molecules are *n*-alkanols that are found in abundance in natural and industrial processes. The structure and dynamics of *n*-alkanols are different on different surfaces because of their hydrophobic backbone. For example, on a graphite surface though the backbone of *n*-alkanols remains parallel to the surface, the head group remains removed from the surface.[11–14] This ordering and structure are also observed in case of a mixture of amphiphilic molecules.[102] On the other hand, non-amphiphilic molecules with a hydrophobic

backbone, such as *n*-alkanes, form highly ordered monolayers on surfaces like graphite[111] and silica.[112] The ordering is dependent on the density of the molecules.[113] It has been observed that on a graphite surface the backbone of the alkanes remains parallel to the surface whereas, on a silica surface it can bind to the surface leading to a tilted structure.[112] Higher alkanes, on a graphite surface, lay itself on the surface to accommodate a larger number of tails. On hydrophilic surfaces, like mica, alkanols form tilted monolayer structures which have been identified in AFM images and it was concluded that the tilting increases with increasing chain length.[109] Recently, Cheng *et al.*[114] demonstrated, using molecular dynamics, that alkanols can form clusters on mica surfaces, in presence of moisture, which is observed for the case of ethanol. Such clusters are found to occur above a certain relative humidity for ethanol which is mainly due to the role of hydrogen bonding in the adsorption.[114]

Surface diffusion of the molecules in the monolayer plays an important role in many kinetic processes such as chemical reactions, catalysis, crystal growth, and is responsible for wetting/dewetting of a surface.[115] On an Ag surface, it is seen that the diffusion of water in monolayers increases with increasing surface coverage whereas, on a Pb surface, the diffusion shows a Λ-shaped variation.[115] Park *et al.*[116] has also reported that the surface diffusivity of water in monolayers on graphite also increases up to a critical coverage, and a further increase in the surface coverage leads to a decrease in diffusivity. On a graphite surface, diffusion of *n*-alkanes in the monolayer also exhibits a Λ-shaped anomaly. The larger alkanes diffuse faster in case of lower coverage. On the contrary, smaller alkanes diffuse faster in case of higher coverage.[117] In a recent work, self assembly behaviour of *n*-alkanols on a mica surface was studied using all-atom molecular dynamics.[118] The authors found that the surface diffusion behaviour of *n*-alkanol monolayers is non-monotonic in nature as a function of surface monolayer coverage. The behaviour was correlated with the lifetime of hydrogen bonding of alkanol molecules on the surface. The authors found a clear correlation of the surface diffusion with the size of alkyl group.

In a recent work, Bhandary *et al.*[118] studied the behaviour of *n*-alkanols on a mica substrate. In order to model a mica substrate, the force-field of Heinz *et al.*[119] was used. Furthermore, standard OPLS-AA was considered for *n*-alkanol.[120] Details of the force-field parameter are summarized in the work of Bhandary *et al.*[118]

Figure 18a presents the monolayer adsorption isotherm of *n*-alkanol on a mica surface at a temperature of 300 K and pressure of 1 atm. A similar isotherm behaviour has been seen for water[121] and ethanol[114] on mica surfaces. To understand the nature of an isotherm, one can fit the data to the Langmuir isotherm, which is defined as

$$\sigma_s = \sigma_m \frac{K \cdot c}{1 + K \cdot c} \qquad (24)$$

where σ_s and σ_m represent the amount adsorbed and the maximum amount of adsorption on the surface, respectively, and K is the Langmuir equilibrium constant and the total surface coverage is denoted by c.
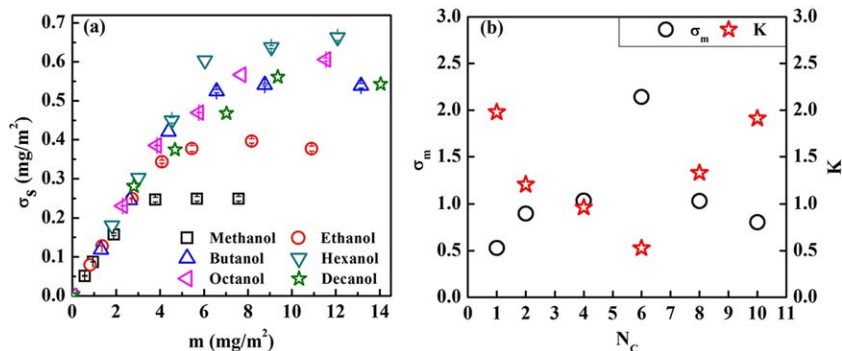
Fig. 18 Adsorption isotherm (panel a) of alkanols on the mica surface at 298 K and 1 atm pressure. The x-axis represents the total surface coverage; whereas the y-axis represents the monolayer surface coverage σ_s. Panel b shows the maximum adsorption (left axis) of alkanols within a monolayer, and the variation of Langmuir equilibrium constant, K (right axis), for different alkanols. Reprinted (adapted) with permission from ref. 118. Copyright (2014) American Chemical Society.

The adsorption isotherm is found to be described very well by eqn (24) indicating strongly that the adsorption of n-alkanols on the mica surface follows Langmuir isotherms. Figure 18b shows σ_m, and K as function of number of carbon atoms, N_c, in n-alkanols. The maximum adsorption increases with increasing N_c until n-hexanol with six carbon atoms (C_6). The maximum at C_6 is well supported by the orientational behaviour, which also peaks at C_6 indicating maximum amount adsorbed for hexanol due to the best packing geometry.[118]

The adsorption of alkanols on surfaces to form monolayers is dependent on the chain length of the backbone, which also holds the key to the orientation of the monolayer. The tilt angle of the monolayer decreases as the chain length of the alkanols increases up to 6, beyond which it increases when increasing the chain length. The adsorption of n-alkanols follows the Langmuir isotherm, and the maximum adsorption within monolayers is found to increase until the case of hexanol, and drops as the backbone is further increased. The surface diffusion coefficient oscillates with increasing surface monolayer coverage. Hydrogen bonding with the surface by the alkanols plays an important role in the diffusion. The longer alkanols diffuse faster at lower concentration whereas the smaller alkanols do so at higher concentrations. The diffusion coefficient varies inversely with the lifetime of the hydrogen bonds with the surface.[118] A similar behaviour is also seen for n-alkanes on graphite surfaces where the Λ-shape behaviour of the diffusion coefficient with surface coverage is attributed to the molecule–surface interactions as well to the rotational motion of the molecules.[23]

While the self-assembled monolayer on surfaces plays an immense role in modifying the substrate nature, direct coating (chemical binding) using soft polymeric matrices can also be used to obtain responsive surfaces. Such a coating with a polymeric matrix provides mechanical flexibility to the system. The behaviour of fluids on a responsive surface where the surface also gets affected by the state condition such as

temperature, pressure or fluid composition is completely different from that seen for the case of a rigid or non-responsive surface.[122] Wetting on the surface having a coating consisting of materials such as polymer chains is entirely dependent on the response of the grafted polymer chain to the fluid and state condition. It has been well established in numerous theoretical,[123–127] computational[128–134] and experimental[135–137] studies that the properties and behaviour of the grafted polymer chains are dictated by the grafting density as well as the molecular weight of the polymer chains.

One of the common polymers used for grafting surfaces, and showing responsive behaviour is poly(N-isopropylacrylamide), PNIPAAm. Since PNIPAAm is an amphiphilic polymer, the coil-to-globule transition is due to the interplay of hydrophobic and hydrophilic factors. It has been shown that PNIPAAm chains contain hydrophobic and hydrophilic domains; above lower critical solution temperature (LCST) the PNIPAAm chains are not completely hydrophobic, while simultaneously they are not completely hydrophilic below LCST.[138] The effect of the grafting density was explored using atomic force microscopy and quartz crystal microbalance techniques.[139] While the conformational changes of grafted PNIPAAm chains became less abrupt with increasing grafting density, these changes start at lower temperatures, well below LCST of free PNIPAAm chains. The influence of both the grafting density and molecular weight of PNIPAAm chains on the temperature-induced conformational transition was reported by Plunkett and co-workers.[140]

Several authors pointed out the existence of abrupt changes in the contact angle measured on the layer of PNIPAAm chains grafted on supporting substrates.[141–143] For example, Sun et al.,[144] demonstrated the existence of reversible switching between superhydrophobic and superhydrophillic using roughness enhanced thermal responsible PNIPAAm modified surfaces. The values of the contact angle reflect only the degree of hydration in the uppermost layer of the grafted chains, whereas the process of dehydration starts in the inner layers.[145] Based on the contact angle measurements, the authors concluded that the change in the contact angle is mainly due to the temperature-induced conformational changes. Molecular modelling of the wetting behaviour of fluids on such responsive surfaces has not been much studied. Recently, Bhandary et al.[146] have studied dynamics of a water droplet on the grafted PNIPAAm chains on a smooth surface, using a combination of molecular dynamics and Monte Carlo simulations.

The details of the molecular model and simulation strategy are given elsewhere.[146] Here we shall instead concentrate on the key results. The grafted PNIPAAm chains and water molecules were simulated by employing a fully flexible atomistic model. The PNIPAAm chains were consisted of 30 N-isopropylacrylamide monomer units, which were anchored to a silicon substrate. The packing density of PNIPAAm on the silicon surface was kept at 1.1493 nm^2 per chain,[147] and was that of the lowest energy. The initial conformation of a simulated system was constructed by placing 36 (12×3) straight PNIPAAm chains perpendicular to the silicon surface (along the z axis). The terminal atom of each PNIPAAm

chain was immobilized. The intra- and intermolecular interactions between the atoms were described by the OPLS force field,[148] and the SPCE model[149] was used for the water molecules. The long-range Columbic interaction was calculated using the Ewald summation technique. The bond lengths and angle of each water molecules were held fixed using the SHAKE algorithm.[150] The position of the Gibbs interface of the PNIPAAm chains was determined using the hyperbolic tangent function of the density (see eqn (19)). A cylinder of water containing 2000 molecules, with the diameter of ~ 2.0 nm and length equal to the width of the grafted PNIPAAm system was equilibrated in a *NVT* ensemble for 2.0 ns at 300 K. The equilibrated water cylinder was placed in the close vicinity of the equilibrated hydrated PNIPAAm layer as shown in Fig. 19a. The systems after 25 ns of MD simulations at different temperatures in the *NVT* ensemble are shown in Fig. 19b–e.

Figure 20 shows the density profile of water and PNIPAAm chains along the normal to the surface at a temperature of 270 K. In order to evaluate the contact angle of the water molecules on the top of a PNIPAAm brush, the interface between the brush and water needs to be defined. Eqn (19) is used to determine the position of the interface of the PNIPAAm layer. This interface is shown in Fig. 20a. The average monomer density is ~ 0.78 g cm^{-3}. The density of water within the PNIPAAM chains is significantly lower (~ 0.375 g cm^{-3}) than the bulk liquid density. However, it swiftly increases across the interface to the bulk value. The density of water sharply decreases to the value of the vapour density with further away from the interface. The Gibbs interface is shifted closer to the silicon surface with increasing temperature as shown in Fig. 20b.

Figure 21 shows the variation of the water contact angle with the temperature. Snapshots of the droplet at different temperatures are also included in the figure. At temperatures below the LCST (~ 32–33 °C), the
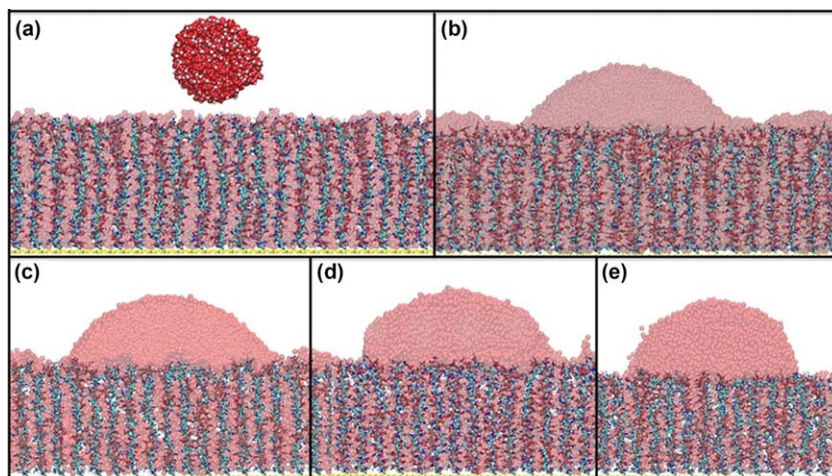
Fig. 19 Snapshots of initial structures of the system of a cylindrical water droplet placed close to a hydrated PNIPAAm layer (a) and of the water droplet morphologies for temperature of 270 K (b), 300 K (c), 310 K (d) and 320 K (e) after 25 ns. Reproduced from ref. 146 with permission from the Royal Society of Chemistry.
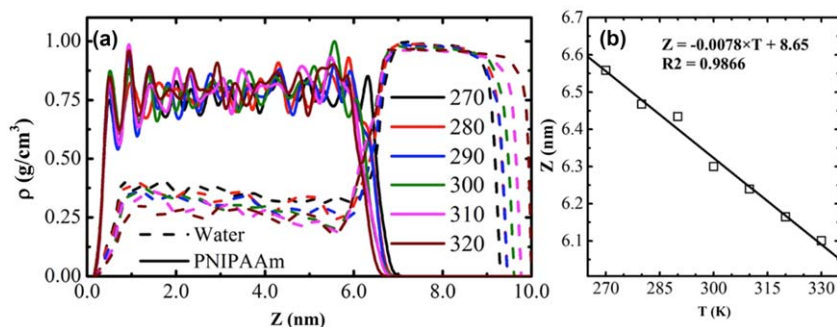
Fig. 20 Density profile of PNIPAAm chains and water along the normal of the silicon surface at different temperatures (a). Solid and dashed colored lines are used to discern the density of PNIPAAm and water, respectively. Figure 3b shows the position of the Gibbs interface as a function of temperature after 25.0 ns. Reproduced from ref. 146 with permission from the Royal Society of Chemistry.
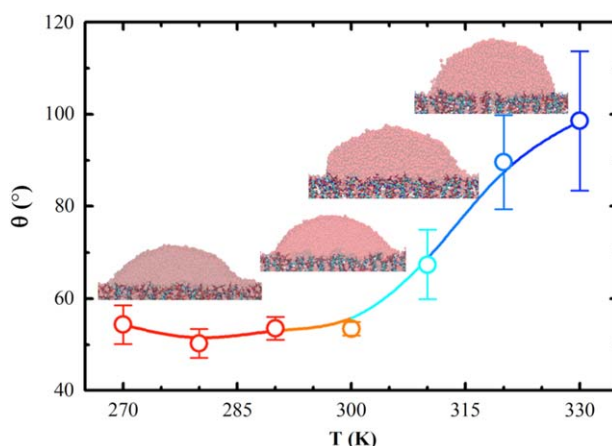
Fig. 21 Variation of the contact angle of a water droplet on the PNIPAAm brush with temperature. Line serves as guide to the eyes. Reproduced from ref. 146 with permission from the Royal Society of Chemistry.

values of the water contact angle on the PNIPAAm interface remain more or less constant and lie in the range of 50°–60°. As the temperature rises above 300 K, the contact angle increases. At 330 K, the water contact angle reaches ∼98°, a value characteristic for hydrophobic materials. Although the temperature-induced variation of the contact angle is more substantial than the temperature-induced conformational changes of the PNIPAAm layer, it is not as steep as has been observed in some experiments.[141,143,151] The experimentally measured values of advancing contact angles of water on the top of terminally grafted PNIPAAm chains increase gradually from ∼50° to almost 90°, within a temperature range of ∼3 °C (21 °C–24 °C), as reported by Takei et al.[141] On the other hand, Balamurugan et al. have reported a transition from 67° to about 78° when increasing the temperature from 31 °C to 32 °C.[151] However, it was noted by Sun et al.[144] that the contact angle behaviour of a water droplet on grafted PNIPAAm chains is sensitive to the underneath roughness of the

base substrate. In case of grafted PNIPAAm chains on a patterned surface, a sharp change in the contact angle of water (∼20° to ∼120°) was observed around LCST by the authors. On the other hand, variation in the contact angle of water was rather more gradual and subdued (60° to 90°) for the case of grafted PNIPAAm chains on a smooth surface. The water contact angles observed for grafted PNIPAAm chains on a smooth surface by Bhandary *et al.*[146] using all-atom MD simulations are in good agreement with the reported experimental values of Sun *et al.*[144]

Bhandary *et al.*[146] used the second virial coefficient, B_2, to understand the nature of the grafted PNIPAAm–water system. For a complex system such as PNIPAAm–water B_2 can be calculated from the potential of mean force (PMF) functions as:

$$B_2 = -\frac{1}{2} \int_0^\infty [\exp(-\beta U_{\mathrm{PMF}}(r)) - 1] 4\pi r^2 \, \mathrm{d}r, \qquad (25)$$

where the PMF calculation is based on the radial distribution function as described elsewhere.[152,153] B_2 quantifies the binary interactions and depends on the potential energy between interacting particles. Favourable binary interactions leading to an attraction are characterized by $B_2 < 0$ (or decreasing trend); whereas $B_2 > 0$ (or increasing trend) indicates repulsive binary interactions. Hence, B_2 values and its behaviour can be used to identify the tendency of the system to agglomerate or disperse. This has been effectively used to quantify the agglomeration/dispersion behaviour of nanoparticles in a polymer matrix in recent works.[152,153] Figure 22 shows the temperature dependence of B_2 for PNIPAAm–water and water–water interactions. B_2 of PNIPAAm–water pair increases with increasing temperature, while B_2 of water–water pair tends to decrease with increasing temperature. This indicates that water molecules prefer to be expelled
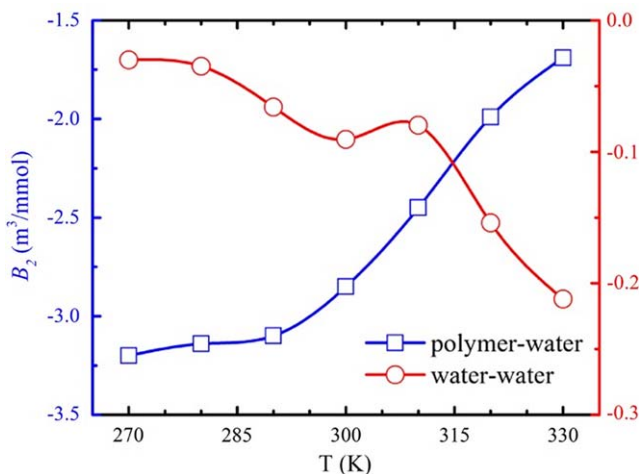
Fig. 22 Temperature dependence of the second virial coefficient, B_2. The left and right ordinates are associated with PNIPAAm–water pairs (squares) and water–water pairs (circles), respectively. All lines serve as guides to the eyes. Reproduced from ref. 146 with permission from the Royal Society of Chemistry.

from the brush moiety with increasing temperature. This is strengthened further by the behaviour of B_2 of water–water, which shows a decreasing trend with increasing temperature, indicating growing attraction between the water molecules. At lower temperatures, the PNIPAAm chains remain hydrated, which is evident from the hydrogen bonding analysis (see ref. 146 for more details), and hence PNIPAAm chains adopt swollen conformations which can also be deduced from the end-to-end distribution (see ref. 146). As the temperature increases, the end-to-end profile shrinks indicating a collapse of the PNIPAAm chains. It also follows that the entropy effects, associated with the release of water molecules from the PNIPAAm brush at elevated temperature, dominate the coil-to-globule transition. On the other hand, water–water pairs contribute to the temperature-induced dehydration of the PNIPAAm brush enthalpically.[146]

7 Conclusion

Molecular simulations have come a long way since the early days of understanding thermophysical properties using simple models. The recent advances in the methods, computing powers and availability of open source codes have fuelled the use of molecular dynamics and Monte Carlo simulations. In this chapter, we first described the prewetting transition using Monte Carlo simulations, and in particularly discussed the possibility of surface phase transitions for molecular associating fluids. Subsequently, we discussed the wetting transition, and in particular highlighted the major difference in the wetting behaviour of water on soft and hard surfaces.

Molecular simulation methodologies to study interfacial properties have matured in recent years, though many groups are still working on sophisticated ways to characterize the interface being hydrophobic or hydrophilic. In addition, with increasing availability of large-scale super-computing platforms, we envisage increasing use of molecular simulations to study the interfacial behaviour of fluids on realistic surfaces.

Acknowledgements

References

1   D. Bonn and D. Ross, *Rep. Prog. Phys.*, 2001, **64**, 1085–1163.
2   J. G. Riess, *Tetrahedron*, 2002, **58**, 4113–4131.
3   M. Kind and C. Woll, *Prog. Surf. Sci.*, 2009, **84**, 230–278.
4   D. Vuillaume, *C. R. Phys.*, 2008, **9**, 78–94.
5   V. Percec, M. Glodde, T. K. Bera, Y. Miura, I. Shiyanovskaya, K. D. Singer, V. S. K. Balagurusamy, P. A. Heiney, I. Schnell, A. Rapp, H. W. Spiess, S. D. Hudson and H. Duan, *Nature*, 2002, **417**, 384–387.
6   L. D. Gelb, K. E. Gubbins, R. Radhakrishnan and M. S. Bartkowiak, *Rep. Prog. Phys.*, 1999, **62**, 1573–1659.

7   S. K. Singh, A. Sinha, G. Deo and J. K. Singh, *J. Phys. Chem. C*, 2009, **113**, 7170–7180.

8   A. Ulman, *Chem. Rev.*, 1996, **96**, 1533–1554.

9   E. Núñez-rojas and H. Domínguez, *J. Colloid. Interface Sci.*, 2011, **364**, 417.

10 J. P. Rabe and S. Buchholz, *Science*, 1991, **253**, 424–427.

11 B. M. Rosen, C. J. Wilson, D. A. Wilson, M. Peterca, M. R. Imam and V. Percec, *Chem. Rev.*, 2009, **109**, 6275–6540.

12 K. Ariga, J. P. Hill, M. V. Lee, A. Vinu, R. Charvet and S. Acharya, *Sci. Technol. Adv. Mater.*, 2008, **9**, 014109.

13 E. Celia, T. Darmanin, E. T. de Givenchy, S. Amigoni and F. Guittard, *J. Colloid Interface Sci.*, 2013, **402**, 1–8.

14 X. Liu, Q. Ye, X. Song, Y. Zhu, X. Cao, Y. Liang and F. Zhou, *Soft Matter*, 2011, **7**, 515–523.

15 W. Barthlott and C. Neinhuis, *Planta*, 1997, **202**, 1–8.

16 C. Ran, G. Ding, W. Liu, Y. Deng and W. Hou, *Langmuir*, 2008, **24**, 9952.

17 Z. Yoshimitsu, A. Nakajima, T. Watanabe and K. Hashimoto, *Langmuir*, 2002, **18**, 5818.

18 D. M. Spori, T. Drobek, S. Zürcher, M. Ochsner, C. Sprecher, A. Mühlebach and N. D. Spencer, *Langmuir*, 2008, **24**, 5411.

19 N. J. Shirtcliffe, G. McHale, M. I. Newton and C. C. Perry, *Langmuir*, 2005, **21**, 937.

20 H. Modaressi and G. Garnier, *Langmuir*, 2002, **18**, 642.

21 F. Heslot, A. M. Cazabat, P. Levinson and N. Fraysse, *Phys. Rev. Lett.*, 1990, **65**, 599.

22 G. O. Berim and E. Ruckenstein, *J. Chem. Phys.*, 2009, **130**, 184712.

23 J. Yaneva, A. Milchev and K. Binder, *J. Chem. Phys.*, 2004, **121**, 12632.

24 M. Lundgren, N. L. Allan and T. Cosgrove, *Langmuir*, 2007, **23**, 1187.

25 G. S. Grest, D. R. Heine and E. B. Webb, *Langmuir*, 2006, **22**, 4745.

26 J. T. Hirvi and T. A. Pakkanen, *J. Chem. Phys.*, 2006, **125**, 144712.

27 V. A. Croutch and R. A. Hartley, *J. Coat. Technol.*, 1992, **64**, 41–52.

28 L. L. Cao, A. K. Jones, V. K. Sikka, J. Z. Wu and D. Gao, *Langmuir*, 2009, **25**, 12444–12448.

29 A. J. Meuler, G. H. McKinley and R. E. Cohen, *ACS Nano*, 2010, **4**, 7048–7052.

30 S. Jung, M. Dorrestijn, D. Raps, S. Das, C. M. Megaridis and D. Poulikakos, *Langmuir*, 2011, **27**, 3059–3066.

31 M. Nosonovsky and V. Hejazi, *ACS Nano*, 2012, **6**, 8488–8491.

32 J. A. van Meel, R. P. Sear and D. Frenkel, *Phys. Rev. Lett.*, 2010, **105**, 205501.

33 G. I. Tóth, G. Tegze, T. Pusztai and L. Gránásy, *Phys. Rev. Lett.*, 2012, **108**, 025502.

34 J. Chen, J. Liu, M. He, K. Li, D. Cui, Q. Zhang, X. Zeng, Y. Zhang, J. Wang and Y. Song, *Appl. Phys. Lett.*, 2012, **101**, 111603.

35 J. K. Singh and F. Muller-Plathe, *Appl. Phys. Lett.*, 2014, **104**, 021603.

36 J. W. Cahn, *J. Chem. Phys.*, 1977, **66**, 3667.

37 C. Ebner and W. F. Saam, *Phys. Rev. Lett.*, 1977, **38**, 1486.

38 R. B. Hallock, *J. Low Temp. Phys.*, 1995, **101**, 31.

39 J. Phillips, D. Ross, P. Taborek and J. Rutledge, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 3361.

40 E. Cheng, G. Mistura, H. C. Lee, M. H. W. Chan, M. W. Cole, C. Carraro, W. F. Saam and F. Toigo, *Phys. Rev. Lett.*, 1993, **70**, 1854.

41 F. Kruchten and K. Knorr, *Phys. Rev. Lett.*, 2003, **91**, 0855021.

42 X. Zhao, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **76**, 0414021.

43 R. Evans and P. Tarazona, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1983, **28**, 1864.

44 T. F. Meister and D. M. Kroll, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1985, **31**, 4055.

45 J. E. Finn and P. A. Monson, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1989, **39**, 6402.

46 W. Shi, X. Zhao and J. K. Johnson, *Mol. Phys.*, 2002, **100**, 2139.

47 M. J. Bojan, G. Stan, S. Curtarolo, W. A. Steele and M. W. Cole, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdisc. Top.*, 1999, **59**, 864.

48 J. R. Errington, *Langmuir*, 2004, **20**, 3798.

49 F. Ancilotto and F. Toigo, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **60**, 9019.

50 J. R. Errington and D. W. Wilbert, *Phys. Rev. Lett.*, 2005, **95**, 226107.

51 M. S. Sellers and J. R. Errington, *J. Phys. Chem. C*, 2008, **112**, 12905.

52 E. M. Grzelak and J. R. Errington, *J. Chem. Phys.*, 2010, **132**, 224702.

53 D. R. Heine, G. S. Grest and E. B. Webb, *Phys. Rev. Lett.*, 2005, **95**, 107801–107804.

54 A. Saha, S. P. Singh, J. K. Singh and S. K. Kwak, *Mol. Phys.*, 2009, **107**, 2189.

55 M. Fitzgerald, R. R. Picard and R. N. Silver, *Europhys. Lett.*, 1999, **46**, 282.

56 B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.*, 1992, **61**, 9–12.

57 A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.*, 1988, **61**, 2635–2638.

58 T. Werder, J. H. Walther, R. L. Jaffe, T. Halicioglu and P. Koumoutsakos, *J. Phys. Chem. B*, 2003, **107**, 1345–1352.

59 T. Werder, J. H. Walther, R. L. Jaffe, T. Halicioglu, F. Noca and P. Koumoutsakos, *Nano Lett.*, 2001, **1**, 697–702.

60 T. Koishi, K. Yasuoka, T. Ebisuzaki, S. Yoo and X. C. Zeng, *J. Chem. Phys.*, 2005, **123**, 204707.

61 W. Mar, J. Hautman and M. L. Klein, *Comput. Mater. Sci.*, 1995, **3**, 481–497.

62 J. Hautman and M. L. Klein, *Phys. Rev. Lett.*, 1991, **67**, 1763–1766.

63 C. F. Fan and T. Cağin, *J. Chem. Phys.*, 1995, **103**, 9053–9061.

64 E. M. Grzelak and J. R. Errington, *J. Chem. Phys.*, 2008, **128**, 014710.

65 M. J. P. Nijmeijer, C. Bruin and A. F. Bakker, *Mol. Phys.*, 1991, **72**, 927.

66 H. Kellay, D. Bonn and J. Meunier, *Phys. Rev. Lett.*, 1993, **71**, 2608–2610.

67 S. Khan and J. K. Singh, *J. Chem. Phys.*, 2010, **132**, 144501.

68 J. E. Finn and P. A. Monson, *J. Chem. Phys.*, 1993, **99**, 6897.

69 J. E. Finn and P. A. Monson, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1989, **39**, 6402.

70 J. R. Errington and D. W. Wilbert, *Phys. Rev. Lett.*, 2005, **95**, 226107.

71 J. K. Singh, G. Sarma and S. K. Kwak, *J. Chem. Phys.*, 2008, **128**, 044708.

72 W. Shi, X. Zhao and J. K. Johnson, *Mol. Phys.*, 2002, **100**, 2139.

73 F. Ancilotto and F. Toigo, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **60**, 9019.

74 J. E. Rutledge and P. Taborek, *Phys. Rev. Lett.*, 1992, **69**, 937.

75 J. A. Phillips, D. Ross, P. Taborek and J. E. Rutledge, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 3361.

76 E. Cheng, M. W. Cole, W. F. Saam and J. Treiner, *Phys. Rev. Lett.*, 1991, **67**, 1007.

77 F. Kruchten and K. Knorr, *Phys. Rev. Lett.*, 2003, **91**, 085502.

78 S. Khan and J. K. Singh, *J. Phys. Chem. C*, 2011, **115**, 17861.

79 S. Khan, D. Bhandary and J. K. Singh, *Mol. Phys.*, 2012, **110**, 1241.

80 D. M. Tsangaris and J. J. d. Pablo, *J. Chem. Phys.*, 1994, **101**, 1477.

81 T. Fukuma, Y. Ueda, S. Yoshioka and H. Asakawa, *Phys. Rev. Lett.*, 2010, **104**, 01601.

82 P. T. Cummings, H. Docherty, C. R. Iacovella and J. K. Singh, *AIChE J.*, 2010, **56**, 842.

83 J. Rafiee, Xi Mi, H. Gullapalli, A. B. Thomas, F. Yavari, Y. Shi, P. M. Ajayan and N. A. Koratkar, *Nat. Mater.*, 2012, **11**, 217.

84   R. C. Dutta, S. Khan and J. K. Singh, *Fluid Phase Equilib.*, 2011, **302**, 310–315.

85   H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.

86   X. Zhao, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **76**, 041402.

87   C. Vega and E. de Miguel, *J. Chem. Phys.*, 2007, **126**, 154707.

88   F. Mugele, T. Becker, R. Nikopoulos, M. Kohonen and S. Herminghaus, *J. Adhes. Sci. Technol.*, 2002, **16**, 951.

89   J. S. Rowlinson and B. W. Widom, *Molecular Theory of Capillarity*, Clarendon Press, Oxford, 1982.

90   T. Werder, J. H. Walther, R. L. Jaffe, T. Halicioglu and P. Koumoutsakos, *J. Phys. Chem. B*, 2003, **107**, 1345–1352.

91   R. Zangi and B. R. Berne, *J. Phys. Chem. B*, 2008, **112**, 8634–8644.

92   A. K. Metya, S. Khan and J. K. Singh, *J. Phys. Chem. C*, 2014, **118**, 4113–4121.

93   J. O. Indekeu, *J. Mod. Phys. B*, 1994, **8**, 309–345.

94   V. Percec, G. Ungar and M. Peterca, *Science*, 2006, **313**, 55–66.

95   M. Sammalkorpi, M. Karttunen and M. Haataja, *J. Phys. Chem. B*, 2007, **111**, 11722–11733.

96   N. Shenogina, R. Godawat, P. Keblinski and S. Garde, *Phys. Rev. Lett.*, 2009, **102**, 1–4.

97   J. G. Riess, *Tetrahedron*, 2002, **58**, 4113–4131.

98   M. Kind and C. Wöll, *Prog. Surf. Sci.*, 2009, **84**, 230–278.

99   D. Vuillaume, *C. R. Phys.*, 2008, **9**, 78–94.

100 D. Wang, R. Kou, D. Choi, Z. Yang, Z. Nie, J. Li, L. V. Saraf, D. Hu, J. Zhang, G. L. Graff, J. Liu, M. A. Pope and I. A. Aksay, *ACS Nano*, 2010, **4**, 1587–1595.

101 E. Núñez-Rojas and H. Domínguez, *J. Colloid Interface Sci.*, 2011, **364**, 417–427.

102 H. Dominguez, *J. Colloid Interface Sci.*, 2010, **345**, 293–301.

103 S. Paria, C. Manohar and K. C. Khilar, *Colloids Surf., A*, 2005, **252**, 221.

104 S. Paria and K. C. Khilar, *Adv. Colloid Interface Sci.*, 2004, **110**, 75–95.

105 C. L. Claypool, F. Faglioni, A. J. Matzger, W. A. Goddard III and N. S. Lewis, *J. Phys. Chem. B*, 1999, **103**, 9690–9699.

106 F. Mugele, S. Baldelli, G. A. Somorjai and M. Salmeron, *J. Phys. Chem. B*, 2000, **104**, 3140–3144.

107 S. C. Biswas and D. K. Chattoraj, *J. Colloid Interface Sci.*, 1998, **205**, 12–20.

108 Z. Bo, W. Chun-Lei, X. Peng and F. Hai-Ping, *Commun. Theor. Phys.*, 2012, **57**, 308–314.

109 L. Wang, Y. Song, B. Zhang and E. Wang, *Thin Solid Films*, 2004, **458**, 197–202.

110 L. Wang, Y. Song, A. Wu, Z. Li, B. Zhang and E. Wang, *Appl. Surf. Sci.*, 2002, **199**, 67–73.

111 K. Morishige, Y. Takami and Y. Yokota, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1993, **48**, 8277–8281.

112 A. B. Sieval, B. V. D. Hout, H. Zuilhof and E. J. R. Sudholter, *Langmuir*, 2000, **16**, 2987–2990.

113 M. W. Roth, M. Kaspar, C. Wexler, L. Firlej and B. Kuchta, *Mol. Sim.*, 2010, **36**, 326–333.

114 T. Cheng and H. Sun, *J. Phys. Chem. C*, 2012, **116**, 16436–16446.

115 J. H. Park and N. R. Aluru, *Appl. Phys. Lett.*, 2008, **93**, 253104.

116 J. H. Park and N. R. Aluru, *J. Phys. Chem. C*, 2010, **114**, 2595–2599.

117 J. H. Park and N. R. Aluru, *Chem. Phys. Lett.*, 2007, **447**, 310–315.

118 D. Bhandary, S. Khan and J. K. Singh, *J. Phys. Chem. C*, 2014, **118**, 6809.

119 H. Heinz, H. Koerner, K. L. Anderson, R. A. Vaia and B. L. Farmer, *Chem. Mater.*, 2005, **17**, 5658–5669.

120   W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

121   A. Malani and K. G. Ayappa, *J. Phys. Chem. B*, 2009, **113**, 1058–1067.

122   T. P. Russell, *Science*, 2002, **297**, 964.

123   S. Alexander, *J. Phys.*, 1977, **38**, 983–987.

124   P. G. d. Gennes, *Macromolecules*, 1980, **13**, 1069–1075.

125   S. T. Milner, T. A. Witten and M. E. Cates, *Macromolecules*, 1988, **21**, 2610–2619.

126   S. T. Milner, T. A. Witten and M. E. Cates, *Europhy. Lett.*, 1988, **5**, 413–418.

127   S. T. Milner, *Science*, 1991, **251**, 905–914.

128   M. Murat and G. S. Grest, *Macromolecules*, 1989, **22**, 4054–4059.

129   P. Y. Lai and K. Binder, *J. Chem. Phys.*, 1991, **95**, 9288–9299.

130   P. Y. Lai and K. Binder, *J. Chem. Phys.*, 1992, **97**, 586–595.

131   A. Chakrabarti, *J. Chem. Phys.*, 1994, **100**, 631–635.

132   D. Bedrov and G. D. Smith, *Langmuir*, 2006, **22**, 6189–6194.

133   Z. Benková and M. N. D. S. Cordeiro, *J. Phys. Chem. C*, 2012, **116**, 3576–3584.

134   F. Goujon, C. Bonal, B. Limoges and P. Malfreyt, *J. Phys. Chem. B*, 2008, **112**, 14221–14229.

135   P. Auroy, L. Auvray and L. Leger, *Macromolecules*, 1991, **24**, 2523–2528.

136   J. Piehler, A. Brecht, R. Valiokas, B. Liedberg and G. Gauglitz, *Biosens. Bioelectron.*, 2000, **15**, 473–481.

137   A. F. Miller, R. W. Richards and J. R. P. Webster, *Macromolecules*, 2001, **34**, 8361–8369.

138   R. Pelton, *J. Colloid. Interface Sci.*, 2010, **348**, 673–674.

139   N. Ishida and S. Biggs, *Macromolecules*, 2010, **43**, 7269–7276.

140   K. N. Plunkett, X. Zhu, J. S. Moore and D. E. Leckband, *Langmuir*, 2006, **22**, 4259–4266.

141   Y. G. Takei, T. Aoki, K. Sanui, N. Ogata, Y. Sakurai and T. Okano, *Macromolecules*, 1994, **27**, 6163–6166.

142   R. Yoshida, K. Uchida, Y. Kaneko, K. Sakal, A. Kikuchi, Y. Sakurai and T. Okano, *Nature*, 1995, **374**, 240–242.

143   J. Zhang, R. Pelton and Y. Deng, *Langmuir*, 1995, **11**, 2301–2302.

144   T. Sun, G. Wang, L. Feng, B. Liu, Y. Ma, L. Jiang and D. Zhu, *Angew. Chem., Int. Ed*, 2004, **43**, 357.

145   H. Yim, M. S. Kent, S. Satija, S. Mendez, S. S. Balamurugan, S. Balamurugan and G. P. López, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2005, **72**, 051801.

146   D. Bhandary, Z. Benková, M. N. D. S. Cordeiro and J. K. Singh, *Soft Matter*, 2016, **12**, 3093–3102.

147   S. G. Lee, T. A. Pascal, W. Koh, G. F. Brunello, W. A. Goddard-III and S. S. Jang, *J. Phys. Chem. C*, 2012, **116**, 15974–15985.

148   W. L. Jorgensen, D. S. Maxwell and J. T. Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

149   H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.

150   J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.

151   S. Balamurugan, S. Mendez, S. S. Balamurugan, M. J. O'Brien-II and G. P. López, *Langmuir*, 2003, **19**, 2545–2549.

152   T. K. Patra and J. K. Singh, *Soft Matter*, 2014, **10**, 1823–1830.

153   T. K. Patra and J. K. Singh, *J. Chem. Phys.*, 2013, **138**, 144901.

Theoretical studies of neutral and charged water clusters

K. D. Jordan* and K. Sen

1   Introduction

Water clusters are found in biological systems, are integral parts of many crystals, and play an important role in atmospheric chemistry. Over the past two decades, much experimental and theoretical research has been done on pure water clusters as well as on water clusters combined with various neutral molecules, ions, and excess electrons. Given the large volume of research that has been carried out on water clusters, it is not possible to provide a comprehensive review of all the work that has been done in this area. Rather, in this review, we will focus on a subset of key developments for three types of water clusters: neutral water clusters, protonated water clusters, and anionic $(H_2O)_n^-$ clusters.

Before addressing each of the three types of water clusters in detail, we first consider some of the common challenges of their experimental and theoretical characterization. To characterize water clusters experimentally, one needs a way of keeping individual clusters isolated from one another, otherwise they will coalesce upon contact. Two common ways of doing this involve trapping the clusters in cold "inert" matrices and using supersonic expansion conditions. Charged clusters have the advantage that clusters of different masses can be separated mass spectroscopically. For each of the three types of clusters, the number of isomers grows rapidly with cluster size, and which isomers are important in a given experiment can depend on growth conditions and the energy content of the cluster at the time it is probed. As the internal energy is increased, the cluster may undergo a finite system analog of a "solid"-to-"liquid" phase transition. For the larger clusters, theoretical identification of the low-energy local minima requires the use of "global optimization" algorithms that avoid being trapped in a particular basin of the potential energy surface. Of particular interest is how the structures, vibrational spectra, thermodynamic and other properties of water clusters of a particular size range are modified by the presence of an excess proton or electron.

2   Neutral water clusters

Research on neutral water clusters has been driven by several considerations. First, as noted above, neutral water clusters play a fundamental role in many biological systems, in certain inorganic and organometallic crystals, and in atmospheric chemistry. Second, water clusters necessarily

*Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, USA.
E-mail: jordan@pitt.edu*

have non-fully H-bonded molecules at their surfaces and thus serve as models of water liquid–vapor interfaces. Third, water clusters can be characterized experimentally and theoretically at a level of detail that is not possible for the bulk, providing important benchmark data for testing and refining more approximate theoretical methods including density functional theory (DFT) and force field methods for use in studying extended water systems.

2.1 Potential energy surfaces and connections with thermodynamics and dynamics

In general, the number of local minima on the potential energy surface (PES) grows exponentially with the number of monomers in the clusters. (The growth is even steeper if one accounts for inversion-permutation symmetry.)[1,2] This has led naturally to several theoretical studies focusing on the relative energies of low-energy minima and reaction pathways for isomerization as well as on the thermodynamical behavior of the clusters. However, with the exception of the smallest clusters (e.g., the dimer and trimer), a detailed knowledge of the potential energy surfaces is only available from calculations using force fields,[1,3] and, even then, most such studies have been carried out using two-body force fields such as TIP4P.[4] While MP2-level[5] calculations with flexible basis sets are feasible for clusters with up to ∼30 monomers, for clusters of this size it would be computationally prohibitive to characterize tens of thousands of minima and transition states as well as the rearrangement pathways at this level of theory.

Detailed explorations of the potential energy surfaces of water clusters, even though typically carried out with relatively simple force fields, have provided considerable insight into isomerization mechanisms. For example, Fig. 1 shows the minimum energy pathway for isomerization
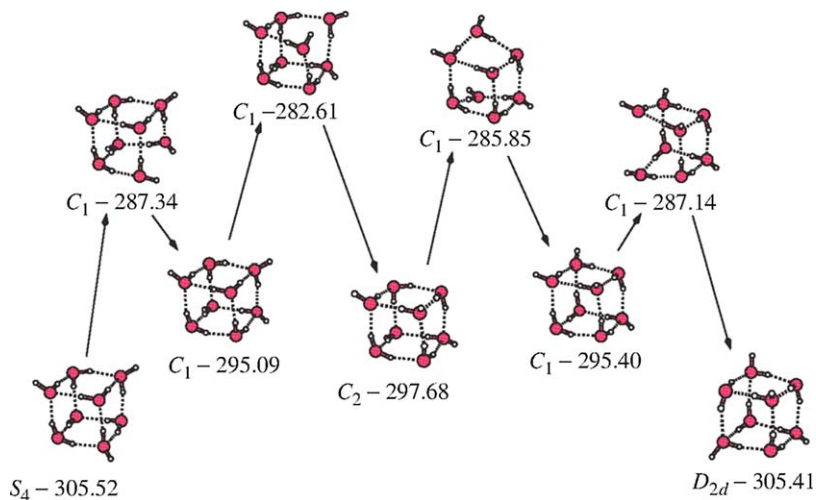
Fig. 1 Minimum energy pathway for interconversion of the S_4 and D_{2d} isomers of $(H_2O)_8$, calculated using the TIP4P model. Energies are in kJ mol^{-1}. Reproduced with permission from ref. 6, Taylor & Francis Ltd., http://www.tandfonline.com.

between the D_{2d} and S_4 cubic structures of $(H_2O)_8$ calculated using the TIP4P potential. These two isomers are known to be nearly isoenergetic and to be energetically well separated from other isomers.[3,7,8] As seen from Fig. 1, the isomerization proceeds *via* a series of monomer rotations with a large ($\sim$5.5 kcal mol^{-1}) overall barrier, consistent with the fact that both isomers have been observed experimentally.[9]

When dealing with clusters with a large number of minima and transition states, it is desirable to have a means of "compressing" the data and presenting it in a form that makes it easy to gain physical insights into the isomerization dynamics and the thermodynamic behavior of the cluster. One of the most useful methods for doing this is the "disconnectivity graph",[10] which makes readily transparent the amount of internal energy required to move from one region of the potential energy surface to another. Figures 2 and 3 display the disconnectivity graphs for $(H_2O)_8$ and $(H_2O)_{20}$,[11] respectively. In these figures, each terminus corresponds to a local minima, and the nodes represent various transition states that have been binned into energy ranges to reduce the number of "nodes". From Fig. 2 it is immediately seen that the two low-energy cubic structures of $(H_2O)_8$ are well separated energetically from any other structures. On the other hand, for $(H_2O)_{20}$, there are several isomers that
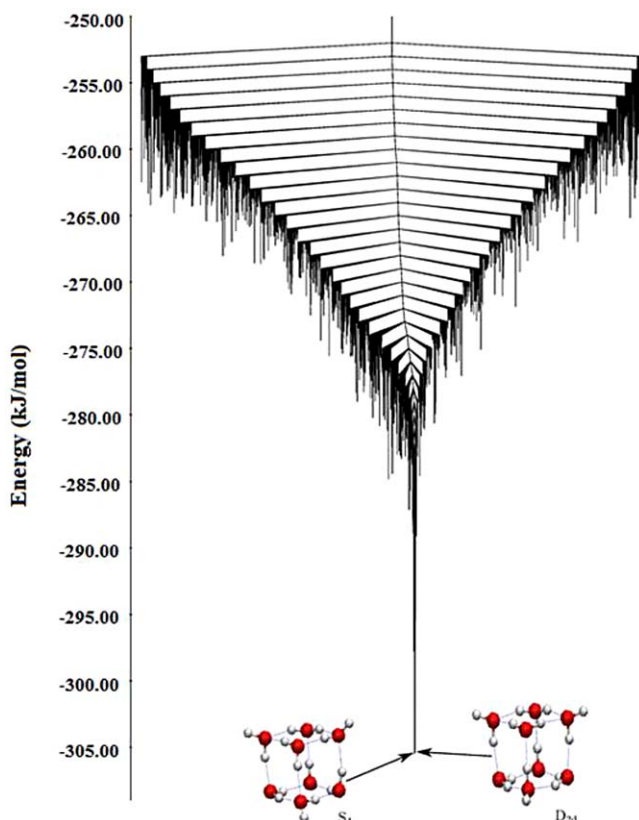
Fig. 2 Disconnectivity diagram for $(H_2O)_8$ as described by the TIP4P potential.
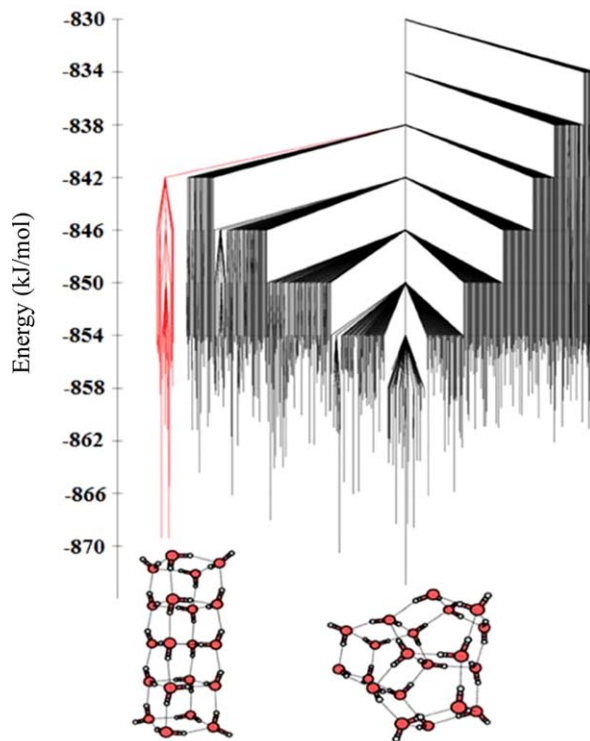
Fig. 3 Disconnectivity diagram for $(H_2O)_{20}$, calculated using the TIP4P model. Reprinted by permission from Macmillan Publishers Ltd: Nature, ref. 11, 2002.

are close in energy but which are separated by high energy barriers. In particular, from Fig. 3, one can immediately see that it requires about 36 kJ mol^{-1} of energy to go from the global minimum to the fused cubic structure.

Inspection of a disconnectivity graph allows one to immediately see whether the system will be challenging for global optimization and for achieving equilibrium in molecular dynamics (MD) or Monte Carlo simulations.[1] If the overall PES is "single funnel-like" in nature, the system will be easy to optimize and to equilibrate, whereas if it is rugged, with multiple funnels separated by high barriers, it will be difficult to locate the global minimum in optimizations and to achieve equilibrium in a finite temperature simulations.[1] Not surprisingly, atomic and molecular clusters are frequently used to test global optimization and simulation methods designed to deal more effectively with quasi-ergodicity issues. In comparing the disconnectivity graphs of $(H_2O)_8$ and $(H_2O)_{20}$, we immediately see that the latter cluster will be much more challenging than the former for global optimization and for achieving equilibrium in Monte Carlo or molecular dynamics simulations. In this context, we note that for $(H_2O)_{20}$ we have found it exceedingly difficult to achieve equilibrium at low temperatures even when using the parallel tempering Monte Carlo method[12] which exchanges configurations
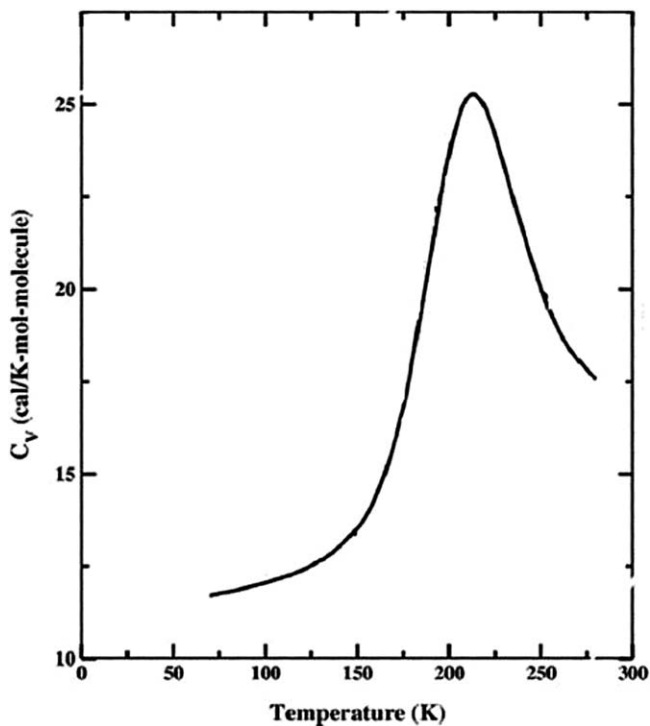
Fig. 4 Heat capacity of $(H_2O)_8$ as a function of temperature calculated using the TIP4P force field and the parallel tempering Monte Carlo method. Reproduced with permission from A. N. Tharrington and K. D. Jordan, *J. Phys. Chem. A*, 2003, **107**, 7380–7389. Copyright 2003 American Chemical Society.

sampled at different temperatures to try to overcome problems associated with trapping in local minima.

Numerous simulations of the properties of neutral water clusters as a function of temperature have appeared.[5,13–19] The major of these simulations have been carried out using force fields and neglecting nuclear quantum effects. Figure 4 reports the heat capacity of $(H_2O)_8$ as a function of temperature, with the results having been obtained using the parallel tempering Monte Carlo algorithm with the TIP4P potential.[14] The heat capacity displays a pronounced peak, indicative of a "solid" to "liquid" phase transition near $T = 210$ K. Analysis of the inherent structures[20] present at different temperatures reveals that the transition is between cubic structures (which dominate at low T) and a high density of non-cubic structures that become important at high T. Both types of structures have significant population near the maximum of the heat capacity curve. In general, it takes only 5–7 kcal mol^{-1} of energy to evaporate a water molecule from an $(H_2O)_n$ cluster. As a result, except at very low temperatures, simulations of water clusters require adoption of algorithms that suppress evaporation.

Nuclear quantum effects might be expected to significantly impact the melting behavior of water clusters due to the fact that the vibrational

zero-point energy (ZPE) can vary appreciably from one isomer to another.[21–24] Indeed, upon inclusion of nuclear quantum effects, the solid-liquid transition in $(H_2O)_8$ becomes much less pronounced and the transition temperature decreases by 10–20 K.[21,22]

2.2 Accurate energies from electronic structure calculations

In this section, we focus on the results for the dimer and hexamer, because these two species have been widely studied and have played a pivotal role in testing and designing force fields for water as well as for assessing the performance of various DFT methods. The global minimum of the water dimer has the structure denoted ''1'' in Fig. 5 in which one water molecule donates an H-bond to the O atom of the other molecule. Based on early experimental work, it was concluded that the binding energy (neglecting vibrational ZPE) is ~ -5.8 kcal mol^{-1},[25,26] and this value was taken into consideration in the parameterization of some early force fields for water. However, as it became possible to carry out near complete basis set (CBS) limit MP2 and coupled cluster singles plus doubles with perturbative triples [CCSD(T)][27] calculations on the water dimer, it became clear that the binding energy (from the potential energy minimum) is $-4.95 - -5.05$ kcal mol^{-1}.[28–31] Subsequently, it was established that the ''experimental'' value of the binding energy from ref. 26 is about 0.8 kcal mol^{-1} too large in magnitude due to the use of harmonic frequencies in correcting for the vibrational zero-point energy.[32–34]

Figure 5 shows for the water dimer nine saddle-point structures in addition to the global minimum (the so-called Smith dimer set).[35]
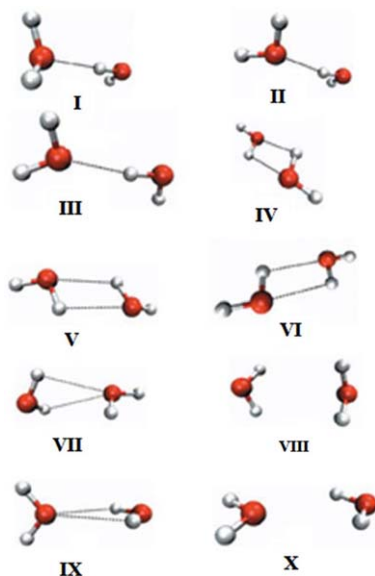
Fig. 5 Global minimum and other stationary states on the water dimer potential energy surface. Reprinted from R. Kumar, F. F. Wang, G. Jenness, and K. D. Jordan, *J. Chem. Phys.*, 2010, **132**, 014309, with the permission of AIP Publishing.
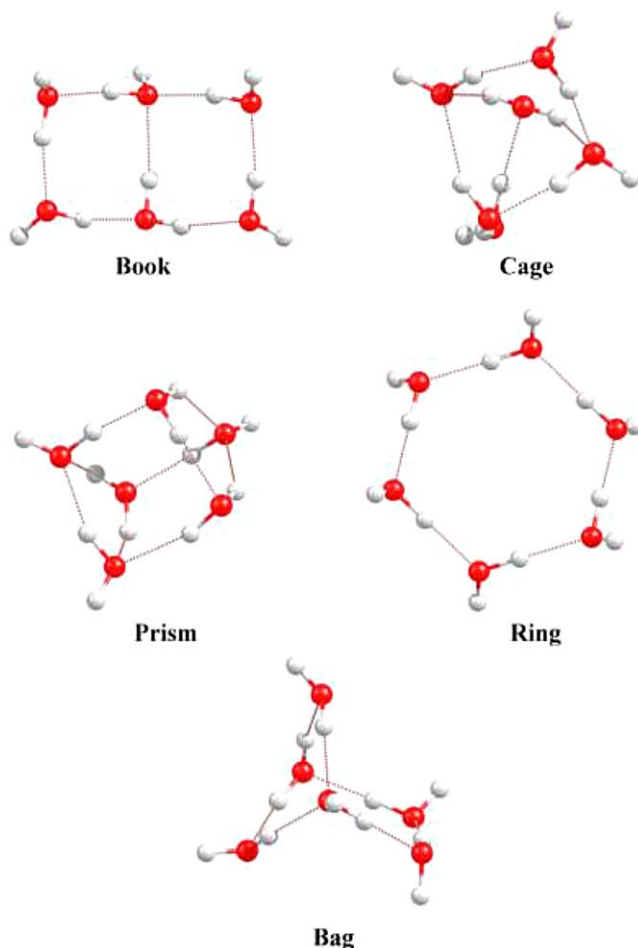
Fig. 6 Low-energy book, cage, prism, ring, and bag isomers of $(H_2O)_6$.

Accurate CCSD(T) energies are available for these stationary points.[36,37] This set of structures has proven especially important in testing force fields and electronic structure methods. Although near CBS-limit calculations with the MP2 and CCSD(T) methods give essentially the same interaction energy of the dimer at its potential energy minimum, the relative energies of the ten stationary points in the Smith dimer set obtained using these two approaches differ by up to 0.2 kcal mol^{-1}.

We now turn our attention to the hexamer which is the smallest water cluster to adopt a three-dimensional structure as its global minimum.[38–42] The majority of the theoretical studies on the hexamer have focused on the lowest energy ring, cage, prism, book, and bag isomers shown in Fig. 6.

Table 1 reports the relative energies of the five selected isomers of $(H_2O)_6$ at the MP2 and CCSD(T) CBS-limits. The CCSD(T) results are presented with and without corrections for vibrational ZPE (calculated using harmonic MP2 frequencies). As is seen from Table 1, the relative

Table 1 Relative energy (kcal mol^{-1}) of five low-energy isomers of $(H_2O)_6$.

	Prism	Cage	Book	Ring	Bag
MP2/CBS[a]	0.00	0.06	0.33	1.21	1.23
CCSD(T)/CBS[a]	0.00	0.25	0.72	1.80	1.62
CCSD(T)/CBS[b] +ZPE	0.00	0.09	0.21	0.51	0.84

[a] From ref. 40.
[b] These results combine the CCSD(T) energies with the harmonic vibrational ZPEs.

energies of the various isomers are impacted by as much as 0.6 kcal mol^{-1} in going from the MP2 to the CCSD(T) method, with the more three-dimensional structures being stabilized relative to the two-dimensional structures (ring and book) upon inclusion of higher-order correlation effects. Moreover, the inclusion of vibrational ZPE corrections compresses the spread in energies of the clusters, leading to the prediction at the CCSD(T)/CBS level of theory that the cage and book structures are, respectively, only 0.09 and 0.21 kcal mol^{-1} less stable than the prism isomer. It is known from the theoretical study of Temelso et al. that inclusion of vibrational anharmonicity makes a negligible impact on the relative energies.[33]

In 2012, Pérez et al.[41] reported the results of charge-pulse Fourier transfer spectroscopy of $(H_2O)_6$ clusters produced in a pulsed supersonic expansion. This study detected the lowest energy prism, cage, and book isomers of $(H_2O)_6$ and provided compelling evidence that the most stable isomer of $(H_2O)_6$ at temperatures near 0 K is the cage isomer which had been detected previously by Liu et al.[43] This is in contrast to the ZPE-corrected CCSD(T) calculations which slightly favor the cage isomer. Foley and Mazziotti[44] have recently reported calculations on the prism and cage isomers of $(H_2O)_6$ using the two-particle reduced density matrix (2PRDM) method[45] which recovers correlation effects missing in CCSD(T). After inclusion of vibrational ZPE effects, theses authors predict the cage isomer to be 0.09 kcal mol^{-1} more stable than the prism isomer consistent with the experiments of Perez et al. The discrepancy between the CCSD(T) and 2PRDM results suggests that the perturbative treatment of triple excitations in the CCSD(T) method may introduce small errors in the relative energies. The low-energy isomers of $(H_2O)_6$ have also been studied using diffusion Monte Carlo (DMC) method.[46] The relative energies from the DMC calculations are in good agreement with the CCSD(T) results, but the statistical errors in the DMC energies are too large to establish definitively the relative stability of the prism and cage isomers. Finally, we note that, due to entropic effects, the book isomer is predicted to be the most stable isomer over the 25–200 K range,[41] which is consistent with it having been detected in earlier experimental studies.[47]

The energies for the water hexamer discussed above were obtained using the supermolecule approach. For water and other cluster systems, there is an alternative method for calculating interaction energies, namely, to make use of the n-body decomposition procedure where the net interaction energy is given by $E_{net} = E_1 + E_2 + E_3 \cdots$, where E_1, E_2, and E_3 represent the one-, two-, and three-body interaction energies,

respectively.[48,49] (E_1 recovers the energy due to geometrical distortion of the monomers and is zero in rigid-monomer calculations). The n-body decomposition method has proven valuable in assessing the performance of various DFT functionals, as well as for testing and parameterizing force fields for water.[36,46,50,51] Force fields for water making extensive use of *ab initio* N-body decomposition data in their parameterization include the DPP2 model of ref. 51 and the CCPOL23+ model of ref. 52.

Analysis of the different N-body contribution reveals that three-body interactions make a sizable contribution (up to ∼20%) of the net interaction energies and that four- and higher-body contributions, while much less important can contribute a few percent to the net interaction energies of water clusters. This means that one can obtain reasonably accurate energies of water clusters (and also of water in condensed phases) from calculation of just the two- and three-body interactions, evaluated at a suitable level of theory. In addition, even more accurate results can be obtained by including an estimate of the four- and, possibly also, higher-body interactions estimated at a low level of electronic structure theory or evaluated with a polarization model. Calculating *ab initio* total energies using such truncated N-body expansions can lead to large reductions in computation effort compared to the more common supermolcule approach. Indeed, Christie and Jordan used this strategy to carry out *ab initio* Monte Carlo simulations of $(H_2O)_6$ as early as 2005.[49] In that work, the two-body interactions were calculated out using the MP2 method, and three-body interactions were carried out using the Hartree-Fock method. More recently, Szalewicz and coworkers used the N-body decomposition method to obtain accurate *ab initio* binding energies for water clusters as large as $(H_2O)_{24}$.[53]

In addition to the methods described above, researchers have generated accurate one-, two-, and three-body multidimensional potentials for water by fitting the energies from CCSD(T) single-point calculations at a large number of strategically chosen geometries on the monomer, dimer, and trimer, respectively.[54,55] In ref. 54 and 55, these have been combined with Thole-type TTM models[56] to describe higher-body polarization contributions. Particularly noteworthy is the WHBB model of Bowman and co-workers[54] which has been used in numerous studies of water clusters and bulk water. This model was subsequently refined by Paesani and co-workers[55] to give the MB-POL model in which the TTM4F polarization model[57] is used to describe the higher-body interactions. Interestingly, Paesani and co-workers found it necessary to include an additional short-range three-body correction term in order to obtain accurate energetics for water clusters.[55] These potentials, while being more complicated than models such as TTM4F (and hence requiring more CPU time to evaluate), have the accuracy needed to predict properties such as anharmonic vibrational spectra.

Recently, embedded N-body expansion methods have been applied to water clusters.[58,59] In these approaches, the energy of each n-mer is calculated in the electric field of the other monomers, which results in the inclusion of polarization effects to infinite order even in the one- and two-body interaction terms.[58,59] In the simplest approach, the charge

distributions of the ''spectator'' molecules are simply represented in terms of atom-centered point charges, although, one can readily employ more accurate descriptions of the charge distributions. The N-body expansion, when recast in this form, converges much more rapidly than the non-embedded variant, with even truncation at the two-body terms giving interaction energies of water clusters quite close to the supermolecule results.[58,59]

2.3   Vibrational spectra

Both microwave[41,43] and vibrational spectroscopy[47,60] have played a major role in identifying the neutral water cluster isomers present under experimental conditions. The majority of the experimental work on water clusters has been carried out using supersonic jet expansions, which generally result in a distribution of cluster sizes and the existence of multiple isomers for each mass. Assigning spectral features to particular isomers under these conditions requires the availability of accurate *ab initio* structures and vibrational frequencies for the low-energy isomers. Unlike charged clusters, where one can readily separate different size clusters, the separation is much harder to accomplish with neutral clusters. In this context, it is interesting to note that there has been limited work using collisions with He atoms to separate different mass water clusters in supersonic jet expansions.[47] In addition, one can accomplish mass separation by introducing a chromophore, such as benzene, which enables obtaining the vibrational spectra through resonance ion dip IR spectroscopy.[61] Indeed, experimental work on benzene-$(H_2O)_n$ clusters was pivotal in establishing that the transition from ring-type to 3D structures occurs at the $n = 6$ cluster.[61] Theoretical calculations of the vibrational spectra of water clusters through $n = 6$ have proven especially important for assigning the observed spectra and for predicting the relative stability of different isomers as a function of temperature.[41,42]

Until recently, most calculations of the vibrational spectra of clusters larger than the dimer were carried out using harmonic approximation. However, in recent years, several studies including anharmonic corrections, generally by means of second-order vibrational PT (VPT2),[62] have been reported for water clusters.[33,63] These allow more reliable assignments of the features in the observed spectra and also have proven valuable for quantifying the role of anharmonic contributions to the zero-point energies.

Over the past few years, several methods have appeared to facilitate calculations of vibrational spectra on large systems. Tschumper and co-workers have used the N-body decomposition procedure to optimize geometries and to calculate vibrational spectra of water clusters.[64] With this approach, one can obtain large basis set CCSD(T) quality harmonic spectra at a fraction of the computational cost of the corresponding supermolecule calculations. This strategy can be readily combined with the various methods for calculating anharmonic corrections to the frequencies. Several groups have introduced local mode approaches

which can also greatly reduce the computational effort of anharmonic vibrational spectra calculations.[65–67] The local mode approaches can be used in conjunction with force constants obtained using standard quantum chemistry methods or using potential energy surfaces such as WHBB for MB-POL described above.

2.4 Force fields for water

More models have been introduced for water–water interactions than for any other substance, and a comprehensive overview of water force fields is beyond the scope of this article. Here we focus instead on some of the most important recent developments in this area. There are multiple ways of classifying water models, including whether they are parameterized (at least in part) to experiment or are totally based on *ab initio* calculations, whether they employ rigid or non-rigid monomers, and whether or not polarization is treated explicitly. Although non-polarizable water models have proven useful for understanding many properties of water, in order to accurately describe water in different environments, *e.g.*, in going from the bulk to the liquid–vapor or at the liquid–protein interface, it is essential to include explicit polarization.

The challenge of accurately predicting the properties of water in different environments (*e.g.* bulk, liquid, liquid–vapor, and liquid–protein interfaces) is reflected, in part, by the problem of correctly predicting the relative energies of different low-lying isomers of $(H_2O)_6$. Table 2 summarizes the interaction energies of the cage, prism, book, and ring isomers calculated using the TIP4P and several polarizable force fields for water.[51,69–72] With the exception of the AMOEBA[71] and TTM3-F models,[70] all force fields listed were designed as rigid monomer; in the case of the AMOEBA and TTM3-F the reported energies were obtained using rigid monomer constraints. The CCSD(T) energies used for reference are essentially aug-cc-pV5Z[73,74] basis set quality and were also obtained using frozen monomers.[51] From Table 2 it is seen that the TIP4P method overbinds all of the isomers, largely as a result of the enhanced dipole

Table 2 Performance of selected force fields for the water hexamer: Binding energies (in kcal mol^{-1}).[a]

Isomer/RMSD	TIP4P	DC[b]	TTM3-F[c]	AMOEBA[d]	CC-POL[e]	DPP2	CSSD(T)
Cage	− 47.27	− 40.85	− 42.80	− 44.90	− 44.75	− 45.90	− 45.12
Prism	− 46.91	− 41.00	− 43.72	− 44.54	− 45.41	− 44.75	− 45.26
Book	− 46.12	− 40.43	− 41.94	− 44.48	− 43.90	− 44.89	− 44.49
Ring	− 44.38	− 39.39	− 41.13	− 43.52.	− 42.91	− 43.58	− 43.64
RMSD	1.53	4.21	2.36	0.38	0.51	0.50	0.00
Range[f]	2.89	1.61	2.09	1.48	2.50	2.32	1.62

[a] TIP4P results from ref. 68; other results from ref. 51.
[b] Ref. 69.
[c] Ref. 70.
[d] Ref. 71.
[e] Ref. 72.
[f] The range is the energy spread between the most stable and least stable of the four isomers considered.

moments on the monomers to effectively account for polarization. More importantly, there are sizable errors (up to 2.9 kcal mol^{-1}) in the relative energies of the various isomers, as described with the TIP4P model, illustrating the limitations of non-polarizable models for describing water in different environments. Surprisingly, both the polarizable Dang-Chang (DC)[69] and TTM3-F models have even larger RMSDs than the TIP4P model. This is largely a consequence of these models significantly underbinding (in magnitude) all the isomers; both the DC and TTM3-F models fare better at predicting relative energies than does the TIP4P model. The other three polarizable models – AMOEBA, CC-POL,[72] and DPP2 – all give much improved values of the interaction energies with RMSDs of 0.51 kcal mol^{-1} or smaller.

Even when polarizability is explicitly included in a water model, it is common to use simple point-charge models to describe the permanent charge distributions of the monomer. In the simplest models, the charges are atom-centered, but in models such as DC, TTM3-F, and DPP2, the negative charge is displaced off the O atom toward the H atoms to better describe the quadrupole moment. Comparison of the electrostatic potentials of the water monomer from point-charge models with that from MP2 calculations shows that the point-charge models fail to accurately describe the electrostatic potential at distances somewhat greater than the van der Waals surface.[75] Thus, the success of such models depends on a cancellation of errors in the different terms in the potential. While this cancellation may be near complete for geometries close to minimum energy structures, it can be far from complete at other geometries. We have encountered this problem with the DPP2 model which fairs poorly for some $(H_2O)_n$ structures extracted from complexes of Cl^- with water clusters.

An improved description of the electrostatics is provided by use of a distributed (atom-centered) multipole expansion[76] of the charge distribution, generally including atomic moments through the quadrupole or higher (and, possibly also, multipole moments at non-atomic sites). Models of this type include AMOEBA, SIBFA[77] and ASP-W4.[78] While the use of distributed multipoles rather than simple point-charge models generally results in considerable improved electrostatics at intermediate distances, it does not address the charge-penetration[76] contribution that results from the overlap of the charge distributions of separate monomers. Charge penetration effects can be sizable. Although most force fields for water do not explicitly include terms for charge penetration, it is implicitly included by scaling of the short-range repulsive terms in the potential. The SIBFA force field does include explicit charge penetration terms and has been proven to provide an accurate representation of the electrostatic interaction between water (or other) molecules over a wide range of distances and orientations.

An even more realistic approach to modeling short-range electrostatic interactions is to use continuous charge distributions (most commonly represented as Gaussians) rather than point multiple moments.[79,80] Such models have the advantage of automatically including charge penetration, and also can incorporate exchange-repulsion (via the density

overlap integral[81]). One of the most sophisticated family of models of this type is the Gaussian electrostatic moment (GEM) models of Piquemal and co-workers.[79,80] In the most elaborate GEM models, the atomic fitting sets contain s-, p-, and d-type Hermite Gaussians. The GEM model can provide very accurate representations of the electrostatic and exchange-repulsion interactions but, obviously, at increased computational cost, compared to models employing point multipoles to describe the electrostatic interactions.

2.5 Tunneling Dynamics and Spectroscopy

A large body of work – both theory and experiment – has been done on the tunneling spectroscopy of small water clusters.[82–86] This work has been covered in other reviews,[83,85] and we note here only that close reproduction of the observed tunneling splittings provides a stringent test of water force fields. One of the most recent and exciting developments concerning tunneling in water clusters is the discovery that the prism form of $(H_2O)_6$ undergoes an isomerization that involves concerted rotation of two water monomers.[87] The observed spectrum displays a pair of triplets due to tunneling. Although the pathway for this process involves a barrier that is appreciably higher than that on the minimum energy pathway (which involves rotation of one water molecule at a time), tunneling "shortcuts" the high barrier on the concerted pathway, making the process observable. This raises the question as to whether such concerted tunneling processes occur in larger clusters or even at water interfaces.

3 Anionic water clusters

The nature of excess electron states of $(H_2O)_n$ clusters has long fascinated researchers. The water monomer has neither a bound anion nor a low-lying temporary anion state associated with electron capture into a vacant valence orbital.[88] Hence water clusters do not possess bound valence-type anions. On the other hand, water clusters with the molecules arranged so as to present an extended region with a sufficiently attractive electrostatic potential can bind an excess electron due to the electrostatics interactions. For example, clusters with dipole moments greater than ~2.5 D will have stable dipole-bound anions.[88,89] Thus, even the water dimer, with a dipole moment of ~2.7 D, has a dipole-bound anion. However, the excess electron in $(H_2O)_2^-$ is very weakly bound, with the vertical electron detachment energy being only ~45 meV.[90] Dispersion-like correlation interactions between the excess electron and the water cluster can lead to electron binding even if the electrostatic interactions are insufficient to give binding.[91]

Experimental confirmation of the stability of $(H_2O)_n^-$, $n \geq 2$, clusters was originally provided mass spectroscopically.[92–94] Figure 7 reproduces a mass spectrum of $(H_2O)_n^-$ clusters from the Johnson group[95] from which it is seen that the intensities of the $n = 2$, 6, 7, and $n \geq 11$ clusters are much greater than for the $n = 3$–5, 8–10 clusters. The intensity pattern for the $n \leq 10$ clusters correlates well with the magnitude of the dipole
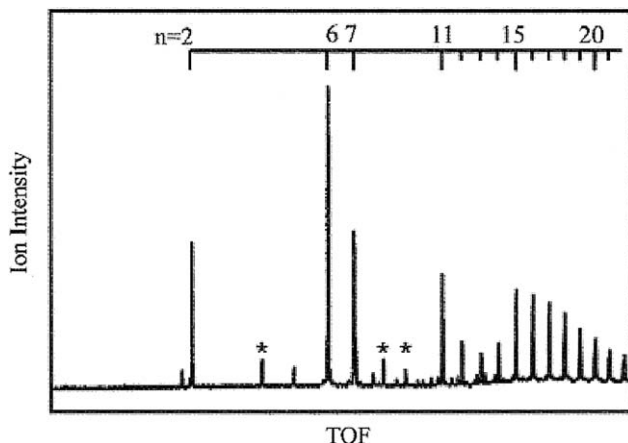
Fig. 7 Time-of-flight mass spectrum of $(H_2O)_n^-$ clusters. The asterisks indicate signals due to complexes containing Ar atoms. Reprinted from J. Kim, I. Becker, O. Cheshnovsky and M. A. Johnson, *Chem. Phys. Lett.*, 297, 90, Copyright (1998), with permission from Elsevier.

moment of the most stable isomers of the corresponding neutral clusters: *i.e.*, neutral clusters with sizable dipole moments have intense signals in the mass spectrum, while those with zero or near zero dipole moments appear only weakly or not at all, in the case of $n = 4$. In a later paper, the Johnson group showed that the missing $(H_2O)_4^-$ ion can be prepared by condensation of a water molecule onto the Ar-tagged $(H_2O)_3^-$ cluster.[96]

Sufficiently large water clusters can bind an excess electron even if the net dipole moment is very small or zero, with the binding coming from a combination of electrostatics and dispersion interactions between the excess electron and the electrons of the water molecules.[88,97] For large water clusters, the excess electron can be accommodated in a variety of ways, with the two limiting cases having the excess electron bound on the surface or in the interior in a so-called cavity state.[98–100] The latter is of particular interest since the prevailing picture of the hydrated electron (e_{aq}^-) in bulk water is that of an electron bound in an approximately spherical cavity with OH groups of the water molecules in the first hydration shell pointing toward the cavity interior.[101,102] Thus the question of the minimum size water cluster for which a cavity-bound anion state can be formed has attracted considerable attention.[103–107] Several of the experimental studies of $(H_2O)_n^-$ clusters were carried out using supersonic jet expansions with the resulting neutral clusters being cold $(T \lesssim 180 \text{ K})$ and solid-like. Electron attachment to such cold clusters may not give rise to detectable cavity-bound anion states even were they thermodynamically stable because the time scale for rearrangement to the cavity-bound anion may exceed that probed in the experiment. Verlet and co-workers have shown that clusters prepared under different expansion conditions can have significantly different vertical detachment energies, and interpreted these results in terms of formation of both surface-bound and cavity-bound anions.[105] However, this interpretation

has been questioned by Turi and co-workers.[104] In still more recent work, Ma and co-workers revisited this problem and concluded that, based on their experimental results, the cavity state first appears in the $n = 25$–30 cluster-size range.[106]

Early theoretical work on $(H_2O)_n^-$ clusters made use of one-electron models due to the fact that accurate *ab initio* treatments would not have been possible for any but the smallest anionic clusters with the algorithms and computers available at the time.[103,108-110] These early studies combined a model Hamiltonian describing the interaction of the excess electron with the water molecules with a two-body model such as the simple point charge (SPC) model[111] for the water–water interactions. In general, the model Hamiltonians used the same atomic charges as employed in the water force field, included polarization terms *via* an adiabatic polarization potential, and included repulsive pseudopotentials to incorporate charge penetration and orthogonalization effects as well as to model exchange interactions in those cases that an explicit term for exchange was not employed. However, at the time these models were developed, there was only limited experimental data and no accurate *ab initio* calculations on $(H_2O)_n^-$ clusters to guide the parameterization. Moreover, the force fields employed for describing the interactions between the neutral water molecules, while proving to be quite successful at describing many properties of liquid water, are now known to fare poorly at predicting relative energies of the neutral water clusters, especially at the geometries to which they are distorted in the presence of the excess electron.[97]

In recent years, several groups[97,112,113] have reported accurate *ab initio* binding energies for small to medium sized $(H_2O)_n^-$ clusters. The availability of these results has facilitated the design of *ab initio* inspired model Hamiltonians for describing excess electron–water cluster systems.[97,114-117] At the present time, two of the most accurate models for describing electron-water cluster systems are those from our group[116] and from the Herbert group.[117] Both of these models allow for consistent treatment of water–water and electron–water polarization and are based on water models that describe the neutral clusters more accurately than simple non-polarizable models. The Schwartz group has also recently introduced a one-electron model (LGS) Hamiltonian for electron–water interactions.[118] Simulations with the LGS model have led Schwartz and co-workers to question the validity of the cavity model for an excess electron in bulk water. However, this conclusion has been questioned by other researchers.[119,120] What is clear from these recent studies is that the structure of the water network in the vicinity of the excess electron depends sensitively on the details of the model Hamiltonian for describing the interaction of the excess electron with the water monomers. We note also that the LGS model uses the SPC force field for water which, as noted above, is not well suited for describing the neutral water clusters, and, therefore, may not be adequate for describing the water structure in the vicinity of the excess electron.

For water clusters for which the Hartree-Fock method does bind the excess electron, the MP2 method generally gives EBE values within

10% of those from higher level (*e.g.*, CCSD(T)) calculations.[97,113] However, for clusters for which the Hartree-Fock method does not bind the excess electron, one needs to use methods that do not depend on the Hartree-Fock method providing a suitable starting wave function for characterizing the anion. Two such methods are the electron affinity equation-of-motion (EA-EOM)[121,122] and algebraic diagrammatic construction (ADC).[123] In the former case, both the EA-EOM-MP2[121] and EA-EOM-CCSD[122] variants have been employed, and, in the latter case the second-order-self energy or ADC2 variant, which is based on a second-order approximation to the self energy, has been employed.[123] The EA-EOM-CCSD method first carries out a coupled cluster singles–doubles (CCSD) calculation on the neutral molecule, with the resulting doubles amplitudes being used to carry out a similarity transform of the Hamiltonian. The transformed Hamiltonian is then used to carry out a one-particle plus two-particle-one-hole CI on the anion. The EA-EOM-MP2 method is similar except that the initial calculation on the neutral molecule is carried out at the MP2 level, and the MP2 doubles amplitudes are used to carry out the similarity transform. Vysotskiy *et al.* have reported EOM and ADC2 EBEs for selected water clusters up to $(H_2O)_{20}$ in size.[111]

One of the best characterized $(H_2O)_n^-$ clusters is $(H_2O)_6^-$, the most stable isomer of which has the structure depicted in Fig. 8.[113,114,124,125] In this isomer the excess electron has considerable charge density near the double acceptor (AA) monomer located on top of a four-membered water ring. From comparison of the structure of the AA anion with those of the isomers of the neutral cluster depicted in Fig. 6, it is seen the excess electron causes an extensive rearrangement of the H-bonding network of the cluster. Specifically, the excess electron causes several of the OH
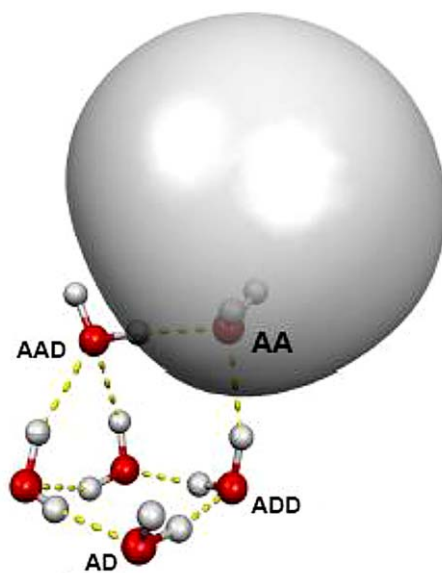

Fig. 8 Dominant isomer of $(H_2O)_6^-$ observed in experimental studies.

groups of the monomers to point roughly in the same direction, generating a large dipole moment. In fact, the neutral cluster at the anion geometry has a ~12D dipole moment. Although it is energetically unfavorable for the neutral cluster to be distorted into a structure with such a large dipole moment, the distortion energy is more than compensated for by the enhanced binding energy of the excess electron. The resulting anion has a vertical detachment energy of about 0.45 eV,[125] but is only slightly more stable than the most stable isomer of the neutral cluster.

In a particularly elegant experiment, using population modulation electron attachment spectroscopy and Ar tagging, Diken et al. have established that the dominant $(H_2O)_6^-$ isomer observed experimentally results from electron attachment to the book isomer of the neutral cluster.[127] Choi and Jordan have used the one-electron polarization potential model Hamiltonian of ref. 114 to map out the minimum energy pathway from the neutral book isomer of $(H_2O)_6$ plus a free electron to the dominant $(H_2O)_6^-$ isomer observed experimentally.[126] These calculations showed that, although an excess electron does not bind to the undistorted book isomer, the interaction of the electron with the cluster can flip free OH groups, to give a sufficiently large dipole that the electron becomes weakly dipole bound. From the distorted book structure, the anion undergoes a sequence of rearrangements leading to the product structure shown in Fig. 8. The minimum energy pathway elucidated by Choi and Jordan has three non-AA local minimum structures between the starting book isomer and the AA product species, with the non-AA structures being calculated to lie lower in energy than the most stable AA species. This is a puzzling result, since the experiment was carried out with Ar-tagged clusters, and Ar atom evaporation would be expected to result in trapping in the non-AA minima. This raises the possibility that pathways other than the minimum energy pathway may be responsible for forming the AA anionic clusters from the neutral book isomer. Further theoretical work is required to determine whether this is indeed the case. The vibrational predissociation studies of Rosciola et al. provide evidence for AA-type anions for clusters as large as $n = 24$.[128]

Excitation of an OH stretch vibration of a $(H_2O)_n^-$ cluster deposits ~0.4 eV of energy into the cluster, which is enough energy to cause evaporation of one or two water monomers, and for small $(H_2O)_n^-$ clusters, is also enough energy to open up the electron autoionization channel. In clusters with Ar atom tags, the evaporation of the Ar atoms suppresses the water loss following vibrational excitation. Using pump-probe experiments on the Ar-tagged cluster, Gauss et al. have been able to follow isomerization of the $(H_2O)_7^-$ cluster following vibrational excitation.[129] It is also possible to use vibration excitation to drive electron transfer and proton-coupled electron transfer reactions. An example of the former is the $(H_2O)_6^- \cdot CO_2$ cluster, in which the $(H_2O)_6^-$ portion of the complex has the structure depicted in Fig. 8 above.[130] Even though the reaction to produce $(H_2O)_6 \cdot CO_2^-$ is exothermic, it does not occur on the time scale of the experiment in the absence of vibrational excitation. Excitation of one quanta of OH stretch or HOH bend of a water

monomer or of the asymmetric CO stretch of the CO_2 permits the electron transfer reaction to be observed experimentally, which indicates that the barrier for rearrangement is small ($\lesssim 4\,\mathrm{kcal\,mol^{-1}}$). Born-Oppenheimer molecular dynamics (BOMD) using DFT energies and forces carried out on $(H_2O)_6 \cdot CO_2$ revealed that the electron transfer reaction is triggered by structural fluctuations that point an OH group toward the CO_2 molecule. Such fluctuations drive the reaction because the H-bond to the CO_2 makes the latter a better electron acceptor. Although these BOMD simulations proved useful in elucidating the reaction mechanism, it would be valuable to repeat them using the MP2 method in light of the problems that DFT methods have in describing anions.[131]

4 Protonated water clusters

The reaction of H_2O and H^+ to form hydronium $(H_3O)^+$ is exothermic by 7.1 eV.[132] In spite of the high stability of the hydronium ion, excess protons are known to be highly mobile in liquid water, and the Zundel ion $(H_5O_2^+)$ with the excess proton shared between the O atoms of two flanking water molecules was introduced to account for the high mobility.[133] The H_3O^+ ion strongly binds water molecules, and, with a complete first solvent shell, one has the so-called Eigen $(H_9O_4^+)$ ion.[134] Although it has proven difficult to unambiguously identify these limiting structures in liquid water, experimental studies on cold $H^+(H_2O)_n$ clusters have provided spectroscopic (vibrational) signatures of the Eigen and Zundel ions and, moreover, have identified species that are intermediate between Eigen and Zundel in character.[135–140] The experimental vibrational spectra of the $n = 2$–11 protonated water clusters are reproduced in Fig. 9.[135] The features associated with the shared proton "stretch" vibration of the $H_9O_4^+$ (Eigen) and $H_5O_2^+$ (Zundel) clusters occur near 2650 and 980 $\mathrm{cm^{-1}}$, respectively. To a large extent, the large spectral red shift of the shared proton vibration in going from the Eigen to the Zundel ion can be understood in terms of the calculated one-dimensional OH stretch potential energy curves for $H_9O_4^+$ and $H_5O_2^+$ depicted in Fig. 10.[140] As expected, the potential energy curve for $H_5O_2^+$ is much more anharmonic than that for $H_9O_4^+$. Moreover, it has a large quartic character a result of which neither the harmonic approximation nor second-order vibrational perturbation theory (VPT2) treatment is adequate for describing the "parallel" stretch vibration of $H_5O_2^+$.

The broad shared proton stretch potential of $H_5O_2^+$ can be viewed as resulting from the mixing of $H_3O^+ \cdots H_2O$ and $H_2O \cdots H_3O^+$ diabatic (or valence bond) configurations. At long OO distances, the resulting adiabatic potential has two minima with the "shared" proton localized on one or the other water molecule, while at short OO distances, including that (~ 2.4 Å) of the minimum energy Zundel structure, the ground state potential energy curve has a single, broad minimum. In a recent paper, McKenzie modeled the shared proton stretch potential of $H_5O_2^+$ in terms of two coupled Morse potentials.[141] The effective valence bond (EVB) and multistate EVB (MSEVB) approaches for treating $H^+(H_2O)_n$ clusters and
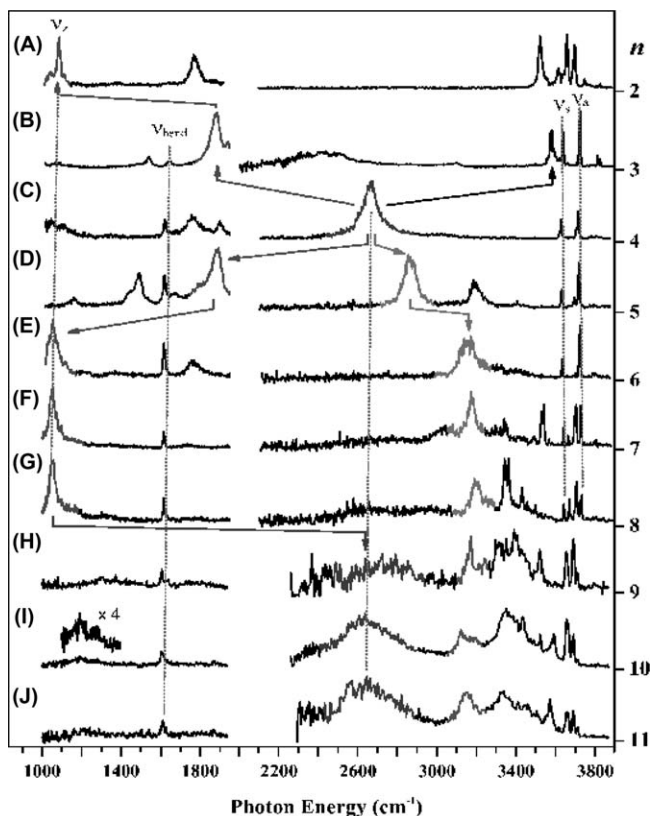
Fig. 9 Ar-atom predissociation spectra of the $n = 2$–11 protonated water clusters. From J. Headrick, E. G. Diken, R. S. Walters, N. I. Hammer, R. A. Christie, J. Cui, E. M. Myshakin, M. A. Duncan, M. A. Johnson, and K. D. Jordan, *Science*, 2005, **308**, 1765–1769. Reprinted with permission from AAAS.

excess protons in bulk water are based on the strategy of associating the excess proton with different valence-bond structures.[142,143]

For the Eigen ion, the potential energy curve for one-dimensional OH stretch (in the local mode approximation) can be viewed as arising from interacting $(H_2O)_2H_3O^+ \cdots H_2O$ and $(H_2O)_2H_2O \cdots H_3O^+$ diabatic states, where the former is significantly more stable. It is interesting to consider the nature of the ions formed by removing one water monomer from or adding one water monomer to the $H_9O_4^+$ Eigen ion. In the former case, the two diabatic states are $(H_2O)(H_2O) \cdots H_3O^+$ and $(H_2O)H_3O^+ \cdots H_2O$, while for the latter case, the two diabatic states are $(H_2O)_2H_3O^+ \cdots H_2O(H_2O)$ and $(H_2O)_2H_2O \cdots H_3O^+(H_2O)$. In both cases, the energy difference between the two diabatic states at their potential energy minima is significantly less than between the corresponding diabatic states for the Eigen ion. As a result, $H_3O^+(H_2O)_2$ and $H_3O^+(H_2O)_4$ display intense shared proton stretch vibrations at energies roughly intermediate between the corresponding vibrations of the Zundel and Eigen ions. By adding two water molecules to one of the water molecules of $H_9O_4^+$, the system evolves to the symmetrical $H_5O_2^+(H_2O)_4$ Zundel-like ion. These
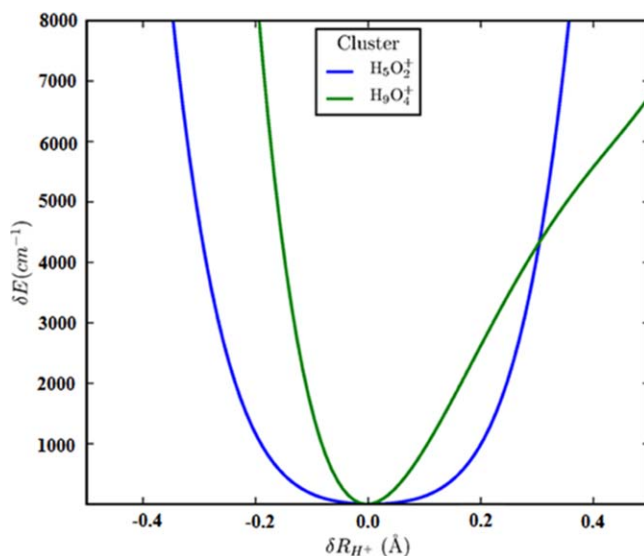
Fig. 10 One dimensional potential energy curves for displacement of the shared proton in $H_5O_2^+$ and of one of the shared protons of $H_9O_4^+$. The potentials were calculated at the MP2/aug-cc-pVDZ[74,75] level, keeping all geometrical parameters other than that of the shared proton at the values they have at the potential energy minima.

results show that, depending on the environment of the cation, there is a continuous evolution from the Eigen to the Zundel ion.[140]

It has been known since 1974 that the mass spectrum of protonated water clusters, $H^+(H_2O)_n$ displays magic numbers, the most pronounced of which is the $n = 21$ cluster.[144,145] Searcy and Fenn, in their 1974 paper, proposed that the magic number $n = 21$ cluster is comprised of an $(H_2O)_{20}$ dodecahedron with the extra water inside the cage and the proton associated with a hydronium ion incorporated in the surface of the cluster.[144] Subsequently, Castleman and co-workers proposed that the excess proton is associated with the water molecule located inside the dodecahedron.[145] Of course, from the mass spectra alone one cannot establish which of these interpretations is correct or even whether the cluster actually involves a dodecahedral cage structure. In 2004, the nature of the $H^+(H_2O)_{21}$ cluster was revisited by two different teams of researchers who used vibrational predissociation spectroscopy combined with electronic structure calculations.[136,137] While the vibrational spectra of the smaller pronated clusters showed two or more transitions in the free OH stretch region, the vibrational spectrum of the $n = 21$ cluster displayed a single feature in this spectral region. This result was interpreted in terms of a structure with all free OH groups being associated with AAD water molecules as would be the case if the magic number cluster has the structure proposed by Fenn and co-workers. Following the 2004 papers of Shin et al.[136] and Miyazaki et al.,[137] there were several experimental and theoretical studies that provided additional support for the Searcy-Fenn structure for the magic number cluster.[146–151] Of course, there is a whole family of isomers with the Searcy-Fenn structure, and, not surprisingly, electronic structure calculations of Xantheas and co-workers[148] revealed a

member of this family slightly lower in energy than the isomer identified in ref. 136.

In spite of the large number of studies supporting the Searcy-Fenn structure for the magic number $H^+(H_2O)_{21}$ cluster, until recently there remained a major puzzle. Namely, while calculated harmonic spectra predicted the most intense features in the IR spectra to correspond to the OH stretch of the embedded H_3O^+ near 2450 cm^{-1}, the experimental spectra did not show any structure in this spectral region.[136,137] This led some researchers to conclude that the excess proton is so mobile on the surface of the cluster that the vibrational structure is washed out.[146] However, in 2011, Torrent-Sucurrat and Anglanda reported VPT2 level anharmonic vibrational frequencies for the low-energy structure reported in the 2004 study of Shin *et al.* and found that the OH stretch vibrations of the H_3O^+ entity of $H^+(H_2O)_{21}$ are strongly red-shifted upon the inclusion of vibrational anharmonicity.[152] Subsequently, the Johnson group revisited this problem, obtaining the vibrational predissociation spectrum of cryogenically cooled and H_2-tagged $H^+(H_2O)_{21}$ over the 600–4800 cm^{-1} range.[153] This new experimental study identified strong features near 1900 cm^{-1} due to OH stretch of the embedded H_3O^+ entity in good agreement with the predictions of the VPT2 calculations.[152]

In the magic number $n = 21$ cluster, the hydronium ion is engaged in three H-bonds. Hence, it is interesting to examine the evolution of the hydronium OH stretch frequencies as the ion is increasingly hydrated. The first hydration shell is completed in the H_9O_4 Eigen ion, and the second in the $H_3O^+(H_2O)_9$ ion. In both harmonic and VPT2 calculations, these vibrations undergo strong red shifts upon addition of the first hydration shell (to form $H_9O_4^+$) and sizable further red shifts upon addition of the second hydration shell to give $H_3O^+(H_2O)_9$.[140] The calculations of ref. 140 show that the red shifts of the hydronium OH stretch vibration in the magic number $H_3O^+(H_2O)_{20}$ cluster are close to those in $H_3O^+(H_2O)_9$, indicating that the spectral shifts due to adding water molecules outside the second hydration shell are relatively small. The red shifts of the hydronium OH stretch vibrations caused by hydration of the ion are greater in the anharmonic than in the harmonic calculations. The large red shifts caused by hydration can be understood in terms of the electric fields imposed on the H_3O^+ by the solvating water molecules.[140]

5 Conclusions

Water clusters, both neutral and charged, with an excess proton or electron, have complex potential energy surfaces. Over the past two decades, spectroscopic studies employing supersonic jet expansion have provided a wealth of experimental data on $(H_2O)_n$, $H^+(H_2O)_n$, and $(H_2O)_n^-$ clusters. This, in turn, has motivated a large body of theoretical studies of these species. These studies have provided new insights into the role of many-body polarization in water clusters and how excess charge (protons or electrons) impact the H-bonding networks of the clusters. This work has also revealed the need for new theoretical tools.

For example, anharmonicity effects are so large in some $H^+(H_2O)_n$ isomers that computational methods such as VPT2 for calculating anharmonic spectra are inadequate. $(H_2O)_n^-$ clusters also present several problems for theory. Namely, due to the need to include correlation effects (at at least the MP2 level) and to employ large, flexible basis sets, calculation of accurate, anharmonic vibrational spectra using *ab initio* methods is computationally prohibitive other than for relatively small clusters. Although model Hamiltonian approaches have been developed for $(H_2O)_n^-$ clusters, these need to be further refined to predict accurate frequencies and IR intensities. Also, although several studies have appeared addressing isomerization pathways of $(H_2O)_n$ clusters, much less is known about isomerization pathways for the charged clusters. One of the grand challenges in the field is to predict the distribution of isomers that are present under supersonic jet conditions. This is especially challenging because of the time scales involved. Thus, it is clear that, in spite of the large body of theoretical work on these clusters, much remains to be done.

Acknowledgements

References

1 D. J. Wales, *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*, Cambridge University Press, Cambridge, 2003.
2 H. Arslan and M. H. Güven, *New J. Phys.*, 2005, **7**(60), 1–22.
3 D. J. Wales and I. Ohmine, *J. Chem. Phys.*, 1993, **98**, 7257–7268.
4 W. L. Jorgenson, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
5 C. Møller and M. A. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.
6 D. J. Wales, *Mol. Phys.*, 2002, **100**, 3285–3305.
7 C. J. Tsai and K. D. Jordan, *J. Chem. Phys.*, 1991, **95**, 3850–3853.
8 C. J. Tsai and K. D. Jordan, *J. Phys. Chem.*, 1993, **97**, 5208–5210.
9 U. Buck, I. Ettischer, M. Melzer, V. Buch and J. Sadlej, *Phys. Rev. Lett.*, 1998, **80**, 2578–2581.
10 O. M. Becker and M. Karplus, *J. Chem. Phys.*, 1997, **106**, 1495–1517.
11 D. J. Wales, M. A. Miller and T. R. Walsh, *Nature*, 2002, **100**, 3285–3305.
12 R. H. Swendsen and J. S. Wang, *Phys. Rev. Lett.*, 1986, **57**, 2607–2609.
13 A. N. Tharrington, Ph.D. Thesis, University of Pittsburgh, 2001.
14 A. N. Tharrington and K. D. Jordan, *J. Phys. Chem. A*, 2003, **107**, 7380–7389.
15 D. J. Wales and I. Ohmine, *J. Chem. Phys.*, 1993, **98**, 7245–7256.
16 J. M. Pedulla and K. D. Jordan, *Chem. Phys.*, 1998, **239**, 593–601.
17 C. J. Tsai and K. D. Jordan, *J. Chem. Phys.*, 1993, **99**, 6957–6970.
18 T. Kaneko, T. Akimoto, K. Yasuoka, A. Mitsutake and X. C. Zeng, *J. Chem. Theory Comput.*, 2011, **7**, 3083–3087.
19 D. Laria, J. Rodriguez, C. Dellago and D. Chandler, *J. Phys. Chem. A*, 2001, **105**, 2646–2651.
20 F. H. Stillinger and T. A. Weber, *J. Phys. Chem.*, 1983, **87**, 2833–2840.

21 P. A. Frantsuzov and V. A. Mandelshtam, *J. Chem. Phys.*, 2008, **128**, 1–7, 094304.

22 E. Asare, A.-R. Musah, E. Curotto, D. L. Freeman and J. D. Doll, *J. Chem. Phys.*, 2009, **131**, 1–11, 184508.

23 S. Shin, W.-J. Son and S. Jang, *J. Mol. Struct.: THEOCHEM*, 2004, **673**, 109–113.

24 P. E. Videla, P. J. Rossky and D. Laria, *J. Chem. Phys.*, 2013, **139**, 1–9, 174315.

25 T. R. Dyke, K. M. Mack and J. S. Muenter, *J. Chem. Phys.*, 1977, **66**, 498–510.

26 L. A. Curtiss, D. J. Frurip and M. Blander, *J. Chem. Phys.*, 1979, **71**, 2703–2711.

27 M. W. Feyereisen, D. Feller and D. A. Dixon, *J. Phys. Chem.*, 1996, **100**, 2993–2997.

28 S. S. Xantheas, C. J. Burnham and R. J. Harrison, *J. Chem. Phys.*, 2002, **116**, 1493–1499.

29 T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall and C. D. Sherrill, *J. Chem. Phys.*, 2010, **132**, 1–5, 144104.

30 E. M. Mas and K. Szalewicz, *J. Chem. Phys.*, 1996, **104**, 7606–7614.

31 K. Raghaachari, G. W. Trucks, J. A. Pople and M. A. Head-Gordon, *Chem. Phys. Lett.*, 1989, **157**, 479–483.

32 B. E. Rocher-Casterline, L. C. Ch'ng, A. K. Mollner and H. Reisler, *J. Chem. Phys.*, 2011, **134**, 1–4, 21101.

33 B. Temelso, K. A. Archer and G. C. Shields, *J. Phys. Chem. A*, 2011, **115**, 12034–12046.

34 A. Shank, Y. Wang, A. Kaledin, B. J. Braams and J. M. Bowman, *J. Chem. Phys.*, 2009, **130**, 1–11, 144314.

35 B. J. Smith, D. L. Swanton, J. A. Pople, H. F. Schaefer III and L. Radom, *J. Chem. Phys.*, 1990, **92**, 1240–1247.

36 M. J. Gillan, D. Alfè, A. P. Bartók and G. Csányi, *J. Chem. Phys.*, 2013, **139**, 1–9, 244504.

37 M. J. Gillan, D. Alfè and A. Michaelides, *J. Chem. Phys.*, 2006, **144**, 1–33, 130901.

38 C. J. Tsai and K. D. Jordan, *Chem. Phys. Lett.*, 1993, **213**, 181–188.

39 J. Kim and K. S. Kim, *J. Chem. Phys.*, 1998, **109**, 5886–5895.

40 D. M. Bates and G. S. Tschumper, *J. Phys. Chem. A*, 2009, **113**, 3555–3559.

41 C. Pérez, M. T. Muckle, D. P. Zaleski, N. A. Seifert, B. Temelso, G. C. Shields, Z. Kisiel and B. H. Pate, *Science*, 2012, **336**, 897–901.

42 Y. Wang, V. Babin, J. M. Bowman and F. Paesani, *J. Am. Chem. Soc.*, 2012, **134**, 11116–11119.

43 K. Liu, M. Brown, C. Carter, R. Saykally, J. Gregory and D. Clary, *Nature*, 1996, **381**, 501–503.

44 J. J. Foley IV and D. A. Mazziotti, *J. Phys. Chem. A*, 2013, **117**, 6712–6716.

45 D. A. Mazziotti, *Phys. Rev. Lett.*, 2008, **101**, 1–4, 253002.

46 M. J. Gillan, F. R. Manby, M. D. Towler and D. Alfè, *J. Chem. Phys.*, 2012, **136**, 1–14, 244105.

47 C. Steinbach, P. Andersson, M. Melzer, J. K. Kazimirski, U. Buck and V. Buch, *Phys. Chem. Chem. Phys.*, 2004, **6**, 3320–3324.

48 S. S. Xantheas, *J. Chem. Phys.*, 1994, **100**, 7523–7534.

49 R. A. Christie and K. D. Jordan, in *Structure and Bonding: Intermolecular Forces and Clusters*, ed. D. Wales, Springer, 2005, vol. 116, pp. 27–42.

50 F.-F. Wang, G. Jenness, W. A. Al-Saidi and K. D. Jordan, *J. Chem. Phys.*, 2010, **132**, 1–8, 134303.

51 R. Kumar, F. F. Wang, G. Jenness and K. D. Jordan, *J. Chem. Phys.*, 2010, **132**, 1–12, 014309.

52 U. Góra, W. Cencek, R. Podeszwa, A. van der Avoird and K. Szalewicz, *J. Chem. Phys.*, 2014, **140**, 1–20, 194101.

53 U. Góra, R. Podeszwa, W. Cencek and K. Szalewicz, *J. Chem. Phys.*, 2011, **135**, 1–19, 224102.

54 Y. Yang, X. Huang, B. C. Shepler, B. J. Braams and J. M. Bowman, *J. Chem. Phys.*, 2011, **134**, 1–12, 094509.

55 V. Babin, G. Medders and F. Paesani, *J. Chem. Theory, Comput.*, 2014, **10**, 1599–1607.

56 G. S. Fanourgakis and S. S. Xantheas, *J. Chem. Phys.*, 2008, **128**, 1–11, 074506.

57 C. J. Burnham, D. J. Anick, P. K. Mankoo and G. F. Reiter, *J. Chem. Phys.*, 2008, **128**, 1–20, 154519.

58 M. J. Gillan, D. Alfè, P. J. Bygrave, C. R. Taylor and F. R. Manby, *J. Chem. Phys.*, 2013, **139**, 1–11, 114101.

59 H. W. Qi, H. R. Leverentz and D. G. Truhlar, *J. Phys. Chem. A*, 2013, **117**, 4486–4499.

60 J. Brudermann, U. Buck and V. Buch, *J. Phys. Chem. A*, 2002, **106**, 453–457.

61 R. N. Pribble and T. S. Zwier, *Science*, 1994, **265**, 75–79.

62 V. Barone, J. Bloino, C. L. Guido and F. Lipparini, *Chem. Phys. Lett.*, 2010, **496**, 157–161.

63 K. Diri, Y. M. Myshaken and K. D. Jordan, *J. Phys. Chem. A*, 2005, **109**, 4005–4009.

64 J. C. Howard and G. S. Tschumper, *J. Chem. Theory Comput.*, 2015, **11**, 2126–2136.

65 Q. Yu and J. M. Bowman, *Mol. Phys.*, 2015, **113**, 1–8.

66 D. P. Taylor, R. Kusaka, P. S. Walsh, T. S. Zwier and E. L. Sibert, *J. Phys. Chem.*, 2015, **119**, 9917–9930.

67 X. Cheng and R. P. Steele, *J. Chem. Phys.*, 2014, **141**, 1–16, 104105.

68 T. James, D. J. Wales and J. Hernandez-Rojas, *Chem. Phys. Lett.*, 2005, **415**, 302–307.

69 L. X. Dang and T.-M. Chang, *J. Chem. Phys.*, 1997, **106**, 8149–8159.

70 S. Fanourgakis and S. S. Xantheas, *J. Chem. Phys.*, 2008, **128**, 1–11, 074506.

71 P. Ren and J. Ponder, *J. Phys. Chem. B*, 2003, **107**, 5933–5947.

72 R. Bukowski, K. Szalewicz, G. C. Groenenboom and A. van der Avoird, *Science*, 2007, **315**, 1249–1252.

73 T. H. Dunning Jr., *J. Chem. Phys.*, 1989, **90**, 1007–1023.

74 R. A. Kendall, T. H. Dunning Jr. and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.

75 A. Defusco, D. P. Schofield and K. D. Jordan, *Mol. Phys.*, 2007, **105**, 2681–2696.

76 A. Stone, *The Theory of Intermolecular Forces*, Oxford University Press, Oxford, UK, 2nd edn, 2013.

77 N. Gresh, G. A. Cisneros, T. A. Darden and J.-P. Piquemal, *J. Chem. Theory Comput.*, 2007, **3**, 1960–1986.

78 M. P. Hodges, A. J. Stone and S. S. Xantheas, *J. Phys. Chem. A*, 1997, **101**, 9163–9168.

79 G. A. Cisneros, J.-P. Piquemal and T. A. Darden, *J. Chem. Phys.*, 2006, **125**, 1–16, 184101.

80 J.-P. Piquemal, G. A. Cisneros, P. Reinhardt, N. Gresh and T. A. Darden, *J. Chem. Phys.*, 2006, **124**, 1–12, 104101.

81 R. J. Wheatley and S. L. Price, *Mol. Phys.*, 1990, **69**, 507–533.

82 P. E. Videla, P. J. Rossky and D. Laria, *J. Chem. Phys.*, 2016, **144**, 1–5, 061101.

83 F. N. Keutsch, J. D. Cruzan and R. J. Saykally, *Chem. Rev.*, 2003, **103**, 2533–2578.

84 J. O. Richardson, S. C. Althorpe and D. J. Wales, *J. chem. Phys.*, 2011, **135**, 1–12, 124109.

85 A. Mukhopadhyay, W. T. S. Cole and R. J. Saykally, *Chem. Phys Lett.*, 2015, **633**, 13–26.

86 C. Leforestier, K. Szalewicz and A. J. van Avoird, *J. Chem. Phys.*, 2012, **137**, 1–17, 014305.

87 J. O. Richardson, C. Perez, S. Lobsiger, A. A. Reid, B. Temelso, G. C. Shields, Z. Kisiel, D. J. Wales, B. Pate and S. C. Althorpe, *Science*, 2016, **351**, 1310–1313.

88 F. Wang and K. D. Jordan, *Annu. Rev. Phys. Chem.*, 2003, **54**, 367–396.

89 K. D. Jordan, *Acct. Chem. Res.*, 1979, **12**, 36–42.

90 J. H. Hendricks, H. L. deClercy, S. A. Lyapustina, C. A. Fancher, T. P. Lippa, J. M. Collins, S. T. Arnold, G. H. Lee and K. H. Bowen, *Structure and Dynamics of Clusters*, Universal Academy Press, 1996, pp. 321-328.

91 M. Gutowski, P. Skurski, A. I. Boldyrev, J. Simons and K. D. Jordan, *Phys. Rev. A*, 1996, **54**, 1906–1909.

92 H. Haberland, H. Langosch, H.-G. Schindler and D. R. Worsnop, *J. Phys. Chem.*, 1984, **88**, 3903–3904.

93 H. Haberland and K. H. Bowen, in *Clusters of Atoms and Molecules II. Solvation and Chemistry of Free Clusters and Embedded, Supported and Compressed Clusters*, ed. H. Haberland, Springer, Berlin, 1994, pp. 134.

94 L. A. Posey and M. A. Johnson, *J. Chem. Phys.*, 1988, **89**, 4807–4814.

95 J. Kim, I. Becker, O. Cheshnovsky and M. A. Johnson, *Chem. Phys. Lett.*, 1998, **297**, 90–96.

96 E. G. Diken, N. I. Hammer and M. A. Johnson, *Chem. Phys. Lett.*, 2004, **399**, 349–353.

97 T. Sommerfeld, A. DeFusco and K. D. Jordan, *J. Phys. Chem. A*, 2008, **112**, 11021–11035.

98 O. Marsalek, F. Uhlig and P. Jungwirth, *J. Phys. Chem. C*, 2010, **114**, 20489–20495.

99 O. Marsalek, F. Uhlig, J. VandeVondele and P. Jungwirth, *Acct. Chem. Res.*, 2012, **45**, 23–32.

100 L. Turi, *J. Chem. Theory. Comput.*, 2015, **11**, 1745–1755.

101 L. Kevan, *Acct. Chem. Res.*, 1981, **14**, 138–145.

102 L. Turi and P. J. Rossky, *Chem. Rev.*, 2012, **112**, 5641–5674.

103 R. N. Barnett, U. Landman, C. L. Cleveland and J. Jortner, *J. Chem. Phys.*, 1988, **88**, 4429–4447.

104 L. Turi, W.-S. Sheu and P. J. Rossky, *Science*, 2005, **309**, 914–917.

105 J. R. R. Verlet, A. E. Bragg, A. Kammrath, O. Cheshnovsky and D. M. Neumark, *Science*, 2005, **307**, 93–96.

106 L. Ma, K. Majer, F. Chirot and B. von Issendorff, *J. Chem. Phys.*, 2009, **131**, 144303.

107 J. Coe, *J. Chem. Phys.*, 2006, **125**, 014315.

108 L. Turi and D. Borgis, *J. Chem. Phys.*, 2002, **117**, 6186–6195.

109 P. Stampfli, *J. Chem. Phys.*, 1994, **101**, 6024–6031.

110 J. Schnitker and P. J. Rossky, *J. Chem. Phys.*, 1987, **86**, 3462–3470.

111 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren and J. Hermans, *Intermolecular Forces*, ed. B. Pullman, Reidel, Dordrecht, 1981, p. 331.

112 V. P. Vysotskiy, L. S. Cederbaum, T. Sommerfeld, V. K. Voora and K. D. Jordan, *J. Chem. Theory Comput.*, 2012, **8**, 893–900.

113 J. M. Herbert and M. Head-Gordon, *PNAS*, 2006, **103**, 4282–14287.

114 F. Wang and K. D. Jordan, *J. Chem. Phys.*, 2003, **119**, 11645–11652.

115   T.-H. Choi, T. Sommerfeld, S. L. Yilmaz and K. D. Jordan, *J. Chem. Theory. Comput.*, 2010, **6**, 2388–2394.

116   V. Voora, J. Ding, T. Sommerfeld and K. D. Jordan, *J. Phys. Chem. B*, 2013, **117**, 4365–4370.

117   L. D. Jacobson and J. M. Herbert, *J. Chem. Phys.*, 2010, **133**, 1–19, 154506.

118   R. E. Larsen, W. J. Glover and B. J. Schwartz, *Science*, 2010, **329**, 65–69.

119   L. D. Jacobson and J. M. Herbert, *Science*, 2011, **331**, 1387-d.

120   L. Turi and A. Madrász, *Science*, 2011, **331**, 1387-c.

121   J. Stanton and J. Gauss, *J. Chem. Phys.*, 1995, **103**, 1064–1076.

122   M. Nooijen and R. J. Bartlett, *J. Chem. Phys.*, 1995, **102**, 3629–3647.

123   J. Schirmer, L. S. Cederbaum and O. Walter, *Phys. Rev. A*, 1983, **28**, 1237–1259.

124   N. I. Hammer, J. R. Roscioli and M. A. Johnson, *J. Phys. Chem. A*, 2005, **109**, 7896–7901.

125   B. M. Elliott, L. R. McCunn and M. A. Johnson, *Chem. Phys. Lett.*, 2008, **467**, 32–36.

126   T.-H. Choi and K. D. Jordan, *Chem. Phys. Lett.*, 2009, **475**, 293–297.

127   E. G. Diken, W. H. Robertson and M. A. Johnson, *J. Phys. Chem. A*, 2004, **108**, 64–68.

128   J. R. Roscioli, N. I. Hammer and M. A. Johnson, *J. Chem. Phys. A*, 2006, **110**, 7517–7575.

129   T. L. Guasco, B. M. Elliott, M. A. Johnson, J. Ding and K. D. Jordan, *J. Phys. Chem. Lett.*, 2010, **1**, 2396–2401.

130   J. Breen, A. F. DeBlase, T. L. Guasco, V. K. Voora, K. D. Jordan, T. Nagata and M. A. Johnson, *J. Phys. Chem. A*, 2012, **116**, 903–9012.

131   M. J. G. Peach, A. M. Teal, T. Helgaker and D. J. Tozer, *J. Chem. Theory Comput.*, 2015, **11**, 5262–5268.

132   S. L. Lias, J. F. Liebman and R. D. Levin, *J. Phys. Chem. Ref. Data*, 1984, **13**, 695–808.

133   G. Zundel, in The Hydrogen Bond - *Recent Developments in Theory and Experiments.II. Structure and Spectroscopy*, ed. P. Schuster, G. Zundel and C. Sandorfy, North – Holland: Amsterdam, 1976, p. 683.

134   M. Eigen, *Angew. Chem., Int. Ed.*, 1964, **3**, 1–19.

135   J. Headrick, E. G. Diken, R. S. Walters, N. I. Hammer, R. A. Christie, J. Cui, E. M. Myshakin, M. A. Duncan, M. A. Johnson and K. D. Jordan, *Science*, 2005, **308**, 1765–1769.

136   J.-W. Shin, N. I. Hammer, E. G. Diken, M. A. Johnson, R. S. Walters, T. D. Jaeger, M. A. Duncan, R. A. Christie and K. D. Jordan, *Science*, 2004, **304**, 1137–1140.

137   M. Miyazaki, A. Fujii, T. Ebata and N. Mikami, *Science*, 2004, **304**, 1134–1137.

138   C.-C. Wu, C.-K. Lin, H.-C. Chang, J.-C. Jiang, J.-L. Kuo and M. L. Klein, *J. Chem. Phys.*, 2005, **122**, 1–9, 074315.

139   G. Douberly, R. Walters, J. Cui, K. D. Jordan and M. Duncan, *J. Phys. Chem. A*, 2010, **114**, 4570–4579.

140   J. A. Fournier, C. T. Wolke, M. A. Johnson, T. T. Odbadrakh, K. D. Jordan, S. M. Kathmann and S. S. Xantheas, *J. Phys. Chem. A*, 2015, **119**, 9425–9440.

141   R. H. McKenzie, *Chem. Phys. Lett.*, 2012, **535**, 196–200.

142   A. Warshel and R. M. Weiss, *J. Am. Chem. Soc.*, 1980, **102**, 6218–6226.

143   U. W. Schmitt and G. A. Voth, *J. Phys. Chem. B*, 1998, **102**, 5547–5551.

144   J. Q. Searcy and J. B. Fenn, *J. Chem. Phys.*, 1974, **61**, 5282–5288.

145   X. Sei, Z. Shi and A. W. Castleman, Jr., *J. Chem. Phys.*, 1991, **94**, 3268–3270.

146 S. S. Iyengar, T. J. F. Day and G. A. Voth, *Int. J. Mass. Spectr*, 2005, **241**, 197–204.
147 F. Calvo, J. Douady and F. Spiegelman, *J. Chem. Phys.*, 2010, **132**, 1–12, 024305.
148 S. S. Xantheas, *Can. J. Chem. Eng.*, 2012, **90**, 843–851.
149 A. Khan, *Chem. Phys. Lett.*, 2000, **319**, 440–450.
150 M. P. Hodges and D. J. Wales, *Chem. Phys. Lett.*, 2000, **324**, 279–285.
151 K. Mizase and A. Fujii, *J. Phys. Chem. Lett.*, 2011, **2**, 2130–2134.
152 M. Torrent-Sucarrat and J. M. Anglada, *J. Chem. Theory Comput.*, 2011, **7**, 467–472.
153 J. A. Fournier, C. I. Johnson, C. T. Wolke, G. A. Weddle, A. B. Wolk and M. A. Johnson, *Science*, 2014, **344**, 1009–1012.

The incremental method – theory and applications in chemistry and physics[†]

Benjamin Fiedler and Joachim Friedrich*

1 Introduction

The aim of *ab initio* quantum chemistry is the description of chemical issues by computational methods, based on quantum mechanics. In many cases it is sufficient to solve the non-relativistic, time-independent Schrödinger equation, using the Born–Oppenheimer approximation. Within this framework single-point energies on the potential energy surface (PES) are obtained from the solution of the electronic Schrödinger equation[1]

$$\underbrace{\left[\hat{T}_{el} + \hat{V}_{el,nuc} + \hat{V}_{el,el}\right]}_{\hat{H}_{el}}|\psi_{el}\rangle = E_{el}|\psi_{el}\rangle \tag{1}$$

The electronic Hamiltonian $\hat{H}_{el}$, which acts on the electronic wave function $|\psi_{el}\rangle$, contains the operators for the electronic kinetic energy ($\hat{T}_{el}$) as well as for the electron-nuclei ($\hat{V}_{el,nuc}$) and electron–electron potentials ($\hat{V}_{el,el}$). There are various methods to solve this equation, *i.e.* to determine the electronic energy E_{el}, which are either based on Density Functional Theory (DFT) or on Wave Function Theory (WFT). Using the electron density in DFT methods provides a low-order scaling with the system size, equivalent to a high computational efficiency, and is therefore the method of choice for modelling reactions in quantum chemistry.[2] However, there is a large variety of density functionals, which feature no strict quality hierarchy, *i.e.* estimating the quality for a specific chemical topic is critical.[2] In this context, wave-function-based methods are advantageous, since they provide such a stringent hierarchy.

The most fundamental of these methods is the Hartree–Fock (HF) theory,[1,3] which uses only one Slater determinant, consisting of one-electron functions (orbitals), as wave function ansatz. Within this formalism, the many-body problem reduces to a set of one-electron equations. They contain the electron–electron repulsion only by means of an average Coulomb potential of the remaining electrons, interacting with one considered electron. Therefore, HF is not capable to describe the correlated motion of two electrons sufficiently.

Thus, post-HF correlation methods, like coupled cluster (CC) or Møller–Plesset (MPn) perturbation theory, are mandatory to solve eqn (1)

Institute for Chemistry, Technische Universität Chemnitz, Straße der Nationen 62, D-09111 Chemnitz, Germany. E-mail: joachim.friedrich@chemie.tu-chemnitz.de
[†]Dedicated to Professor Hermann Stoll.

in a more exact way. They are generally based on the Hartree–Fock orbitals, but the wave function includes more Slater determinants, that represent the substitution of occupied orbitals by virtual ones.[1,3]

Among these methods, CCSD(T), *i.e.* coupled cluster with singles, doubles and perturbative triples, is known as the "gold standard" of quantum chemistry, since it provides a very high accuracy.[4] However, CCSD(T) unfortunately exhibits a high scaling (O^3V^4) with respect to the number of occupied (O) and virtual (V) orbitals[5] and is therefore only feasible for rather small chemical systems.

Therefore, a lot of local correlation methods with a reduced scaling have been developed over the last three decades. Electron correlation is a short-range interaction and its magnitude decreases rapidly with increasing distances. Hence, the usage of localized orbitals is a suitable basis. That way, the locality of the electron correlation can be exploited to neglect the correlation of several orbital pairs.

However, these methods are predicated on different fundamental ideas. At first, the locality of the electron correlation can be exploited by specific construction of the occupied and virtual orbital spaces. Based on the essential work of Pulay and Saebø about projected atomic orbitals (PAOs),[6–10] Werner, Schütz and co-workers developed several local correlation methods, like LMP2, LCCSD or LCCSD(T).[11–18] Neese *et al.* implemented the highly efficient local pair natural orbital (LPNO) approach[19–24] and further improved it by combination with the concepts of PAOs, giving the domain-based LPNO (DLPNO) method.[25–27] Furthermore, Yang *et al.* proposed orbital-specific virtuals (OSVs).[28,29]

A second group is based on the idea of splitting the whole system into small fragments. Thereby, correlation calculations are performed for those fragments and the results are recombined to get a total value. Some exemplary methods are the cluster-in-molecule (CIM) approach of Li,[30–36] natural linear-scaling coupled cluster[37–39] by Bartlett, the divide-and-conquer scheme,[40–42] fragment molecular orbital (FMO) approaches,[43–45] the local CCSD method of Head-Gordon,[46–48] the divide-expand-consolidate (DEC) scheme by Jørgensen[49–51] as well as the electrostatically-embedded many-body (EEMB) expansion of Truhlar.[52]

Incremental methods, which are focused in this review, may be mainly classified into the second category, but contain also aspects of specific orbital space construction, *e.g.* PAOs have been used to truncate the virtual space.[53] In the next two sections we will present some fundamental and theoretical aspects of incremental methods and take into account the various implementations of different groups. Afterwards we will give an overview of applications to molecular as well as periodic systems.

2 General formalism

2.1 Brief historical overview

First, we give a short overview on the origins of the incremental methods. The concept of increments has been used in chemistry for a long time and in various ways to compute different properties. In 1992, Stoll was

the first to combine this concept with the framework of local correlation methods,[54] which is the "birth" of the incremental methods, presented in this review. However, the fundamental equations for the incremental ("many-body") expansion root back to Nesbet's work[55–59] on the Bethe–Goldstone equations. The latter have originally been developed by Bethe and Goldstone, who modified the Schrödinger equation for a system of two nucleons,[60] based on the work of Brueckner.[61–64] Nesbet generalized these equations to an N-fermion system with n particles embedded in a Fermi sea of the other $(N - n)$ particles. That way he introduced the so-called atomic Bethe–Goldstone equations to quantum chemistry, calculating correlation energies of second-[55–57] and third-period[59] elements by means of increments. Fundamental for incremental expansions were also the Faddeev equations, which have been used in scattering theory, including all interactions in three-[65] and several particle systems.[66]

Originally, Stoll applied the incremental method to correlation energy calculations for periodic systems, like diamond,[54] silicon[67] or graphite.[68] Since then a lot of effort has been spent to advance incremental calculations, especially by Stoll, Paulus, Fulde, Dolg, Doll and many others. Thus, calculations for various systems, including bulk semiconductors,[54,67–74] insulators,[75–79] metals,[80–83] open-shell[84,85] and multi-reference systems[86–88] as well as adsorbing surfaces[89–99] and atomic respectively molecular crystals,[100–105] were performed. Furthermore, starting in 2007, the applicability of the incremental scheme has been successfully extended to finite molecules by Dolg, Friedrich and Hanrath.[53] Beside closed-shell molecules and clusters, this framework was also implemented for open-shell systems[106–108] and the computation of molecular properties.[109–112]

2.2 Incremental expansions

In this section, we want to have a look at the general formalism, that characterizes all incremental methods, including the equations for incremental expansions.

Post-HF methods for the calculation of correlation energies are based on the HF solution of the Schrödinger equation, *i.e.* on SCF orbitals. In incremental methods these orbitals need to be localized first. Next a partitioning of the set of localized orbitals is performed, giving small disjoint subsets, which we call (one-body) domains. We will have a more detailed look on suitable localization and partitioning schemes in Section 2.3.

In most cases the incremental methods have been applied to weakly-correlated systems, *i.e.* those with single-reference character, and therefore only the occupied SCF orbitals are localized. However, multi-reference cases have also been treated in context of periodic systems, especially metals. More information is provided in Section 4.3.

In a first-order approximation the total correlation energy may be calculated as the sum of correlation energies (ε_i, ε_j, ε_k,...) of the disjoint (one-body) domains, identified by $i, j, k,...$.

$$E_{(1)}^{\text{corr}} = \sum_i \varepsilon_i \tag{2}$$

To get such a correlation energy, the occupied orbitals, not belonging to a specific domain, are kept frozen during the post-HF calculation and therefore only the active orbitals of this domain are correlated. That way the "intradomain" correlation is described, which would be sufficient for non-interacting domains. But since "interdomain" interactions occur, one needs to take into account also non-additive correlation contributions. Therefore, for a second-order approximation, two-body domains (ij, ik, jk,...) have to be formed as set unions of the LMOs in the corresponding one-body domains and their correlation energies (ε_{ij}, ε_{ik}, ε_{jk},...) must be calculated. These energies contain not only the non-additive contributions, the so-called two-body increments ($\Delta\varepsilon_{ij}$, $\Delta\varepsilon_{ik}$, $\Delta\varepsilon_{jk}$,...), but also the one-body correlation energies (with $\Delta\varepsilon_i = \varepsilon_i$):

$$\varepsilon_{ij} = \Delta\varepsilon_i + \Delta\varepsilon_j + \Delta\varepsilon_{ij} \Leftrightarrow \Delta\varepsilon_{ij} = \varepsilon_{ij} - (\Delta\varepsilon_i + \Delta\varepsilon_j) \tag{3}$$

The simple summation of all two-body energies is not proper, due to the multiple counting of the one-body energies. Due to index symmetry we have $\Delta\varepsilon_{ij} = \Delta\varepsilon_{ji}$ and may either take into account all combinations and prevent double countings by a prefactor of (1/2) or avoid double countings *a priori* by restricting the second summation index j. Thus, we get

$$E_{(2)}^{corr} = \sum_i \Delta\varepsilon_i + \frac{1}{2}\sum_{i,j} \Delta\varepsilon_{ij} = \sum_i \Delta\varepsilon_i + \sum_{i<j} \Delta\varepsilon_{ij} \tag{4}$$

To get a higher accuracy, also third- or even higher-order contributions should be taken into account. The corresponding increments as well as the incremental expansion, are considered similarly to the two-body increments. The third-order equations read,

$$\Delta\varepsilon_{ijk} = \varepsilon_{ijk} - (\Delta\varepsilon_{ij} + \Delta\varepsilon_{ik} + \Delta\varepsilon_{jk}) - (\Delta\varepsilon_i + \Delta\varepsilon_j + \Delta\varepsilon_k) \tag{5}$$

$$E_{(3)}^{corr} = \sum_i \Delta\varepsilon_i + \frac{1}{2!}\sum_{i,j} \Delta\varepsilon_{ij} + \frac{1}{3!}\sum_{i,j} \Delta\varepsilon_{ijk} = \sum_i \Delta\varepsilon_i + \sum_{i<j} \Delta\varepsilon_{ij} + \sum_{i<j<k} \Delta\varepsilon_{ijk} \tag{6}$$

Next the previous considerations are extended to the O-th order and generalized descriptions of the increments as well as the incremental expansion are derived. Therefore, we define $\mathbb{X}$ as index for the increments $\Delta\varepsilon_{\mathbb{X}}$, representing the set of LMOs in the according N-body domain. First, $\mathbb{X}$ is an element of the power set $\mathscr{P}(\mathbb{D})$, which contains all possible subsets of the set of one-body domains ($\mathbb{D}$). Second, the cardinality $|\mathbb{X}|$, representing the number of one-body domains in the N-body domain, runs up to O, which is the truncation order of the incremental expansion. Thus, we get

$$E_{(O)}^{corr} = \sum_{\mathbb{X} \in \mathscr{P}(\mathbb{D}) \wedge |\mathbb{X}| \leq O} \Delta\varepsilon_{\mathbb{X}} \tag{7}$$

For a specific increment with index $\mathbb{X}$, we subtract all lower-order increments from $\varepsilon_{\mathbb{X}}$ to avoid their multiple counting. Thus, we introduce

$\mathbb{Y}$ as index for all lower-order increments. This set runs over all possible subsets of $\mathbb{X}$ and is therefore an element of the power set $\mathscr{P}(\mathbb{X})$. The exclusion of $\mathbb{X}$ itself occurs due to the restriction of the cardinality $|\mathbb{Y}|$. Therefore, the general increment $\Delta\varepsilon_{\mathbb{X}}$ reads

$$\Delta\varepsilon_{\mathbb{X}} = \varepsilon_{\mathbb{X}} - \sum_{\mathbb{Y}\in\mathscr{P}(\mathbb{X})\wedge|\mathbb{Y}|<|\mathbb{X}|} \Delta\varepsilon_{\mathbb{Y}} \qquad (8)$$

The exact canonical correlation energy is reached, if the incremental expansion runs up to the $|\mathbb{D}|$-th order, where $|\mathbb{D}|$ is the total number of one-body domains. In this case all lower-order increments cancel exactly in the incremental expansion and the correlation energy of the total system remains.

Another condition for the convergence of the incremental series to the correct correlation energy arises from the fact, that $\varepsilon_{12\ldots|\mathbb{D}|}$ is received, using a local occupied orbital space. Since the canonical orbitals are unitary transformed to generate localized ones, either the implementation of the considered energy expression must be invariant with respect to these transformations (e.g. for CCSD) or a set of pseudocanonical orbitals has to be built from the localized ones. The latter is necessary for the canonical formulations of the MP2 energy and the perturbative triples contribution in CCSD(T), which are most common and computationally cheapest, but do not feature the required invariance. For the transformation into a pseudocanonical basis,[113] the Fock matrix in the local basis is needed first. Since only the orbitals in the main region of the domain (d) as well as those in the environment (e), but not the frozen core (c) and virtual (v) orbitals are localized, the rearranged local Fock matrix is

$$\mathbf{F}(\mathbf{C}_L) = \begin{pmatrix} \mathbf{F}_{cc} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{ee} & \mathbf{F}_{ed} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{de} & \mathbf{F}_{dd} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{F}_{vv} \end{pmatrix} \qquad (9)$$

with diagonal blocks $\mathbf{F}_{cc}$ and $\mathbf{F}_{vv}$, but non-diagonal ones $\mathbf{F}_{ee}$ and $\mathbf{F}_{dd}$. As the environment is frozen in the subsequent correlation calculation, only the $\mathbf{F}_{dd}$ block needs to be diagonalized to obtain pseudocanonical orbitals. The matrix $\mathbf{U}_{dd}$, which diagonalizes the $\mathbf{F}_{dd}$ block, is used to build matrix $\mathbf{U}$ (see eqn (10)), which transforms the local MO matrix $\mathbf{C}_L$ into the pseudocanonical one $\tilde{\mathbf{C}}_L$ (see eqn (11)).

$$\mathbf{U} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}_{dd} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix} \qquad (10)$$

$$\tilde{\mathbf{C}}_L = \mathbf{U}^T\mathbf{C}_L \qquad (11)$$

2.3 Localization and domain partitioning

As a first step in all incremental methods, those orbitals, which shall be correlated, need to be localized. In combination with the incremental method different localization algorithms, like the Boys[114] or the Pipek–Mezey[115] procedure, have been used. The Boys localization minimizes the spatial extent of all N_{orb} orbitals (eqn (12)), which is equivalent to the maximization of the sum of pairwise distances between the orbital centers (eqn (13)). Here, integrals of the orbitals (ϕ_i, ϕ_j) over the position operator $\mathbf{r}$ need to be considered.

$$\langle \Omega \rangle_{Boys} = \sum_{i=1}^{N_{orb}} \langle \phi_i \phi_i | (\mathbf{r}_1 - \mathbf{r}_2)^2 | \phi_i \phi_i \rangle \tag{12}$$

$$\langle \Omega' \rangle_{Boys} = \sum_{i>j}^{N_{orb}} (\langle \phi_i | \mathbf{r} | \phi_i \rangle - \langle \phi_j | \mathbf{r} | \phi_j \rangle)^2 \tag{13}$$

However, the Pipek–Mezey algorithm maximizes the sum of partial atomic charges ($|\rho_A|^2$), generally based on Mulliken charges (eqn (14)). They are obtained from the overlap matrix elements of AO functions (α, β) and the LCAO coefficients, generating the MO i.

$$\langle \Omega \rangle_{PM} = \sum_{A=1}^{N_{atoms}} |\rho_A|^2 \quad \text{with} \quad \rho_A = \sum_{i=1}^{N_{orb}} \sum_{\alpha \in A}^{N_{basis}} \sum_{\beta=1}^{N_{basis}} c_{\alpha i} c_{\beta i} S_{\alpha \beta} \tag{14}$$

Please note, that a special localization procedure is necessary for metallic systems. For further information see Section 4.3.

As a second step, the partitioning of the total system into disjoint domains is essential in all incremental methods. It is obvious, that compact domains improve the convergence of the incremental series, since then the predominant part of the short-range correlation energy is caught by low-order increments.

After receiving the orbitals ϕ_i in the local basis, their charge centers $\mathbf{R}_i$ are obtained from the diagonal dipole integrals

$$\mathbf{R}_i = \langle \phi_i | \hat{r} | \phi_i \rangle \tag{15}$$

Next the orbital groups can be generated on different ways. On the one side, it is possible to partition the localized orbitals by hand. A first simple case are clusters and crystals, that are consisting of single atoms, ions or small molecules and kept together by metallic, strongly ionic, van der Waals as well as hydrogen bonding or other weak interactions. Here the partitioning is obvious. Due to the absence of covalent bonding, every localized orbital is unambiguously assigned to one fragment.

If covalent bonding occurs in a periodic or a molecular system, the assignment to an atom is ambiguous and another scheme must be chosen. A frequently-used approach is to partition the localized orbitals, considering the bonds. In this way a single bond generates a domain, consisting of only one orbital, a double bond a domain with two orbitals,

and so on. Furthermore, all lone pairs and correlated core electrons at one atom may also form their own domain.

There are many periodic systems with an intermediate character, *i.e.* with parts of ionic and covalent bonding. Thus, one has to decide between a partitioning by ions or by bonds. However, such a partitioning by hand ensures individuality and flexibility.

On the other side, it is very useful to get the most suitable partitioning, relating to compact domains. Therefore, one should use an algorithmic approach, if possible. During the development of the fully-automated incremental scheme different partitioning schemes have been utilized.

The first one[53] was based on METIS graph partitioning,[116] where the nodes of the graph represent the LMOs with their charge centers. Furthermore, the edges represent the pairwise "connectivity" of two LMOs, which is defined by the matrix **C** with

$$
C_{ij} = \begin{cases} 10^8 & \text{if } D_{ij} \leq t \land \dfrac{10^4}{D_{ij}} \geq 10^8 \\[2mm] \dfrac{10^4}{D_{ij}} & \text{if } D_{ij} \leq t \land \dfrac{10^4}{D_{ij}} > 10^8 \\[2mm] 0 & \text{if } D_{ij} > t \end{cases} \tag{16}
$$

Here $D_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$ is the distance of the two charge centers of the LMOs i and j and t a distance threshold. METIS also requires the number of domains $|\mathbb{D}|$, which is given indirectly by the number of occupied LMOs $|\mathbb{O}|$ and the domain size parameter (dsp), an integer input parameter corresponding to the aspired number of LMOs in a domain.

$$
|\mathbb{D}| = \frac{|\mathbb{O}|}{\text{dsp}} \tag{17}
$$

The METIS algorithm now produces $|\mathbb{D}|$ parts, minimizing the sum of the cut edge weights. The sum of the remaining "connectivities" is therefore maximized, implying domains with LMOs very close to each other.

Later this partitioning scheme has been extended by an additional pairwise refinement.[117] Here two domains $\mathbb{D}_i$ and $\mathbb{D}_j$ are successively selected. Considering their set union $\mathbb{D}_{ij}$, those two orbitals i' and j' with the largest distance mark the beginning of both refined domains $\mathbb{D}'_i$ and $\mathbb{D}'_j$. The remaining LMOs are classified depending on whether they are closer to $\mathbf{R}_{i'}$ or $\mathbf{R}_{j'}$. The maximum number of orbitals in a domain is defined by dsp and the refinement runs until the domains are consistent.

Another automated approach to generate the domains is the so-called K-means clustering,[118,119] implemented by Dolg and co-workers.[120] In this algorithm the function

$$
J = \sum_{k=1}^{|\mathbb{D}|} \frac{\displaystyle\sum_{\mathbf{R}_{k_i} \in \mathbb{D}_k} |\mathbf{R}_{k_i} - \overline{\mathbf{R}}_k|^2}{|\mathbb{D}_k|} \tag{18}
$$

is minimized with respect to the partitioning into $|\mathbb{D}|$ domains with index k. Here the $\mathbf{R}_{k_i}$ indicate all charge centers in the k-th domain and $\bar{\mathbf{R}}_k$ represents their geometrical center. Hence, the algorithm minimizes the sum of all "intradomain" distances. Starting from a random initial guess, one charge center is picked and for test purposes put into all domains. It is then assigned to that domain, for which J is minimal. This step is iteratively done for all charge centers until the partitioning is consistent.

For the same order of expansion, using larger (determined by dsp) respectively less (determined by $|\mathbb{D}|$) domains increases the accuracy.

2.4 Reducing the number of increments

Localized orbitals are useful, since the short-range correlation energy decreases rapidly with increasing orbital respectively domain distances. Therefore, increments of two-, three-body,... domains can be neglected and therefore need not to be calculated, if their distance is large enough (see left side of Fig. 1). There are two different truncation functions, that have been tested. Both depend on the order i, since the increments decrease also significantly with increasing order and thus the minimal distance for screening an increment can be lowered. In the first works, the order-dependent distance threshold $t_{\text{dist}}(i)$ was determined by[121]

$$t_{\text{dist}}(i) = \frac{f}{i} \tag{19}$$

To demonstrate the efficiency of a distance truncation, Fig. 2 shows the dependency of the local errors and computational savings on the f-value, using eqn (19) for a $(H_2O)_{11}$ cluster.[121] Since the local error is affected very little, chemical accuracy is provided, except for a second-order expansion with a very small f (12 Bohr). In contrast, the computational effort (*i.e.* the number of calculations N and the total CPU time) is reduced tremendously.

Later, the following truncation function has been well established and is used within the current automated incremental scheme.[113]

$$t_{\text{dist}}(i) = \frac{f}{(i-1)^2} \tag{20}$$
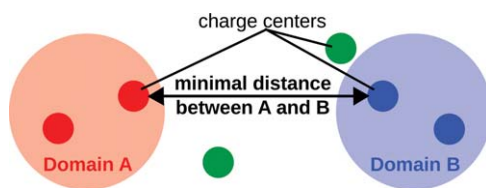
Fig. 1 Schematic representation for the distance truncation. If the minimal distance between two domains, *i.e.* between two charge centers of them, is larger than a given threshold $t_{\text{dist}}(2)$, the two-body increment $\Delta\varepsilon_{AB}$ is negligible. The same applies to higher orders i with $t_{\text{dist}}(i)$.
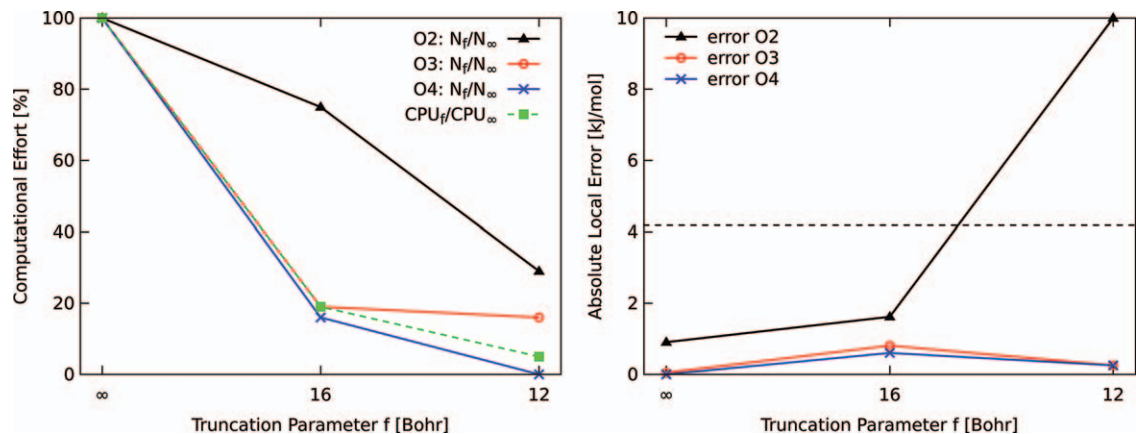
Fig. 2 Dependency of the computational savings (left) relative to the untruncated values (*i.e.* $f = \infty$) and the local errors (right) with respect to the truncation parameter f, using eqn (19). The local errors are shown for expansions up to the n-th order ($n = 2,3,4$). The computational savings are depicted in terms of numbers of calculations for n-th order contributions and as total CPU time for a fourth-order expansion.[121]

Evaluating the parameter f, used in eqn (20), by a statistical analysis for a large test set yielded a recommended value of $f = 30$ Bohr.[117] Depending on the system, too small values of f provoke an apparent error, *e.g.* $f = 15$ Bohr for a $(H_2O)_{13}$ cluster[113] or $f = 20$ Bohr for a test set.[112]

Within the method of increments for periodic systems there is the principle of "nearest neighbours" to limit the necessary calculations by a distance truncation. Thus, a specific number of these neighbours is taken into account, determining the number of two- and three-body increments. Considering *e.g.* a two-body domain, consisting of two atoms A and B, as well as an n-th nearest neighbour approach, the n closest lying atom or ion pairs A–B need to be calculated. An example for this approach is depicted in Fig. 3 in terms of the one-body, nearest-neighbour (NN) and next-nearest-neighbour (NNN) two-body increments for GaAs.[70]

	increment	weight factor
(a)	one-body	4
(b)	NN Ga-As-Ga two-body	6
(b')	NN As-Ga-As two body	6
(c)	NNN two-body, 60° dihedral	24
(d)	NNN two-body, 180° dihedral	12
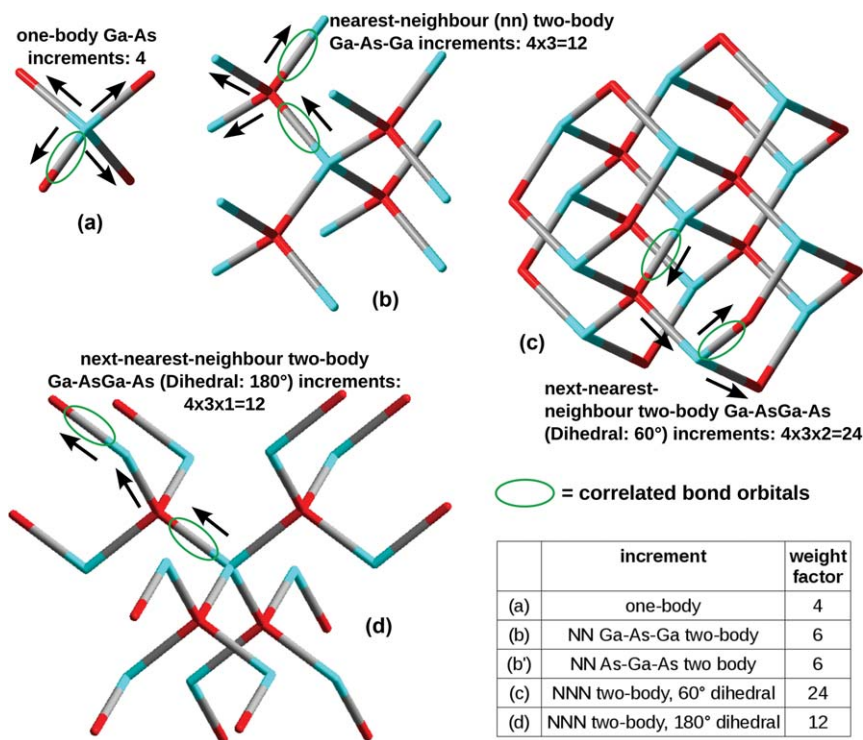
Fig. 3 Presentation of the nearest-neighbour approach and the evaluation of weight factors, exploiting symmetry, for the 5 largest increments of GaAs. In the fragments, the branching points represent the alternating Ga and As atoms, starting with Ga as central one. Furthermore, the black arrows refer to the counting of symmetry-equivalent increments, starting from the central atom, *e.g.* for (b): For the first "counting step", *i.e.* the bonds directly connected to the central Ga, only 1 of the 4 possibilities is marked with an arrow (all 4 are presented for (a)). For the second "counting step" all 3 subsequent As-Ga bonds are tagged and overall there are $4 \times 3 = 12$ counted increments. Additional ellipses mark the correlated bond orbitals. (a),(c),(d): Only the increments, starting with Ga, are counted, as those with As as central atom would provide the same increments, *i.e.* a double counting would occur. (b): To obtain the weight factor (6), the number of counted increments (12) must be divided by 2, as otherwise those increments would be double counted. (b'): The counting of As-Ga-As is not presented, as it is similar to (b), but with changed atoms. However, the As-Ga-As increments are obviously different from the Ga-As-Ga ones in (b) and thus need to be taken into account.

Beside the distance-dependent screening of small increments, there were also efforts to use an energy screening within the incremental scheme.[53,121] This means, a lower level method may be used to compute all increments and afterwards all N-body domains are neglected, for which the lower-level energy is smaller than a given threshold E_{thresh}. It has been shown, that, even for rather small molecules (hydrocarbons, $(H_2O)_{11}$), the number of calculations can be reduced significantly when using an appropriate value of E_{thresh}. Since mainly the small, but time consuming higher-order increments are screened off, the benefit in the computation time should be even higher. However, for molecules with rather nonlocal MOs and compact shape, like naphthalene, only few increments could be screened off.

Furthermore, it is possible to reduce the number of calculations by the use of symmetry. Within the automated incremental scheme, an algorithm for symmetry-adaption and the further treatment of the symmetry-equivalent groups has been implemented. For details see ref. 122, especially the corresponding supplementary material. Since specific increments are symmetry-equivalent within a given symmetry (point or space) group, they must be counted, giving the corresponding weight factor for this increment. Thus, from the set of symmetry-equivalent increments, only one has to be calculated and afterwards multiplied by the weight factor. Especially for periodic systems, the consideration of symmetry is essential. In Fig. 3 every ''cluster'' represents symmetry-equivalent domains of a covalently described GaAs crystal and the corresponding weight factors are considered.

2.5 Getting correlation and total energies

Considering the correlation energy, often pure valence correlation contributions describe the predominant part. Thus, in many cases the frozen core approximation can be applied within the incremental scheme.

However, some work has been spent to calculate also core and core-valence contributions for molecular systems.[123,124] Due to the additional core electrons and the requirement of special larger basis sets, the computational effort may increase significantly. But since the core orbitals are clearly more localized than the valence ones, the correlation between core orbitals of different atoms is negligibly small, offering the screening of many increments. Thus, a different truncation of the incremental expansion for valence and core orbitals is necessary, which excludes mixed core-valence domains. Hence, the valence domains were obtained after the conventional graph partitioning of the localized valence orbitals. A core domain was formed for every atom (except for H). Then, correlation contributions between two core domains are negligible and all N-body domains, containing two or more core domains, are screened off. This special treatment of core- and core-valence correlation effects within the incremental scheme has been implemented and applied for CCSD[123] as well as CCSD(T) and MP2.[124] As for the valence contributions, in most cases a third order incremental expansion was sufficient to reach chemical accuracy for the core/core-valence part. Furthermore, both the cc-pVXZ (X = D,T) and the cc-pCVXZ (respectively

cc-pwCVTZ) basis sets were tested. The latter ones are necessary to yield the core correlation contributions to a sufficient (and in several cases non-negligible) amount. Beside absolute energies, also relative ones were investigated in terms of binding energies for $Na(H_2O)_n^+$ clusters ($n = 2$–4) and the share of core-valence correlation in the total binding energy was 3–13%.[124]

For systems with heavier atoms it is often necessary not only to freeze the core orbitals, but also to replace some of them by effective core potentials (ECPs). ECPs have been used within the molecular incremental scheme *e.g.* for mercury[122] and gold clusters,[123] Au–S interactions[121] or cisplatinum.[123] Furthermore, ECPs have been especially employed for the treatment of periodic systems. Additionally there have been several efforts to include core-valence correlation by the core polarization potential (CPP) method.[71,72] Further Information is provided in Section 4.1.

For the computation of the total energy in most cases only the correlation energy is obtained from an incremental method, whereas the HF part is received from a standard calculation. Considering the quantum chemical descriptions of periodic systems, the incremental calculation of the correlation energy is based on finite-cluster models, often combined with an appropriate embedding. The HF energy was in most cases computed by an infinite model, using periodic boundary conditions.

A common incremental expansion of the HF energy with an appropriate, efficient truncation is critical, since there are usually non-negligible long-range HF contributions. Though, the method of increments has been used to calculate the HF binding energy of rare-gas crystals with small errors of *e.g.* only 0.7% for an Ar crystal.[103] This is due to the long-range effects, which are usually essential within the HF energy, but negligibly small for rare-gas crystals. In contrast, for crystals of polar molecules (*e.g.* CO_2, HCN) an incremental expansion of the HF energy failed.[104] Though, Mata and Stoll applied the method to compute SCF energies of hydrogen-bonded clusters by using special (and in some cases even iteratively optimized) embedding potentials.[125]

3 Incremental methods for molecules

3.1 Virtual space truncation

3.1.1 Projected atomic orbitals.
For an efficient local method it is not only important to correlate small "portions" of occupied orbitals, but also to truncate the virtual orbital space. A first approach, which was used within the automated incremental scheme, is based on projected atomic orbitals (PAOs).[53] Using PAOs has been well established within the local MP methods, developed by Boughton and Pulay[126] as well as Werner, Schütz, *et al.*[11–18] In this approach, a virtual space $\mathbb{V}_{\phi_a}$ for every occupied orbital ϕ_a is spanned by a set of PAOs $\{\phi_i^{PAO}\}$. A PAO is constructed by projecting an AO function ϕ_i^{AO} according to

$$\left|\phi_i^{PAO}\right\rangle = \left|\phi_i^{AO}\right\rangle - \sum_{j=1}^{n_{occ}} \left\langle\phi_j^{MO}\middle|\phi_i^{AO}\right\rangle\left|\phi_j^{MO}\right\rangle \tag{21}$$

with n_{occ} as the number of occupied MOs. Since we use local orbitals, only a subset of AOs is relevant for a given MO ϕ_a^{MO} and thus needs to be projected. An AO ϕ_i^{AO} is important, if the distance R_{ai} of its center to the charge center of ϕ_a^{MO} is rather small. Thus, the AOs are ordered by increasing R_{ai}, i.e. ϕ_1^{AO} is the closest AO. Assuming, there are n_{AO} AOs, the original AO representation of ϕ_a^{MO} is

$$\phi_a^{MO} = \sum_{i=1}^{n_{AO}} c_i \phi_i^{AO} \tag{22}$$

Neglecting some of the AOs, we get an approximate description of the MO with $\bar{n}_{AO}$ AOs remaining:

$$\bar{\phi}_a^{MO} = \sum_{i=1}^{\bar{n}_{AO}} c_i \phi_i^{AO} \tag{23}$$

The number $\bar{n}_{AO}$ of AOs, that are projected, is determined as the minimal value, for which the equation

$$\int \left(\phi_a^{MO} - \bar{\phi}_a^{MO} \right)^2 d\tau \leq t_{dens} \tag{24}$$

is fulfilled for a given density threshold t_{dens}. This means, the difference between the densities of the original and the approximate MOs is smaller than this threshold. Then the virtual space $\mathbb{V}_{\phi_a}$ contains $\bar{n}_{AO}$ PAOs, obtained by projecting the chosen $\bar{n}_{AO}$ AO functions. Furthermore, the virtual space for a domain is the set union of the virtual spaces of all included localized MOs.

With this approach, t_{dens} has been introduced as an additional parameter. If this threshold is fixed to loose (e.g. 0.1), the virtual space is restricted too much. This results in a rather worse convergence behaviour, as shown for hexane[53] and a water cluster.[121] With an appropriate value (e.g. 0.001), the PAO approach works well within the incremental scheme.

3.1.2 Domain-specific basis set approach.

Another method for the truncation of the virtual space, the so-called domain-specific basis set (DSBS) approach, was implemented by Friedrich and Dolg in 2008[127] and has been successfully established since then. Within this scheme a main region is assigned to every domain, which is generated as follows (see Fig. 4). For every localized orbital a sphere around its charge center is determined, using a radius equivalent to a fixed distance threshold t_{main}. Unification of the spheres of all localized orbitals, belonging to a specific one-, two-, three-body domain, yields the main region for this domain. Now all atoms inside this main region are treated with the large target basis set (B) and all outlying atoms, determining the environment, are described by a small basis set (B'). This mixed (or dual) basis set is called domain-specific basis ($B_{\mathbb{X}}$).

Hence, the total number of virtual orbitals as well as the computational cost are reduced in comparison to a treatment of all atoms with the larger
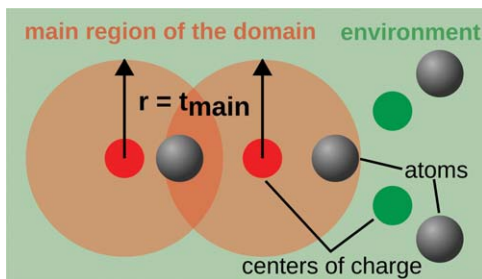
Fig. 4 Schematic representation for the construction of the domain-specific basis set. The centers of charge of the considered domain and the radius t_{main} define the domain's main region. The atoms in the main region are treated with a larger basis set than those in the environment.

basis, but an additional error is introduced. The accuracy of the DSBS approximation within the incremental scheme has been analyzed in detail, also with respect to the value of t_{main}, for which 3 Bohr turned out to be an optimal value (see left side of Fig. 5).[117] However, even for this optimized value a DSBS error remains. To correct for this error, an MP2-based scheme has been established.[128] It has been shown, that the local error due to DSBS approximation is similar for a third-order incremental CCSD/CCSD(T) and MP2 expansion (see right side of Fig. 5). Therefore, this local error can be reduced significantly by estimating it at the MP2 level of theory (eqn (25)). The figure shows also, that the MP2 correction is meaningless, if no DSBS approximation is used (*i.e.* $t_{main} = \infty$). Thus, the correction scheme shall be viewed as an exclusive correction for the DSBS error,[112] which is in contrast to a recent interpretation of eqn (25) as a method correction.[104]

$$E^{\text{corr}}(\text{i3-CC}|\text{MP2}) = E^{\text{corr}}(\text{i3-CC}) + [E^{\text{corr}}(\text{MP2}) - E^{\text{corr}}(\text{i3-MP2})] \quad (25)$$

It is possible to perform the initial HF calculation of the incremental scheme in the small basis B', as the charge centers of the localized orbitals, which are required in this step, are insensitive with respect to the basis set. However, the total HF energy must be computed in the large target basis set to obtain a sufficient accuracy and thus both calculations can be executed parallel. Since $B_{\mathbb{X}}$ differs from B', a unique mapping of the localized orbitals of $B_{\mathbb{X}}$ to those of B' is necessary for every domain to identify the orbitals to be correlated. In a first approach, this mapping was done by the charge centers, which is not unique, if the Boys algorithm may converge to different localizations for $B_{\mathbb{X}}$ and B', *i.e.* for systems with two or more Boys maxima.[127] This problem was solved by implementation of the so-called template localization.[129] Here the localized orbitals in B' represent the template orbitals and those of $B_{\mathbb{X}}$ shall be adapted. An optimal similarity of the localized orbitals in both basis sets is reached by maximizing the functional

$$D = \sum_k \langle \phi_k^{B_{\mathbb{X}}} | \phi_k^{B'} \rangle. \quad (26)$$
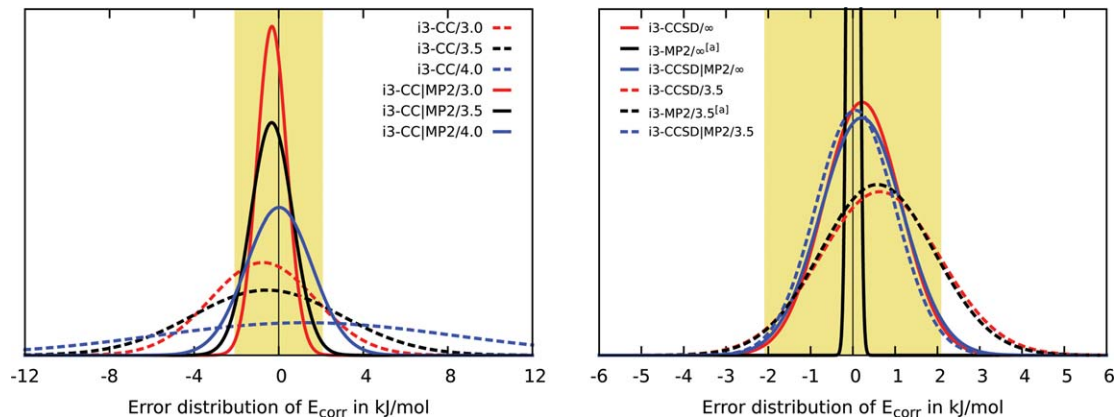
Fig. 5 Left: Gaussian error distributions for incremental CCSD(T)(F12*)/cc-pVDZ-F12 absolute energies with respect to the canonical calculations. The DSBS parameter t_{main} is varied (3.0, 3.5, 4.0 Bohr) and the success of the MP2 correction is shown.[117] Right: Gaussian error distributions for incremental CCSD/cc-pVDZ and MP2/cc-pVDZ absolute energies with respect to the canonical CCSD calculations. The key specifies in this sequence: method, t_{main}. The figure shows the errors of the incremental CCSD and MP2 results in the full target basis and in the DSBS, as well as the CCSD results with MP2 correction.[112] (a) canonical MP2 as reference to obtain local error at the MP2 level

i.e. the sum of overlaps between two assigned orbitals. The orbitals in $B_{\mathbb{X}}$ are transformed iteratively by orthogonal (2×2) rotations. For details see ref. 129.

A recent work also attended to the improvement of the environmental basis ($B' = $ DSBSenv), concerning the reduction of the computational effort.[130] Previously, due to an ad hoc decision, the environment of a domain has been described by an SV orbital basis set as well as an SVP auxiliary basis sets for density fitting of MP2 integrals and as complementary auxiliary basis set (CABS). To decrease the CPU times, smaller contraction patterns in comparison to the original basis sets were generated by recontraction and reoptimization of the contraction coefficients, using appropriate optimization techniques. Furthermore, even cheaper environmental basis sets ($B' = $ rDSBSenv) were generated by removing the respective basis functions with the highest angular momentum, leaving the exponents and previously optimized coefficients unchanged. While the accuracy of absolute energies was only slightly reduced by the new environmental basis sets, it was even enhanced for reaction energies (*cf.* Fig. 9 in Section 3.3.1). The reduced new environmental basis was recommended for the future use within the domain-specific basis set approach as it decreases the CPU time on average by a fifth, considering the test set of small- to medium-sized systems, and by two third for the large $Al(H_2O)_{25}{}^+$ cluster. For this structure, a second-order incremental expansion was both faster and more accurate than the tight-DLPNO approach, which is currently one of the most efficient local correlation methods.[25–27,131] The results will be discussed more detailed in Section 3.3.2.

3.1.3 Dual-basis set zero-buffer approach. Another approach to reduce the virtual space by a dual basis treatment has been implemented by Zhang and Dolg in 2013.[120] If the correlation energy shall be computed with a basis B, the initial SCF calculation, providing the MOs to be localized and partitioned, is performed with a smaller basis set B'. Here B' is formed from B by removing all basis functions with an angular momentum higher than p. The basis set $B_{\mathbb{X}}$ for the main region of a domain $\mathbb{X}$ is then obtained by adding all those removed basis functions, belonging to the main region. Thus, the initial occupied space of B' is maintained in $B_{\mathbb{X}}$ and no mapping is necessary. The virtual space of $B_{\mathbb{X}}$ is obtained by augmentation of B' and a subsequent Löwdin orthogonalization. Furthermore, a new procedure of generating the main region, based on K-means clustering,[118,119] was established. Using the (fixed) geometric centers $\bar{\mathbf{R}}_k$ of the $|\mathbb{D}|$ orbital domains, the complete set of atoms with coordinates $\mathbf{x}_{k_i}$ is also partitioned into $|\mathbb{D}|$ atomic domains, denoted as $\mathbb{M}_k$. Again, the functional

$$J = \sum_{k=1}^{|\mathbb{D}|} \frac{\sum_{\mathbf{x}_{k_i} \in \mathbb{M}_k} \left| \mathbf{x}_{k_i} - \overline{\mathbf{R}}_k \right|^2}{|\mathbb{M}_k|} \tag{27}$$

is minimized, providing atomic domains close to the orbital domains. For the calculation of a domain, consisting of one-body domains $\mathbb{D}_k$, only

the atoms of the corresponding $\mathbb{M}_k$'s are treated with the large basis B, *i.e.* with an augmented virtual space. Since no large-basis buffer region around the $\mathbb{M}_k$'s is applied, this is the so-called dual basis-set zero-buffer (db-B0) approach.

The effect of the zero-buffer approximation has been investigated.[120] First, CCSD and CCSD(T) feature similar local errors at third order. But more important, the neglect of a buffer region barely affects the accuracy of the absolute error, if the number of domains is not too high, *i.e.* the domain size is sufficiently large. Thus, the zero-buffer approximation is recommended for an efficient treatment of the virtual space. However, larger domains (compared to the usual dsp = 4 within the DSBS approach) were used in some cases to provide a sufficient accuracy without a buffer region and thus the efficiency due to the segmentation of the system may be lower.

An exemplary summary of the two described dual-basis set implementations (Sections 3.1.2 and 3.1.3) as well as of the corresponding domain partitionings is given in Fig. 6.

3.2 Approaching the complete basis set limit

Highly accurate wave-function-based methods are especially useful to provide results on a benchmark level, *e.g.* for the assessment of DFT functionals. As described in the introduction, their accuracy is on the one side determined by the size of the *N*-particle basis, *i.e.* the number of Slater determinants. In this context, CCSD(T) has been mentioned as the "gold standard" in quantum chemistry. On the other side, also the size of the one-particle basis set is crucial.[1,3] Thus, if highly accurate calculations are required, one has to get as close as possible to the complete basis set (CBS) limit. This means, the one-particle basis set error, also known as basis set incompleteness error (BSIE), should be reduced to a negligible amount in comparison to other errors.

In this context, one may think of the Coulomb cusps,[132] which occur in the exact wave function for vanishing interparticle distances r_{12}. Close to these points (*i.e.* for $r_{12} \rightarrow 0$) the wave function becomes linear with respect to r_{12}. With the commonly used Gaussian basis sets, which are smooth (*i.e.* continuous) even at $r_{12} = 0$, it is difficult to describe these cusps. Thus, a large number of basis functions is needed and, especially for the correlation energy, a slow convergence to the CBS limit with respect to the ζ-level of the basis sets is observed.[132,133] Due to hardware limitations and the high computational effort, coupled cluster calculations with sufficiently large basis set are not suitable for many cases. At this point, we would like to remark, that with the incremental scheme a larger basis set can be applied in comparison to a common canonical calculation, even at the same hardware. This is a significant benefit, which makes benchmark CCSD(T) calculations feasible, even for medium-sized systems. Nevertheless, it is convenient and often even necessary to accelerate the CBS convergence. Thus, there are different
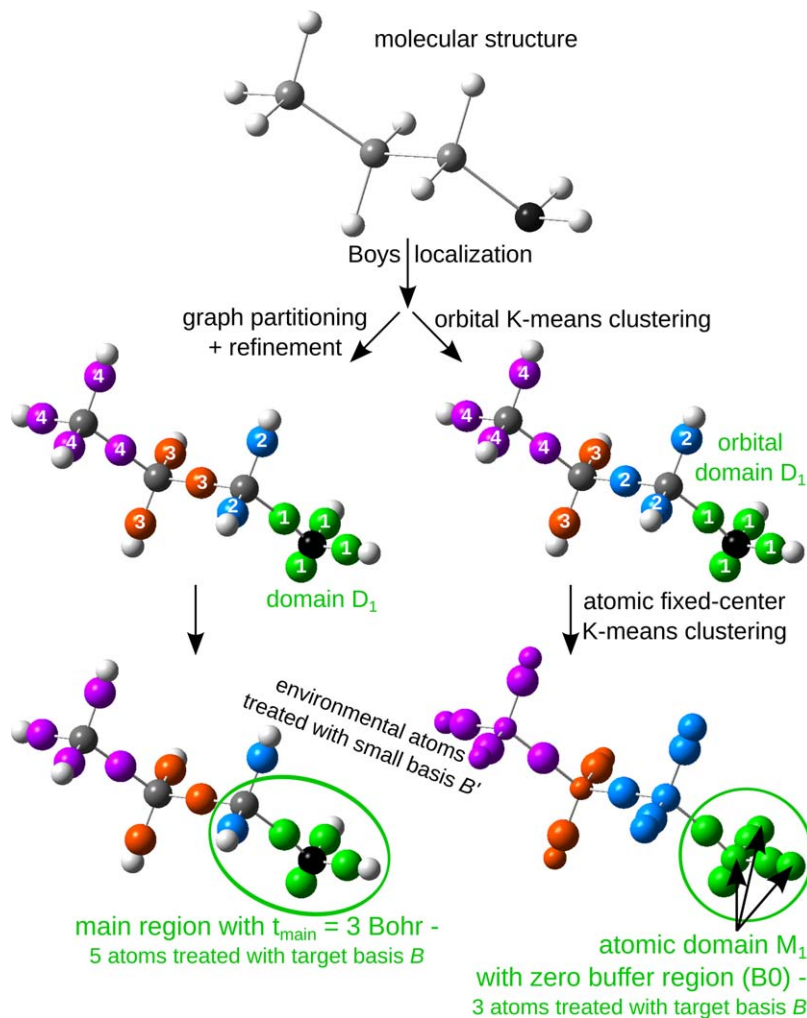
Fig. 6 Schematic representation of two ways of domain partitioning and dual-basis set treatment, represented for *n*-propylamine (atoms are depicted unnumbered and smaller than charge centers; H – white, C – grey, N – black). The 4 domains (with numbered charge centers) from graph partitioning with a subsequent refinement differ slightly from those obtained *via* K-means clustering. Defining the main region of the first domain by the distance parameter t_{main} (here: 3 Bohr) leads to 5 atoms to be treated with the larger (target) basis set, while after the fixed-center K-means clustering of the atoms, only 3 atoms are treated with basis B, when applying no buffer region (B0 approach).

techniques to improve this convergence. First, one may estimate the CBS limit by two-point extrapolations. Very useful are the CBS extrapolations for the HF energy

$$E_{XY}(HF) = \frac{e^{-a\sqrt{Y}}E_X(HF) - e^{-a\sqrt{X}}E_Y(HF)}{e^{-a\sqrt{Y}} - e^{-a\sqrt{X}}} \qquad (28)$$

of Petersson and co-workers[134] with the recommended parameter value $a = 6.3$ as well as the one for the correlation energy

$$E_{XY}^{\text{corr}} = \frac{X^3 E_X^{\text{corr}} - Y^3 E_Y^{\text{corr}}}{X^3 - Y^3} \qquad (29)$$

by Halkier *et al.*[135] Here E_X, E_Y are the respective energies, using a basis set at the ζ-levels X, Y.

Another approach to converge faster to the CBS limit is the use of explicitly-correlated methods, which are focused in the next section, considering also their implementation for the incremental scheme.

3.2.1 Explicitly-correlated methods. As suggested above, the description of the wave function may be improved, especially around the Coulomb cusps, if an explicit dependency on the interelectronic distances r_{12} is included.[133,136–139] To account for this, pair correlation effects are additionally described by geminal functions in explicitly-correlated methods. Thus, the "normal" doubles, used *e.g.* in CCSD or MP2 theory, are now supplemented by the operator $\hat{T}_2'$, which is defined as

$$\hat{T}_2' = \frac{1}{8} c_{xy}^{ij} w_{\alpha\beta}^{xy} \tilde{a}_{ij}^{\alpha\beta} \qquad (30)$$

Here $\tilde{a}_{ij}^{\alpha\beta} = a_\alpha^\dagger a_\beta^\dagger a_i a_j$ describes the double excitations from the occupied (or valence within the frozen-core approximation) HF orbitals $(i, j, \ldots)$ into a formally complete virtual space $(\alpha, \beta, \ldots,)$. The c_{xy}^{ij} are the amplitudes for these double excitations, similar to conventional double amplitudes t_{ab}^{ij}. Furthermore, the explicit correlation is introduced *via* the integrals

$$w_{\alpha\beta}^{xy} = \langle \alpha\beta | \hat{Q}_{12} f(r_{12}) | xy \rangle \qquad (31)$$

over the correlation factor $f(r_{12})$. Due to the near linearity of the wave function for $r_{12} \to 0$, f_{12} was made linearly dependent on r_{12} in the original R12 methods. A better description for also larger values of r_{12} is provided by an exponential Slater-type function,[140] as used within the F12 methods:

$$f_{12} = -\frac{1}{\gamma} e^{-\gamma r_{12}} \qquad (32)$$

In practical use, this correlation function is represented as linear combination of Gaussian-type functions.[141] Furthermore, $\hat{Q}_{12}$ is a projection operator ensuring orthogonality. There are different ansätze for $\hat{Q}_{12}$, denoted as 1, 2, 3.

However, there are a lot more approaches and approximations for the F12 methods to reduce the computational effort significantly. Within the general F12 model, both the conventional (CC or MP2) and the F12 amplitudes are optimized. As a first approximation, there is the SP ansatz, using fixed F12 amplitudes.[142] One advantage of this fixed-amplitude approach is the prevention of an additional basis set super-position error (BSSE) due to the geminals, which occurs if the F12 amplitudes are optimized. Furthermore, there are several implementations of CCSD-F12 with neglected higher-order contributions, *e.g.* CCSD(F12),[143,144] CCSD(F12*),[145] CCSD-F12a and CCSD-F12b.[146,147]

Due to the use of the geminal functions, higher integrals are introduced within the explicitly-correlated methods. Thus, it is especially important to use density fitting for the evaluation of these integrals, which is done by means of the complementary auxiliary basis sets (CABS) for the three- and four-electron integrals[148–150] as well as with the CBAS auxiliary basis sets for F12 and MP2 Coulomb integrals.[151,152]

Considering explicitly-correlated CCSD(T), the perturbative triples correction is in general the same as in the conventional method, *i.e.* the (T) energy is simply added to the CCSD-F12 energy.

For more information about explicitly-correlated methods in general as well as the specific ansätze and approximations we recommend the reviews in ref. 153 and 154.

In combination with the incremental scheme, the F12 versions of MP2, CCSD and CCSD(T) have been used. In a first work, incremental MP2- and CCSD-F12 have been tested critically by using one-body domains, consisting of single orbitals, for a test set of 15 small-molecule reactions.[155] The MP2-F12 energy is simply the sum of the MP2 energy and the F12 contribution, *i.e.* the convergence of the incremental expansion can be investigated separately for both parts. Even for the critical case of single-orbital domains the local errors of the F12 contribution are negligible already at the second order, it converges significantly faster with respect to the incremental order, while the local errors of CCSD(F12) are remarkably higher than those of MP2. Furthermore, a more efficient and similarly accurate small-basis-set approximation has been tested in this context, *i.e.* a TZ basis set is used for the first and second order and the DZ one for higher increments. The described errors are summarized in Gaussian distributions, shown in Fig. 7. However, the rather large errors
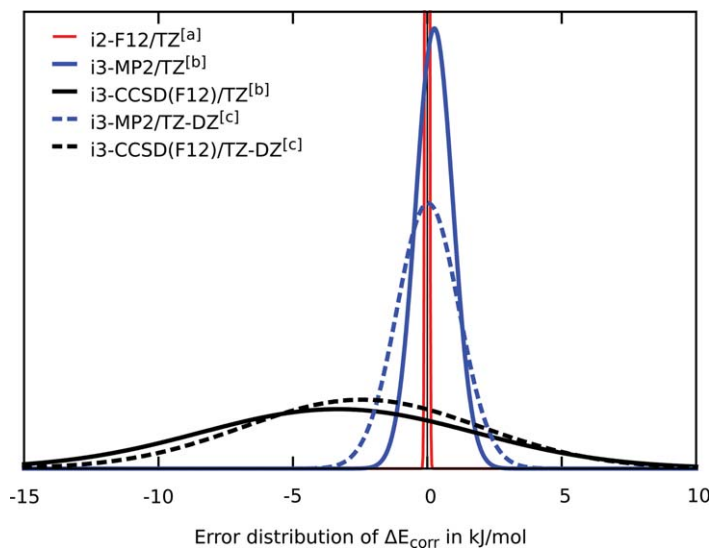
Fig. 7 Gaussian error distributions for the incremental MP2–F12 corrections, MP2 energies and CCSD(F12) energies with respect to the corresponding canonical cc-pVTZ-F12 calculations for 15 reaction energies.[155] (a) F12 contribution within the MP2–F12 method, TZ = cc-pVTZ-F12. (b) TZ = cc-pVTZ-F12. (c) TZ-DZ = cc-pVTZ-F12 for first and second order; cc-pVDZ-F12 for third order.

at third incremental order can be significantly reduced by using larger domains, *e.g.* with dsp = 3 or 4.

Later, CCSD(T)-F12 results have been presented, including also the DSBS approximation in combination with the CCSD(T)(F12)[156,157] and CCSD(T)(F12*)[117] methods as well as the db-B0 approach with CCSD(T)-F12x (*x* = a,b).[158] We will discuss the accuracy and efficiency, *i.e.* of these methods, in the next section.

3.3 Incremental CCSD(T) as benchmark method

3.3.1 Accuracy. Applying the previously described techniques of approaching the CBS limit to the "gold standard" method CCSD(T), the resultant methods may be used for benchmark calculations. However, prior to that the accuracy of incremental CCSD(T) has to be checked. A first ingredient to obtain a benchmark with the incremental scheme are negligible errors due to the local approximations. The main part of the incremental error comes from the DSBS approximation, but it can be significantly reduced by the described MP2 correction (see Section 3.1.2).

Figure 8 shows again the effect of the MP2 correction, which is based on the similarity of the MP2 and CCSD(T) local error. The behaviour for conventional as well as explicitly-correlated methods is comparable, as the test sets in ref. 159 and 157 are also similar. The (F12*) approximation performs slightly better than (F12), considering the local errors of the incremental scheme. However, comparison of the results is difficult, since they are received with different test sets. Though, the similarity of the both i3-CCSD(T)(F12*)|MP2-F12 curves, obtained from completely differing test sets,[117,130] demonstrates the "transferability" of the local
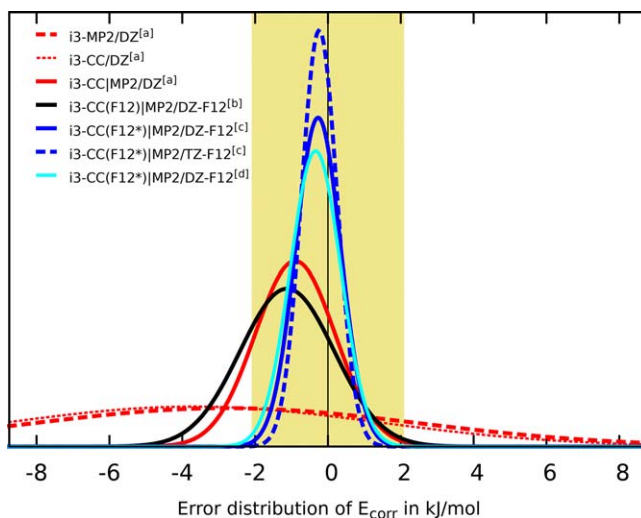
Fig. 8 Gaussian error distributions of third-order incremental absolute energies with respect to the corresponding canonical calculations. The abbreviation CC refers to CCSD(T), while the MP2 correction in the (F12) or (F12*) calculations is performed as MP2-F12. The applied target basis sets are cc-pVDZ (DZ), cc-pVDZ-F12 (DZ-F12) and cc-pVTZ-F12 (TZ-F12). Different test sets were used, according to the following references. (a) Ref. 159, 34 molecules. (b) Ref. 157, 27 molecules. (c) Ref. 117, 81 molecules. (d) Ref. 130, 37 molecules.

error values. Furthermore, the local errors are rather insensitive with respect to a change of the basis set, as shown here for cc-pVDZ-F12 and cc-pVTZ-F12.[117] The error decreases even slightly, when increasing the basis set quality, *i.e.* the basis set can be extended further within the incremental scheme without a loss of accuracy.

For reaction energies (Fig. 9) an enlargement of the basis set again decreases the local errors slightly.[117] The introduction of the recently optimized environmental basis (rDSBSenv) increases the accuracy, although it provides reduced computational efforts.[130] Finally, considering the local error distributions for both absolute and relative energies, i3-CCSD(T)(F12*)|MP2-F12 is almost completely within a range of 1 kcal mol^{-1}, the so-called chemical accuracy (see highlighted area in Figs. 8 and 9). Therefore, the incremental scheme provides a sufficiently high accuracy, since the local errors are negligible in comparison to *e.g.* method errors.

The i3-CCSD(T)(F12*)|MP2-F12 method has been compared exemplarily to i3-db-B0-CCSD(T) for the cyclization reaction[160] in Fig. 10.[117,120] For a proper comparison not only the local errors, but also the computational effort should be taken into account. Thus, also the number of basis functions and the CPU times are given in Table 1. In general, also the domain sizes respectively number of domains is decisive. However, for the considered system the influence of the domain partitioning is not essential, as both methods use 6 one-body domains, each with 4 ± 1 correlated orbitals, and are therefore better comparable. Using the DSBS approach, the application of the MP2 correction lowers the local errors clearly, while the additional CPU time is negligible. The local errors for
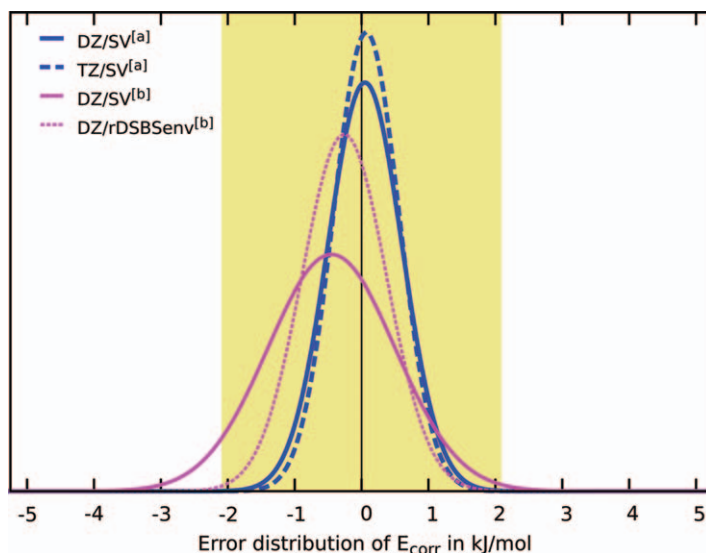
Fig. 9 Gaussian error distributions of third-order incremental CCSD(T)(F12*)|MP2-F12 reaction energies with respect to the corresponding canonical calculations. As target basis sets cc-pVDZ-F12 (DZ) and cc-pVTZ-F12 (TZ) were used and SV respectively rDSBSenv refer to the environmental basis sets. Different test sets were used, according to the following references (a) ref. 117, 51 reactions. (b) Ref. 130, 17 reactions.
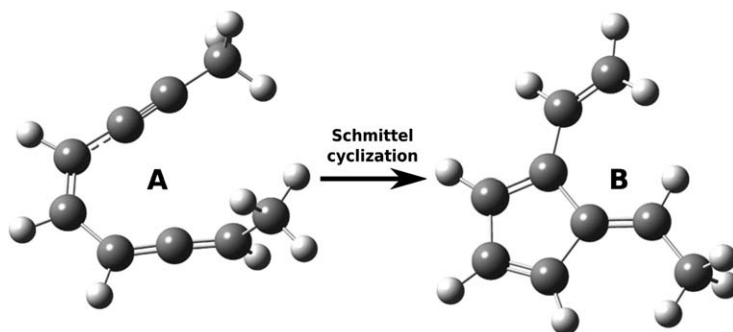
Fig. 10 Schmittel cyclization,[160] as computationally investigated in ref. 120 and 117.

Table 1 Total CPU times and local errors for educt (**A**), product (**B**) and reaction energy (ΔE) of the Schmittel cyclization, using third-order incremental expansions. The dual-basis set zero-buffer approach (dbB0) as well as the domain-specific basis set approach are compared and the basis with the number of basis functions (N_{bf}) is given for a better assessment. CC refers to CCSD(T), while the used basis sets are cc-pVDZ (DZ), cc-pVDZ-F12 (DZ-F12) and cc-pVTZ (TZ).[117,120]

Method	Target basis/ N_{bf}	CPU time [d] A	B	Local Error [kJ mol^{-1}] A	B	ΔE	ΔE [kJ mol^{-1}]
dbB0-CC	DZ/176	0.1	0.1	1.84	0.50	-1.34	-196.7
CC(F12*)	DZ-F12/360	1.71	2.12	-1.10	-2.55	-1.45	-191.9
CC(F12*)|MP2	DZ-F12/360	1.72	2.13	-0.26	-0.97	-0.71	-191.2
dbB0-CC	TZ/410	2.5	2.5	0.76	-1.18	-1.94	-188.8

i3-db-B0-CCSD(T) are then slightly higher, even for the larger TZ basis set, which might be due to the absent buffer region.

Next, there are some useful approaches to either increase the accuracy of the incremental results or improve the efficiency.

Considering the explicitly-correlated methods, the F12 contributions lead to a remarkable decrease of the BSIE. However, they do not influence the HF energy. To improve also the HF part, one may apply the complementary auxiliary basis sets (CABS) singles correction, which describes single excitations into the formally complete virtual space by means of perturbation theory.[147,161] Furthermore, the perturbative triples energy is not affected by explicit correlation, but an increase of the accuracy is also desired for this contribution. Therefore, it is possible to scale the triples by e.g. a constant[162] or the following simply-estimated factor as introduced by Marchetti and Werner[163]

$$E^{(T)}(F12) = \frac{E(MP2\text{-}F12)}{E(MP2)} \cdot E^{(T)} \tag{33}$$

as well as to extrapolate the triples by e.g. the Schwenke type extrapolation[164,165]

$$E^{(T)}(CBS(XY)) = E^{(T)}(XZ\text{-}F12) + F_{XY}(E^{(T)}(YZ\text{-}F12) - E^{(T)}(XZ\text{-}F12)) \tag{34}$$

The latter approach has been used in a recent work to obtain an incremental reference for benchmarking binding energies of molecular clusters.[166] In the same work, a more efficient way of reaching the CBS limit by explicit correlation has been discussed. Therein the CCSD(T) calculation (CC) is only performed in a small basis, while the resulting BSIE is estimated at a lower level of theory (e.g. MP2) by using a larger basis set with nearly CBS accuracy[23,167–173]

$$E^{\text{corr}}(\text{CC/CBS}) = E^{\text{corr}}(\text{CC/SB}) + [E^{\text{corr}}(\text{MP2/CBS}) - E^{\text{corr}}(\text{MP2/SB})] \quad (35)$$

More precisely, for such an incremental focal-point method (foX-i n-CC/Y with Y = SB), the MP2/CBS calculation was performed with MP2-F12/cc-pVXZ-F12. Since no explicitly-correlated coupled-cluster calculation is necessary, this variation is very efficient, but features also the problem of introducing an error due to the different convergence behaviour of CC and MP2 with respect to basis set size. To account for this error, an interference correction ΔE^{INT} [174–176] according to

$$E(\text{CCSD(T)/CBS}) \approx E(\text{CCSD(T)}) + \underbrace{\sum_{ij} \Delta e_{ij}}_{\Delta E^{F12}} + \underbrace{\sum_{ij}(f_{ij} - 1)\Delta e_{ij}}_{\Delta E^{\text{INT}}} + \Delta E^{\text{CABS}}$$

$$(36)$$

with

$$\Delta e_{ij} = e_{ij}(\text{MP2-F12}) - e_{ij}(\text{MP2}) \quad (37)$$

has been applied together with the incremental scheme.[159] The summation of the first two terms in eqn (36) is equivalent to eqn (35) and the individual factors f_{ij} and orbital pair energies e_{ij}, obtained from the first-order MP wave function, account for interference effects.[177,178] This scheme was successfully assessed (see Fig. 8 for the local errors) and applied for the calculation of highly accurate binding energies in MOF-CO_2 structures as well as a subsequent DFT benchmarking.[159]

With negligible local errors, different techniques of approaching the CBS limit and further corrections, the incremental scheme can now be used to receive highly accurate energies. Finally, the incremental scheme should also be compared to other methods. A very extensive study has been performed for the binding energies of molecular clusters[166] with results, shown in Fig. 11. This comparison shows, that all three techniques to approach the CBS limit with the help of the incremental scheme provide very accurate results, especially for the respective larger basis set. Amongst them, the smallest errors are obtained by the focal-point method, which might also be the most efficient of these three methods, since a large basis set with explicit correlation is only necessary at the MP2 level. The errors of the common incremental scheme, which are, by the way, shown for a second-order truncation, are larger due to the BSIE, when using DZ/TZ basis sets without explicit correlation. Compared to this method, DLPNO performs slightly worse, while the errors of
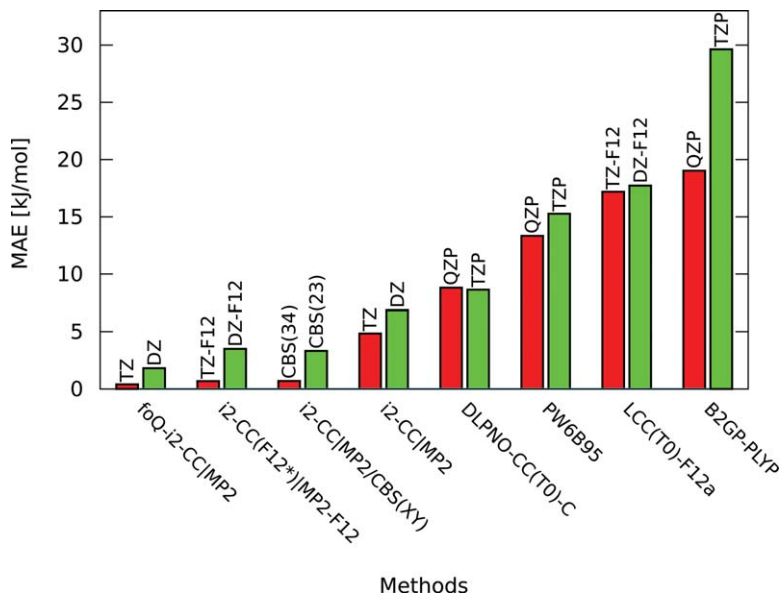
Fig. 11 Mean absolute errors of binding energies for molecular clusters with respect to the benchmark energies in ref. 166 obtained with different quantum chemical methods and basis sets. The following abbreviations are used: foQ = focal-point method with MP2-F12/cc-pVQZ-F12 as MP2/CBS (see eqn (35)). CC = CCSD(T), CC(T0) = CCSD(T$_0$). XZ(-F12) = cc-pVXZ-(F12), CBS(XY) = extrapolation cc-pVXZ→cc-pVYZ, XZP = def2-XZVPP.

LCCSD(T$_0$)-F12a are even larger than those of the best DFT method, using the PW6B95 functional.[179] All the other DFT methods (*e.g.* B2GP-PLYP) in this study provided a lower accuracy.

3.3.2 Efficiency. At the end of this section, we want to discuss also the efficiency and the computational efforts of the incremental scheme. It is well-known, that the formal scaling of CCSD(T)[5] with the number of basis functions ($\mathbb{N}$) is $\mathbb{N}^7$, or more precisely $\mathbb{O}^3\mathbb{V}^4$, if the number of occupied ($\mathbb{O}$) and virtual ($\mathbb{V}$) orbitals is considered. Therefore, a similar scaling should also be observed, if the molecular system is enlarged, keeping the same basis set.

The aim of local correlation methods is a significant reduction of this scaling. If no distance truncation would be performed within the incremental scheme, the scaling of the number of calculations N with the number of one-body domains $|\mathbb{D}|$, representing the system size, were[121,180]

$$N = \sum_{i=1}^{O} \binom{|\mathbb{D}|}{i}$$

Applying the distance truncation and assuming a system, which is already large enough, the addition of further domains would increase the total number of calculations approximately linear. With every new domain there is only a small number of new increments, which may be fluctuating, but is neither dependent on $|\mathbb{D}|$ nor increasing in another

way. Concerning the computational effort for a single increment, the formal scaling with respect to $\mathbb{O}$ is removed, since only a smaller number (approximately dsp) of occupied orbitals is correlated in every domain. Though, the dependency on $\mathbb{V}$ remains, as there is no distance truncation for the virtual space, but due to the dual-basis set approaches it is based on the size of the small environmental basis.

For an efficient method, the incremental scheme has to be computationally less demanding in comparison to the standard calculation. Therefore, the relative total CPU times with respect to those of the canonical CCSD(T) computations are shown in Table 2 for some systems, method settings and basis sets. It can be seen, that a higher efficiency of the incremental scheme is reached, if the size of the system (in terms of correlated electrons) as well as the size of the basis set (in terms of the number of basis functions) increases. Furthermore, a partitioning into smaller (and therefore more) domains reduces the computational effort of an incremental computation, as shown for $C_{10}H_{22}$.[127] The necessary MP2 calculations for the DSBS error correction require a negligible additional amount of time (see Table 1).[117] Also timings for both open-shell implementations of the method, which will be presented in Section 3.4, are denoted. Here the implementation of an unrestricted treatment within the DSBS approach[107] causes larger computation times as the spin-restricted db-B0 method[108] and a similar efficiency is reached only with a higher number of correlated electrons respectively basis functions. For the smallest systems with only a DZ basis, the total CPU time of all incremental calculations is even longer than standard one. However, a

Table 2 Efficiency of the incremental scheme for different systems and methods. Furthermore, CC abbreviates CCSD(T), aVXZ the aug-cc-pVXZ basis (X = D,T) and DZ(-F12) the cc-pVDZ(-F12) basis. N_{bf} and $N_{corr.e-}$ refer to the number of basis functions respectively correlated electrons. The total time of the incremental method (t_{inc}) is shown relative to the one of the standard calculation (t_{st}).

System	Ref.	Method	Basis set	N_{bf} ($N_{corr.e-}$)	t_{inc}/t_{st} [%]
Closed-shell					
$(H_2O)_6$, prism	158	db-B0-CC-F12	aVDZ	246 (48)	133.3
$C_{10}H_{22}$	127	CC(dsp = 4)	DZ	250 (62)	91.5
$C_{10}H_{22}$	127	CC(dsp = 3)	DZ	250 (62)	73.3
$Na(H_2O)_6^+$	130	CC(F12*)|MP2	DZ-F12	327 (56)	21.7
$Na(H_2O)_6^+$	130	CC(F12*)|MP2	DZ-F12[a]	327 (56)	16.3
$(H_2O)_{10}$	157	CC(F12)|MP2	DZ-F12	480 (80)	41.7
$(H_2O)_6$, prism	158	db-B0-CC-F12	aVTZ	552 (48)	53.4
$Na(H_2O)_6^+$	130	CC(F12*)|MP2	TZ-F12	596 (56)	9.3
$Na(H_2O)_6^+$	130	CC(F12*)|MP2	TZ-F12[a]	596 (56)	7.2
Open-shell					
DPC[b] (t)	108	db-B0-CC	DZ	232 (62)	135.3
dHBDI[c] (d)	108	db-B0-CC	DZ	279 (81)	65.7
DPC[b]-CH$_3$OH (t)	108	db-B0-CC	DZ	280 (76)	62.9
$Na(H_2O)_6^{\bullet}$(d)	107	CC	(a)VTZ	478 (49)	68.1

[a] Optimized rDSBSenv environmental basis instead of SV.
[b] DPC = diphenylcarbene.
[c] dHBDI = 4-hydroxybenzylidene-2,3-dimethylimidazolinone.

significant benefit of the incremental scheme, regarding its efficiency, is the possibility to parallelize the computation of the single increments as much as desired. Therefore, the wall time may be a small fraction of the standard one, depending on the number of used cores. Furthermore, it is difficult to show the efficiency for large systems, as the standard calculations are often not feasible in these cases. Thus, the incremental scheme ensures the feasibility in case of an enlargement of the system on the one side and additionally features a significantly increasing efficiency on the other side. Examples for the large systems, treated with the incremental scheme are shown in Fig. 12.

Finally, the high efficiency of the incremental scheme in combination with a very high accuracy has been demonstrated in a recent work for the interaction energy of a large cluster $Al(H_2O)_{25}^{3+}$ (see left side of Fig. 12).[130] Some results of this investigation are shown in Table 3. Using a second-order CCSD(T) expansion with MP2 correction, we could receive a benchmark energy, as a CBS(45) extrapolation was possible. Considering the same calculation with a TZ basis set, the CPU time was reduced by 63%, when applying the described new environmental basis (rDSBSenv). The deviation of the latter calculation is about 2 kJ mol^{-1} from the old environmental basis (SV) and only 4 kJ mol^{-1} from the
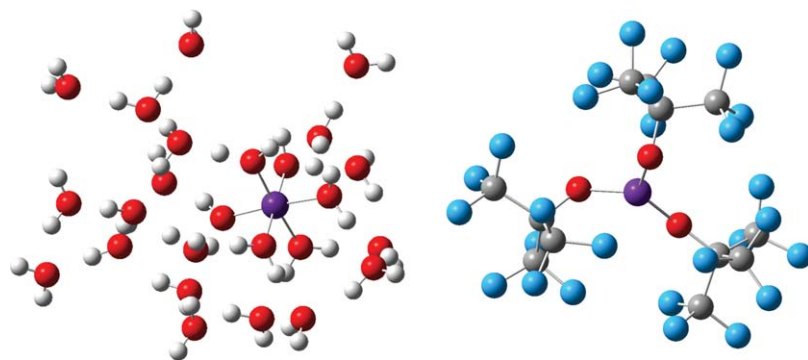
Fig. 12 Two large systems, treated with the incremental scheme. The binding energy of the $Al(H_2O)_{25}^{3+}$ cluster (left, 208 correlated electrons, 76 atoms) was calculated up to a 5Z level basis set (6106 basis functions) with a second-order incremental expansion.[130] A reaction of $Al(OC(CF_3)_3)_3$ (right, 258(!) correlated electrons, 43 atoms) was investigated by a third-order incremental expansion, using the def2-TZVP basis set (1339 basis functions).

Table 3 Correlation contributions to the interaction energy of $Al(H_2O)_{25}^{3+}$ and corresponding total CPU times for the second-order incremental scheme with MP2 correction (i2CC|MP2) and the DLPNO-CCSD(T$_0$) method.[130] SV and rDSBSenv refer to the environmental basis, tight and normal to the parameter settings for the PNOs. The benchmark energy is obtained at the CBS(45) level with SV environmental basis.

	i2CC|MP2		DLPNO-CCSD(T$_0$)		i2CC|MP2
Basis	SV	rDSBSenv	Tight	Normal	Benchmark
ΔE [kJ mol^{-1}]	−532.0	−529.7	−497.6	−476.3	−525.8
CPU time [d]	36.0	13.4	19.4	2.3	—
Wall time [d]	0.53	0.24	—	—	—

benchmark energy. In contrast, the DLPNO-CCSD(T_0) method of Neese and co-workers with tightPNO, as recommended for interaction energies,[27] provides an error of 28 kJ mol^{-1} at the TZ level. Furthermore, the CPU time for this approach is larger than for the second-order incremental calculation with reduced environmental basis. Thus, the incremental scheme provides a significant higher accuracy as well as a better efficiency. Using the DLPNO method with the normalPNO parameters[27] would reduce the CPU time, but introduces an error of about 50 kJ mol^{-1}. Furthermore, the incremental scheme provides an essential benefit in terms of its parallelizability due to the calculation of the increments, which is represented by the significantly reduced wall times.

3.3.3 Applications: reactions and non-covalent interactions.

The incremental scheme has been used in several different applications to compute CCSD(T) reaction or non-covalent interaction energies on a benchmark level. The incremental scheme is especially useful for systems, which are too large to be treated with standard CCSD(T) at a given hardware, *i.e.* it is fully-automated, strongly parallelizable, computationally efficient and provides negligible local errors, when using the optimized parameter settings.

On the one hand, the incremental scheme was applied to compute highly accurate reaction energies supplementary to DFT values. A first example was the reaction of methyl methacrylate (MMA) with 2-phenyl-2-oxazoline (POX), yielding two different intermediate structures.[181] Comparing the incremental CCSD(T)(F12) reaction energies with DFT, the PBE0 functional[182] performed best, while the other functionals produced significantly different results. As the most sufficient functional for this system, PBE0[182] was applied to model the outstanding reaction steps.

A similar proceeding was employed in another work on the mechanisms for rearrangement reactions of tritylcarbenes.[183] Due to the system size the incremental scheme was used for CCSD(T)(F12*) calculations to benchmark the corresponding DFT reaction and activation energies. For this system the PW6B95 functional[179] in combination with the D3 dispersion correction of Grimme[184] turned out to perform best. Finally, a deeper understanding of the reaction mechanisms could be achieved.

Considering non-covalent interactions, the incremental scheme has been used in a multiscale approach, computing highly accurate binding energies for different hydrogen fluoride clusters.[156,185] These energies were necessary input parameters for a subsequent QM/MM scheme, the quantum cluster equilibrium (QCE) method,[186,187] to study hydrogen fluoride in condensed phase. The importance of using a high-accuracy method like CCSD(T) at the CBS limit has been shown, since binding, and therefore the density in liquid phase, is overestimated by DFT as well as underestimated by correlated methods (MP2, CCSD(T)) in combination with too small basis sets. In this work, the CBS limit was approximated by a CBS(34) extrapolation, using the aug-cc-pVXZ basis sets. The incremental scheme was necessary to keep the aug-cc-pVQZ calculations for the $(HF)_n$ rings ($n = 6,8,12$) feasible.

In another study, the QCE model was extended and refined for the description of condensed phase *N*-methylformamide-water mixtures.[188] Here, CCSD(T)(F12*) was used to calculate binding energies of the required molecular clusters and for benchmarking the corresponding DFT results. The highly-accurate binding energies were again necessary to compute parameters for the QCE model. The largest clusters could be only treated with the help of the incremental scheme.

Furthermore, incremental results have been used to assess a new and improved parametrization of the semiempirical polarized molecular orbital (PMO) model.[189] In this context, the method has been applied to six clusters, that contain sulfuric acid, ammonia as well as dimethyla-mine and represent atmospheric aerosol. The benchmark energies were obtained from CCSD(T)(F12)/cc-pVTZ-F12 calculations.

The same method in combination with the incremental scheme was utilized for $(H_2O)_{26}$ clusters, received from Monte Carlo simulations.[190] Such large clusters are necessary to represent the condensed phase in a sufficient way. With the obtained benchmark energies, many-body expansions were assessed with respect to the use of an electrostatical embedding as well as the truncation order.

In another study on four $(H_2O)_{20}$ clusters, which are part of the WATER27 test set,[191] new benchmark values for the binding energies were determined.[192] Therefore, MP2 and CCSD(T) calculations were performed up to the 5Z basis set level for CBS(45) extrapolations. Furthermore, MP2-F12 and CCSD(T)(F12*) energies with basis sets up to cc-pVQZ-F12 provided another way to reach the CBS limit. Additionally, the effect of other contributions (BSSE correction by the site–site function counterpoise (SSFC) method, core-valence correlation) at the MP2 level has been investigated. Close to the CBS limit, which is necessary for benchmark energies, the CP correction is small (0.3 kcal mol^{-1}), but core-valence correlations (1.2 kcal mol^{-1}) should be taken into account. The results from a third-order incremental CCSD(T)(F12*)|MP2-F12 expansion, complemented by CP and core-valence contributions at the MP2 level, were taken as reference values for the assessment of DFT functionals. But also explicitly-correlated CCSD(T)(F12*) and even MP2-F12 provide very accurate binding energies, when using a QZ-F12 basis. However, the previous benchmark values[193] vary significantly from new and more accurate ones, obtained with the incremental scheme. The errors of the DFT methods, which have been larger within the old benchmark,[194] turned out to be better than expected. Especially the PW6B95 functional[179] in combination with the D3 dispersion correction[184] was convincing.

In two works, Dolg and co-workers calculated Gibbs free energies to investigate the hydration of lanthanoid(III)[195] and actinoid(III)[196] ions. Herein the gas phase binding energies of the Ln(H$_2$O)$_n$ and An(H$_2$O)$_n$ ($n = 8,9$) clusters were computed by means of the db-B0 incremental method. By adding solvent and entropic contributions, obtained at the DFT level (with COSMO model respectively the freeh module), as well as an explicit hydrogen bond correction, the Gibbs energies for these hydrations were received. The lanthanoid(III) clusters were computed at

different levels of theory, where the binding energies are overestimated by DFT and underestimated by MP2, both in a significant way. Beside CCSD and SCS-MP2 also HF performs well, indicating an convenient error compensation. However, these first principle hydration Gibbs free energies were obtained with less than 2% deviation from the experimental values.

3.4 Open-shell systems

As shown above, the incremental scheme has been applied to various chemical problems, dealing with closed-shell molecules. But since there is also a large necessity of treating *e.g.* radicals or high-spin transition-metal molecules, the incremental scheme has been extended to open-shell systems. In a first work it has been shown, that a third-order expansion provides also small local errors, using a spin-restricted RCCSD approach for these systems.[106] Therefore, the modified "active space" incremental expansion has been implemented, based on the ideas of Mödl *et al.* for an incremental multi-reference treatment of open-shell thioferrate(III) complexes (see also Section 4.3.2).[84] The active space MOs, *i.e.* the singly-occupied ones, are in general significantly more delocalized and should therefore be correlated with all double-occupied ones. Hence, the active space MOs are not included in the partitioning scheme and afterwards assigned to every one-, two-, three-body,... domain. To avoid a multiple counting of the correlation within the active space, the incremental series is expanded up to the O-th order as

$$E_{(O)}^{\text{corr}} = E_{\text{active}}^{\text{corr}} + \sum_{\mathbb{X} \in \mathscr{P}(\mathbb{D}) \wedge |\mathbb{X}| \leq O} \Delta \varepsilon_{\mathbb{X}} \qquad (38)$$

$$\Delta \varepsilon_{\mathbb{X}} = \varepsilon_{\mathbb{X}} - E_{\text{active}}^{\text{corr}} - \sum_{\mathbb{Y} \in \mathscr{P}(\mathbb{X}) \wedge |\mathbb{Y}| < |\mathbb{X}|} \Delta \varepsilon_{\mathbb{Y}} \qquad (39)$$

Within the incremental RCCSD method, the mechanism of a 4-*exo* cyclization, which had been focused in some previous computational works,[197,198] was studied on a post-HF level. Furthermore, the singlet–triplet gap of an Arduengo carbene as well as the electron affinity of a water hexamer cluster were investigated by means of the incremental method.[106]

Recently, the dual-basis set zero-buffer approach of Dolg *et al.* as well as the domain-specific basis set approach of Friedrich *et al.* have been implemented for the treatment of open-shell molecules on the CCSD(T) level of theory. The former one operates in a spin-restricted scheme, just like in the first work on incremental RCCSD, but extended to a dual-basis set treatment.[108] The method was applied to compute the singlet–triplet of diphenylcarbene (DPC) in context of explicit solvation. While the free DPC is more stable in a triplet state, solvation by one methanol molecule generates a singlet state. Furthermore, the vertical detachment energy (VDE) of the green fluorescent protein (GFP) chromophore was investigated by means of model systems, representing a gas phase as well as a

protein-like structure. It was found, that a positively charged protein fragment increases the VDE significantly in comparison to the gas phase, *i.e.* from 278 to 586 kJ mol^{-1}, which was obtained with the i3-db-B0-CCSD(T) method and a CBS(23) extrapolation.

However, the open-shell method with the DSBS approach has been implemented in an unrestricted scheme and thus the assignment of α and β orbitals was a crucial step.[106] Therefore, two approaches (α domain partitioning, template equalization) were introduced, which provide approximately the same local errors (see Fig. 13). The open-shell method was assessed by means of a test set, *i.e.* on a statistical basis. Again, the MP2 correction works very well. Furthermore, the local errors are similar to those of the closed-shell incremental scheme.

3.5 Molecular properties

Beside the calculation of absolute and relative energies, there is a broad variety of other properties of interest. Some of them characterize the interaction of molecules with an external field, *i.e.* the dependence of the energy on its strength. Therefore, it is obvious to expand the energy in a Taylor series, which has the following form,[2] when considering a uniform electric field with the strength $\mathbf{F}$ and an expansion relative to the unperturbed energy $E(0)$:

$$E(\mathbf{F}) = E(0) + \underbrace{\left(\frac{\mathrm{d}E}{\mathrm{d}\mathbf{F}}\right)_0}_{-\boldsymbol{\mu}_0} \mathbf{F} + \frac{1}{2!} \underbrace{\left(\frac{\mathrm{d}^2E}{\mathrm{d}\mathbf{F}^2}\right)_0}_{-\boldsymbol{\alpha}} \mathbf{F}^2 + \frac{1}{3!} \underbrace{\left(\frac{\mathrm{d}^3E}{\mathrm{d}\mathbf{F}^3}\right)_0}_{-\boldsymbol{\beta}} \mathbf{F}^3 + \cdots \quad (40)$$

The variation of the energy with the electric field strength corresponds to the dipole moment

$$\boldsymbol{\mu} = -\frac{\mathrm{d}E}{\mathrm{d}\mathbf{F}} = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\mathbf{F} + \frac{1}{2}\boldsymbol{\beta}\mathbf{F}^2 + \cdots \quad (41)$$

where $\boldsymbol{\mu}_0$ is the permanent dipole moment and the other terms represent the induced part of the polarization. Within these the polarizability $\boldsymbol{\alpha}$ corresponds to a linear dependence on $\mathbf{F}$ and the hyperpolarizabilities (*e.g.* the first hyperpolarizability $\boldsymbol{\beta}$) cause non-linear effects. Considering *e.g.* the interaction of an electromagnetic light wave with matter, the (hyper)polarizabilities lead to (non-)linear optical properties.

Since the described molecular properties are derived energy expressions, they are usually tensors. Thus, every element of this tensor is expanded separately according to eqn (7). The suitability of computing molecular properties from a truncated incremental expansion was focused in several works. First, Yang and Dolg applied the Incremental Scheme to $Ga_4As_4H_{18}$ as a model system for the non-linear optical (NLO) material GaAs.[109] Whereas the CCSD polarizabilities converged very well with the order of the expansion, the hyperpolarizabilities were more critical. Furthermore, the domain decomposition was varied and found to be very crucial. It was shown, that the increments for the hyperpolarizabilities are not only affected by the size and distance of the
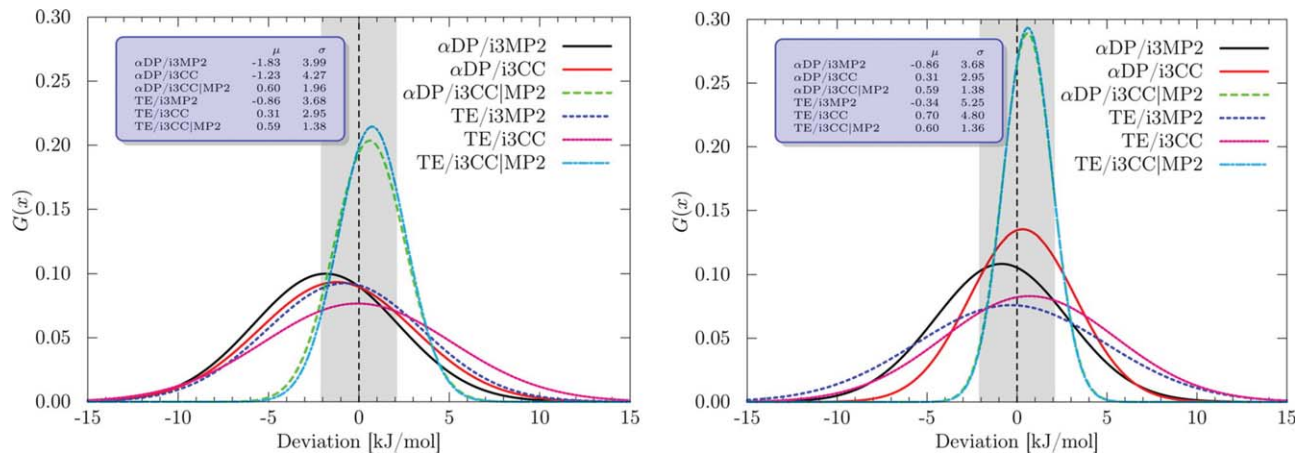
Fig. 13 Gaussian distributions of the third-order incremental local errors, using aug-cc-pVXZ (left: X = D, right: X = T) basis sets.[106] αDP refers to the α domain partitioning, TE to the template equalization and CC to CCSD(T). Reprinted with permission from ref. 107. Copyright 2016 American Chemical Society.

domains, but also significantly by the orientation and distance of the dipole moments situated at the domains. For example, a strong coupling between two domains with parallel dipoles has been observed.

Later, the fully-automated incremental scheme has been used to compute orbital-unrelaxed first-order properties,[110] precisely dipole and quadrupol moments, for a test set, containing molecules of different chemical nature (polar and non-polar, aromatic and non-aromatic). When using appropriate domain sizes (dsp = 3 in this work) the incremental series is well converged, even at low orders. Herein the convergence is slightly worse for aromatic compounds.

Recently, also the domain-specific basis set (DSBS) approach has been implemented for the calculation of dipole moments.[112] In this context, the optimal set of truncation parameters was found to be similar to the common one in incremental energy calculations. Due to the failure of the MP2 correction scheme for unrelaxed properties, a similar CC2-based one was successfully introduced to avoid the DSBS error and decrease the local errors, obtaining a high accuracy of the incremental scheme (see left side of Fig. 14).

In another work, the incremental DSBS approach has been utilized for frequency-dependent polarizabilities.[111] Here, the local errors were found to be negligible for a third-order incremental expansion with both augmented (aug-cc-pVDZ = AVDZ) and nonaugmented (cc-pVDZ = DZ) basis sets. However, the importance of the augmentation was demonstrated, as DZ leads to significant deviations of averaged 16%. Thus, it was also tested to compute the lower-order increments with augmentation and the higher ones without (see right side of Fig. 14). A sufficient reduction of the error is already obtained, if only the first-order increments are calculated with AVDZ. Furthermore, the errors in this work were found to be smaller for linear and weakly bound structures compared to globular ones, i.e. the accuracy of the polarizabilities depends significantly on the shape of the molecules.

4 Incremental methods for periodic systems

4.1 Basis sets and effective core potentials

A crucial aspect, especially for solids, are the basis functions, in which the wave function is expanded. For infinite solids with translational symmetry plane-wave basis sets are a suitable choice. However, as the method of increments deals with cluster models and local effects, the use of more localized basis functions (e.g. Gaussian-type atomic orbitals) is appropriate. The common Gaussian-type basis sets are usually constructed for molecular systems and therefore contain diffuse functions to describe also the edge of these structures in a sufficient way. However, in solids a quite dense packing occurs and every atom is completely surrounded by neighbouring atoms, which describe also this atom with their basis functions. Therefore, diffuse functions are usually not necessary. They are even problematic, since they provoke linear dependencies due to the dense packing. Thus, special basis sets for crystals with eliminated diffuse functions are needed and in most applications of the incremental
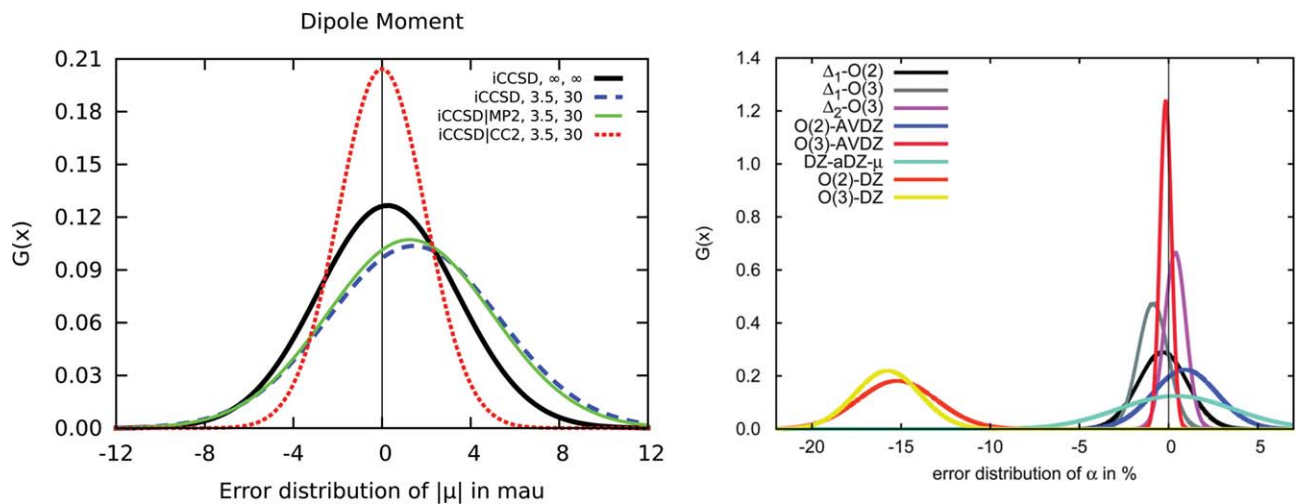
Fig. 14 Left: Gaussian error distributions of CCSD/cc-pVDZ absolute dipole moments for third-order incremental expansions with respect to the canonical results. The key specifies in this sequence: method, t_{main}-value, f-value. The figure shows the local errors in the full target basis, the domain-specific basis and using an MP2 or (in the same way) a CC2 correction. Reprinted with permission from ref. 112. Copyright 2016 American Chemical Society. Right: Gaussian distributions of local and basis set errors for CCSD polarizabilities with respect to the canonical AVDZ calculations. O(2) and O(3) refer to the incremental orders. Δ_1/Δ_2 denote the computation of increments up to first/second order with AVDZ, whereas the other increments are obtained with the DZ basis. AVDZ refers to the aug-cc-pVDZ basis, DZ to cc-pVDZ. Reproduced from ref. 111 with permission from the PCCP Owner Societies.

method for periodic systems, the basis sets were previously optimized. However, the neglect of the diffuse functions is not that drastic for the accuracy as neighbouring atoms are described by each others basis functions, due to the dense packing.[199]

Another problem associated with the basis is the polynomial scaling of the computational effort with the number of basis functions. Concerning most of the properties, the core electrons are definitely not as decisive as the valence electrons and can therefore be described on a less sophisticated level. For this purpose, effective core potentials are very suitable, as they enable also a treatment of heavier atoms. Within the method of increments different energy-consistent small- and large-core pseudopotentials have been applied, which were fitted to scalar-relativistic data in many cases. The corresponding basis sets were also used, but modified. In some cases the contraction was changed or too diffuse s/p functions were removed. Furthermore, polarization functions (e.g. 1d, 2d1f) were often added. Due to the modifications the coefficients and exponents had to be optimized subsequently. For first- and second-period elements with an all-electron treatment the correlation-consistent basis sets of Dunning and co-workers were generally used.

Due to the description of the core electrons by pseudopotentials, also static and dynamic core polarization effects are excluded. Whereas the first one is the core polarization at the HF level, the latter one corresponds to the core-valence correlation. One way to account for this, are core polarization potentials (CPPs), which describe the interaction of the core dipole with the electric field, generated by its surroundings (e.g. valence electrons). Within test calculations for GaAs it was found, that the contributions from CPPs are very local as even two-body increments of them are negligibly small.[70,199] Therefore, the cohesive energy is barely affected by the core polarization. In contrast, the CPPs are essential for other ground-state properties, especially for the bulk modulus (see also Fig. 17 in Section 4.4).[69,71,72]

4.2 Embedded cluster models for insulators and semiconductors

The introduction of a cluster model for calculating the correlation energy of an infinite system implies a significant approximation. Thus, it is crucial to choose a suitable cluster, especially concerning its size. Furthermore, it is essential, to work with localized orbitals, which are similar to the Wannier orbitals in the corresponding infinite system.[199] For this purpose, an embedding of the considered cluster is in general necessary. There are different approaches for such an embedding, depending on the binding properties.

4.2.1 Covalent systems. The first applications of the method of increments provided correlation energies for elemental semiconductors (C, Si, Ge, Sn), all with diamond structure.[54,67,69] Later also III,V-[70,72,73] (e.g. BN, GaAs, InSb) and IV,IV-semiconductors[73] (SiC) have been treated. Although especially the III,V-semiconductors are somewhat polar, all these three-dimensional periodic systems are dominated by covalent bonding. Furthermore, also two- (graphene[68]) and

one-dimensional covalent systems (*e.g.* linear polymer chains[200–203]) as well as (finite) Buckminsterfullerenes[204] have been focused. Since the charge centers of the localized MOs are located between two covalently bound atoms, it is obvious to define the one-body domains "by bonds". Thus, one single-bond orbital corresponds to one domain. The generation of cluster models for covalent systems proceeds by cutting them out of the crystal structure without a relaxation, *i.e.* all geometry parameters remain unchanged. However, cutting-out leads to unsaturated exterior atoms. An efficient and simple, but in most cases sufficient approach is the saturation with hydrogen atoms.[54] Furthermore, it is often necessary not only to form the cluster from the atoms with adjacent correlated bond orbitals, but also to embed them with additional atoms. Thus, another important issue is the size of the clusters, *i.e.* it is essential to screen a required increment with respect to the fragment size. As the HF orbitals of such an embedding are frozen in the correlated calculation, this may be called Hartree–Fock embedding. In this context, it is also possible to treat the atoms of the HF embedding with a smaller basis set than the "main region".

Two material properties, which are connected to electron localization and correlation, are conductivity and band gap (E_g). For example, the local treatment of electron correlation in metallic systems is rather sophisticated. Therefore, the changes from the insulator diamond ($E_g \approx 5.5$ eV) *via* the semiconductors silicon and germanium with smaller band gaps to the semi-metal α-tin have been studied, concerning the performance of the incremental method.[69] First, the localization procedure has been found to perform well in all cases with even more localized orbitals for α-Sn. Furthermore, a similar convergence behaviour of the incremental series was observed for the four systems. Considering the necessary cluster size, significant differences occur, as shown in Table 4 for the first-order increment. For diamond the smallest possible cluster, *i.e.* a direct H saturation, seems to be sufficient, which is due to the similarity of the C–C and C–H bonds. In contrast, a larger cluster is definitely necessary for α-Sn. Furthermore, the convergence with the cluster size for the polar semiconductor GaAs[70] is even better compared to tin. However, to ensure an appropriate accuracy, the smallest possible cluster should preferably be enlarged by at least one additional shell of atoms.

Table 4 One-body increment (in kJ mol^{-1}) of the central X–Y-bond for different sizes of $X_n\,Y_n\,H_{4n+2}$ clusters (X = Y = C, X = Y = Sn, X = Ga/Y = As).[69,70] Basis sets at VDZ level, CEPA-0 method for C and Sn, CCSD for GaAs. Number in parentheses refers to percental deviation from the value of the largest cluster ($n = 4$).

Cluster size n	C	Sn	GaAs
1	−66.2 (+0.2)	−48.7 (−4.0)	−52.4 (−3.5)
2	−66.3 (± 0.0)	−47.2 (−2.6)	−49.4 (−0.5)
3	−66.4 (−0.1)	−45.9 (−1.3)	−49.3 (−0.4)
4	−66.3	−44.6	−48.9

Beside the ground-state properties, which are discussed in Section 4.4, the method of increments has been applied to study the zinc-blende wurtzite polytypism of SiC and III,V nitrides.[73] The cluster models, obtained from the HF-optimized crystal structures, were treated at the CCSD level. For SiC, the zinc-blende structure is only 1.8 kJ mol^{-1} more stable than the wurtzite structure, corresponding to the experimentally found mixed structures. In contrast, for the III,V nitrides the wurtzite structure is more stable with *e.g.* an energy difference of 3.4 kJ mol^{-1} for GaN, for which a zinc-blende could yet be observed in an experiment. For AlN (9.5 kJ mol^{-1}) and InN (6.0 kJ mol^{-1}) only wurtzite structures occur. This example shows the sound agreement between theory and experiment.

Furthermore, also systems with multiple bonds were investigated. Herein σ and π orbitals may either be assigned to different domains, as done for graphene,[68] or to the same domain, like in the works on poly-acetylene[200] and the fullerene C_{60}.[204] It was shown, that the incremental expansion also works well for such delocalized systems, even if the convergence is in general slightly worse and π-bonds require basis sets with larger ζ-levels. The correlation energy (per C atom) of C_{60} is between those of diamond and graphene. Another phenomenon, which represents this intermediate character, is the bond alternation. While a uniform bond length occurs in diamond as well as in graphene, the bonds in C_{60} (and even more in polyacetylene) feature two different lengths. In the fullerene, a larger bond alternation is supported by the HF part and the one-body correlations, whereas it is counteracted by correlations within the π-system.

The incremental method has also been applied to covalent one-dimensional systems, more precisely to polymer-chains, *e.g.* poly-acetylene,[200] beryllium hydride polymer[202] or polymethineimine.[201] The correlation energies of oligomers $[M]_n$ were calculated incrementally for increasing chain sizes n. The correlation energy of a monomer unit M_∞ within an infinite polymer chain is then[200]

$$E_{M_\infty}^{\text{corr}} = \lim_{n \to \infty} \frac{E_{[M]_n}^{\text{corr}}}{n} \tag{42}$$

or

$$E_{M_\infty}^{\text{corr}} = \lim_{n \to \infty} \left[E_{[M]_{n+1}}^{\text{corr}} - E_{[M]_n}^{\text{corr}} \right] \tag{43}$$

The latter approach features a faster convergence with respect to n, as finite-size effects (in terms of chain terminations) are approximately cancelled. The incremental method was applied to calculate cohesive as well as polymerization energies, which were used to study *e.g.* the polymerization mechanism in the specific case of polyhydrophosphazenes.[203]

4.2.2 Transition to ionic systems. A lot of strongly ionic materials, which are therefore insulators, have been focused by means of the incremental method. Examples are earth alkali oxides (MgO, CaO),[75,76] alkali halides (*e.g.* LiF, NaCl, RbBr, KI)[77,79] or transition-metal crystals (*e.g.* NiO, AgCl, cubic AuCl),[78,205] for which also ground-state properties

have been calculated (see Section 4.4). When cutting out an appropriate cluster from such an ionic crystal, it is obvious to apply a point-charge embedding with the formal charges of the ions. That way, the electrostatic interactions between the cluster fragment and its environment are taken into account. Since this type of interaction has a long-range character, several shells of point charges may be necessary. Furthermore, it has been found, that in such strongly ionic systems the correlation of the cation is nearly independent of its explicit surroundings, *i.e.* it can be simply embedded in the point-charge grid.[75] In contrast, the anion must always be embedded in one shell of cation pseudopotentials to describe the surrounding Madelung field of the solid in an appropriate way.[75] A direct embedding in hard point charges would be not sufficient, as the distortion of the electron-rich and thus easily polarizable anion would be too intense, resulting maybe even in a non-ionic situation. By means of the pseudopotentials the surrounding positive charges are softer, which is closer to the real solid.

Considering the two-body increments, anion–anion correlations are remarkably larger than their cation–cation counterparts. Therefore, the former ones have to be evaluated for more nearest neighbours, while the latter ones may even be neglected for some cations (*e.g.* Mg^{2+}, Li^+). First, the two-body increments should obviously decrease with increasing distances, *i.e.* lattice constants. However, this tendency is superposed (or even inverted) due to the significant influence of the polarizability of the involved ions. An increasing polarizability of both cations (*e.g.* from Li^+ to Rb^+) and anions (*e.g.* from F^- to I^-) enlarges also the respective ion-pair two-body increment. Considering *e.g.* the alkali chlorides (XCl), the Cl–Cl increments decrease from LiCl to KCl, since the lattice constants increase. In contrast, the X–Cl increments are increasing due to the rising polarizabilities of X^+. Thus, in LiCl the Cl–Cl increments are larger and in KCl the X–Cl increments.[77,79]

In a work on purely ionic NiO, the method of increments has been successfully extended to transition-metals with partially filled d bands using the quasidegenerate variational perturbation theory (QDVPT).[205] Here, the two-body Ni–Ni increments were additionally computed for high-spin (ferromagnetic) and low-spin (antiferromagnetic) coupling. While next-neighbouring Ni^{2+} ions tend to the former one, second-nearest neighbours prefer the latter. Therefore, the overall splitting is very sensitive towards the cluster size as well as to the embedding and was therefore not discussed quantitatively.

However, there are many systems with a binding character between covalent and ionic, like II,VI-semiconductors (*e.g.* ZnS, CdSe, CdTe).[71] In these cases, one has to decide, whether a point-charge embedding and the definition of the domains by ions is appropriate. Thus, a covalent treatment, as described in the previous section, should also be taken into account. Such an investigation has been performed for the series ZnS, ZnSe, ZnTe, using an X_4Y_4 cluster, embedded (I) in only hydrogen atoms for saturation, (II) in another shell of X/Y atoms and the corresponding H saturation, (III) in 56 point charges, *i.e.* an overall cell of $4 \times 4 \times 4$ size. As shown in Table 5 embedding (I) performs significantly worse,

Table 5 CCSD valence correlation energy (in kJ mol^{-1}) per unit cell for Zn(II),VI-semiconductors, using different embeddings. Basis sets at a VDZ level.[71]

Embedding	ZnS	ZnSe	ZnTe
(I)	−497.0	−461.6	−446.3
(II)	−484.9	−433.2	−390.9
(III)	−476.0	−429.8	−372.8

especially for ZnTe. However, both embedding (II) and (III) are sufficient for these intermediate systems, but the ionic description (III) may be favored due to the computational efficiency.

The method of increments has also been applied to other transition-metal materials, which feature no fully ionic, but rather an intermediate character with covalent parts. Examples are rutile (TiO_2)[74] or AuCl,[78] both with a tetragonal structure. The established ionic description with point-charge-embedding and pseudopotentials for the first-shell cations has been applied also in these cases, but the transferability of the same increment between clusters of different sizes (in these cases: one, two and three ions) had to be checked. It was found, that this transferability is worse than for pure ionic crystals. This is due to the partial charge transfer in such intermediate materials, which is represented very differently for various, especially small cluster sizes. However, larger clusters and basis sets would be required. A representative example is AuCl, which has been considered in a theoretical cubic as well as in the experimental tetragonal structure.[78] The former one is purely ionic with a good transferability, whereas it is worse for the latter one, which features a partial covalent character in terms of the Au(I)–Au(I) interactions. These closed-shell aurophilic interactions had been reported in experimental works and could be found by means of the method of increments. While the nearest-neighbour Au–Au increment of the cubic structure provides only −3 kJ mol^{-1}, the one of the experimental structure is about −15 kJ mol^{-1}.

Furthermore, also rare earth oxides $(CeO_2$[206,207]) and nitrides (CeN, LaN,[208] GdN[209]) were treated, extending the incremental method to systems with f shells. For a second-order incremental expansion it was found, that the correlation energy for LaN and GdN is nearly 100% of the experimental value, whereas only 72% are obtained for CeN. The 4f shell is empty for La^{3+} as well as half-filled for Gd^{3+}, so that the applied pseudopotential is sufficient. In contrast the single 4f electron of Ce^{3+} is more delocalized and therefore the incremental series is not well converged at second order. However, an incremental expansion with partially-filled f bands is also suitable, but more sophisticated. The treatment of CeO_2 is easier due to the $4f^0$ configuration of the Ce^{4+} ion. The three-body contribution is negligible and different embedding approaches were tested.[207] First, a Ce_4O_7 cluster, embedded in ECP54 pseudopotentials for the first-shell Ce^{4+} ions and point charges apart from that, was used to calculate all increments for the same system and to obtain a HF embedding. Furthermore, as in most previous cases, individual clusters for every increment with a point-charge or a

point-charge + pseudopotentials (ECP46/ECP54) embedding, were used. Supplying the first-shell Ce^{4+} ions by a contracted [1s1p1d1f] basis set was also tested. The latter approach as well as the pure point-charge embedding overestimate the correlation energy significantly, while the Ce_4O_7 cluster provides proper results. This is due to nearly charge-neutral Ce_4O_7 fragment in the infinite solid, representing the importance of similarity (and therefore transferability) of the cut-out cluster.

4.2.3 Weakly bound systems. Beside the described strongly bound crystals, the method of increments was also applied to periodic systems with weak interactions, like atomic (*e.g.* rare gas[100,103]) and molecular (*e.g.* dihalogen,[105] CO_2/HCN[104]) van-der-Waals crystals as well as hydrogen-bonded systems (*e.g.* H_2O/NH_3/CH_3OH,[125] infinite chains of HF/HCl[101,102]). In these cases, the one-body domains are obviously formed from single atoms or small molecules.

In two studies on infinite zig–zag HX chains (X = Cl,F), the basis set convergence was investigated. As the binding energies in weakly bound systems are small, the application of large basis sets and/or BSSE corrections are decisive to reach an adequate accuracy. While a pure 5Z basis set treatment provided deviations between CP-corrected and non-corrected binding energies up to 20%, a sufficient accuracy was only obtained by CBS extrapolations ($3 \rightarrow 4$, $4 \rightarrow 5$).[101] As these hydrogen-bonded systems contain also a slightly ionic character, an embedding with point charges has been tested. If the oligomers $(HX)_n$, used to compute the increments, have an adequate length ($n = 10$ in this work) no point-charge embedding is necessary. Beside the increments, representing short-range correlations, also long-range correlations have been estimated by means of two-body van der Waals dispersion terms $\varepsilon_{0n}^{vdw} = - C/R_{0n}^6$. This function was fitted to the incremental two-body increments, calculated for different monomer distances R_{0n}.[102] Similar approaches to estimate the long-range correlations have also been performed for other systems.[103,105]

Even less sophisticated in terms of the embedding are rare-gas crystals.[103] Here in a simpler approach compared to the HF embedding, the cluster is embedded into ghost atoms. That means only the basis functions of the embedding atoms remain, but their electrons/charges are not taken into account. Furthermore, even bare cluster models, *i.e.* clusters without embedding, were studied. It has been found, that all three embedding approaches can provide a sufficient accuracy. However, the less sophisticated the embedding is, the more increments must be considered, *i.e.* the convergence of the incremental series is slower.

In an approach for molecular crystals of CO_2 and HCN, the main contribution of the correlation energy was obtained with the periodic LMP2 method and only correction terms were evaluated incrementally.[104] Thus, a basis set correction

$$\Delta\Delta E(CBS) = \Delta E(inc\text{-}LMP2/CBS) - \Delta E(inc\text{-}LMP2/(A)VTZ) \quad (44)$$

as well as a post-MP2 correction

$$\Delta\Delta E(post\text{-}MP2) = \Delta E(inc\text{-}CCSD(T)) - \Delta E(inc\text{-}MP2) \quad (45)$$

were used to obtain cohesive energies at a CCSD(T)/CBS level. For this correction, both a bare cluster approach as well as an embedding with nearest neighbour molecules and point charges, representing the partial charges of more distant molecules, have been tested. The bare cluster model was found to provide an appropriate accuracy with a higher efficiency, so that also three-body increments could be taken into account.

4.3 Metallic systems and multi-reference treatment

4.3.1 Special treatment of metals. In contrast to the insulators and semiconductors, conducting metallic materials are more challenging systems for incremental methods. However, many metallic solids were focused in various works, *e.g.* Li,[86,210] Hg,[80,211,212] Mg,[81,212] Zn, Cd,[82,212] Be,[83] Ca, Sr.[87]

When considering a metallic system, two problems arise from the electron delocalization. For an appropriate basis set (*i.e.* more than the minimal one), a localization procedure, which is required within the incremental method, generates poorly localized orbitals with long-range tails.[211] Thus, one has to expect a bad convergence behaviour of the incremental series. Furthermore, another problem, concerning the delocalization, arises from the embedded finite-cluster model. Whereas the delocalized electrons are distributed homogeneously in a real, infinite metallic crystal, they are transferred to the ''surface'' in a cluster.[211] Since only the resulting positively charged central atoms are correlated, the description is no longer balanced. In order to solve the two problems, a special treatment of metals within the method of increments has been developed:[211,213]

1. *Cluster model:* Proper cluster models have to be generated, similar to the common incremental method. As the metallic systems consists of atoms, they are the obvious choice of one-body domains. For a specific increment, the involved atoms are assigned to the central part and embedded in additional atom shells of various sizes.

2. *Localization:* Both the atoms in the central part and in the embedding must be described by a minimal basis set. Such a basis features no virtual orbitals, which ensures that no charge transfer from the central part to the surface of the embedding takes place. Therefore, the surroundings of the central part are described sufficiently, concerning electrostatic interactions. After applying one of the common localization procedures to the central and embedding orbitals, the minimal basis provides also well-localized occupied orbitals and thus both problems, mentioned above, are solved.

3. *Delocalization in the central part:* However, due to the absent possibility of electron delocalization within the minimal basis, no metallic character is provoked. Thus, for a proper description the basis set in the central part needs to be extended to a larger one, decreasing the BSIE and generating a metallic character. The latter obviously occurs only in the central part, giving the corresponding increment. Therefore, the metallic character of the total system is expanded successively, according to the order of the incremental expansion. For the purpose of introducing a

larger basis, recalculations of the integrals as well as a HF calculation with reoptimization of the central orbitals and freezing of the embedding ones have to be performed. That way the reference wave function for the subsequent correlation calculation is obtained.

4. *Correlation calculations:* Finally, the reoptimized orbitals are correlated within a post-HF calculation, again with frozen embedding orbitals and, in most cases, large-core pseudopotentials at the embedding atoms. It is also possible to correlate different parts of the "outer" electrons, which has been performed for Mg as well as the group-12 metals Zn, Cd, Hg. In these cases the valence-only contributions of s^2 electrons were determined, but also the semi-core d^{10} and underlying core s^2p^6 electrons were additionally included to identify the influence of the different shells. Considering the calculation of a two-body increment $\Delta\varepsilon_{ij}$ according to eqn (3), ε_i and ε_j should be computed within the same cluster model as ε_{ij}, due to the high sensitivity of the increments with respect to the cluster size and embedding for metallic systems. However, using the calculated increments of all required domains as well as the corresponding weight factors the total correlation energy is expanded incrementally.

As the treatment and character of metallic systems differ significantly from those of insulators and semiconductors, the different parameters within the method of increments need to be screened again. First, there is the truncation of the incremental series. Due to the stronger delocalization, a slower convergence behaviour may be possible. It was shown, that an embedding of the central part is decisive for a fast convergence, since a free cluster approach provides significant four- and even higher-body increments, as shown especially for Mg, but also for Hg.[80,81] However, using a proper embedding the three-body increments are necessary to get a suitable accuracy and fourth-order contributions are small, but were taken into account for several systems. To find out the size of such an adequate embedding, the number of atom shells has been varied for different systems. Thus, only one (for Hg) or two (for Mg) shells of embedding atoms were applied. In other works a cut-off parameter for the maximum distance of central part and embedding was found to provide a sufficient accuracy and efficiency for 1.5 times the nearest-neighbour distance.

Similar to Hg, the cohesion of Zn and Cd is only due to electron correlation, as the HF cohesive energy is positive for all three cases.[80,82,212] Furthermore, a distance screening of two-body increments was performed for Zn and Cd, yielding a polynomial $-C/r^m$ function with m slightly smaller than 6, which would represent the frequently assumed van-der-Waals like decrease of the correlation energy.[82] The method of increments was also applied to compute the potential energy surface of Zn and Cd with respect to the parameters a and c of the hexagonal crystal structure. For both elements including the correlation contributions of the full d shell was essential. In case of Zn, three local minima could be identified,[214] with one very close to the experimental structure, featuring an unexpected large c/a ratio. But a more stable and denser packing is

obtained for a smaller c/a value. For Cd the PES provides a single and rather flat minimum region close to the experiment.[215] That way the experimental cohesive energies could be nearly obtained with the method of increments, providing also an insight into the influence of different correlation contributions.

In contrast to the previous cases, Be features a more covalent character.[83] To account for this, another embedding scheme was applied with the inner embedding shell, described by a [2s1p] basis set instead of the common minimal [2s] basis set, used apart from that. Furthermore, the near degeneracy of 2s and 2p provides a multi-reference case, but single- and multi-reference results were found in good accordance, *i.e.*. the dynamic correlation is predominant and a single-reference treatment possible. However, there are several metals with a more distinct multi-reference character (*e.g.* Li, Ca, Sr), which will be discussed in the next section.

4.3.2 Multi-reference cases. There are many systems with quasi-degenerate orbitals, *e.g.* around the Fermi surface of metals or in strongly-correlated materials. In these cases, the mean-field description of the ground state by only one (HF) determinant, *i.e.* a single-reference (SR) method, is also qualitatively wrong. Thus, more configurations need to be taken into account by using a multi-reference (MR) approach, for which the incremental method has also been implemented. According to this implementation, increments are first obtained from a (size-extensive) multi-configurational self-consistent field (MCSCF) method like CASSCF (complete active space self-consistent field). Therefore, an active space of the N_o highest occupied molecular orbitals (HOMOs) and the N_v lowest virtual ones (LUMOs), regarding the SR HF determinant, is defined and afterwards all possible distributions (configurations) of the N_e HOMOs electrons in the active space are taken into account as determinants. This is refered to as CAS(N_e, $N_o + N_v$). Within the MR incremental method for metals,[87,210] an active space is defined for every atom. For example, in case of n atoms of Ca, Sr[87] or Be[83] a CAS($2n,4n$) calculation, consisting of the fully-occupied valence s orbitals as well as the lowest virtual p orbitals, was sufficiently accurate. Next, not only the occupied orbitals, but also the virtual orbitals in the active space of a domain are localized. Then the CASSCF calculation can be performed and the resulting energy, relative to the HF energy, yields a first correlation increment for every domain.

Afterwards a calculation with a MR correlation method is performed on top of the CASSCF wave function. As the incremental approach relies on the size-extensivity of a method, the choice of a suitable method is critical. A method, which is only approximately size-extensive, is the multi-reference averaged coupled pair functional (MRACPF),[216,217] which was always used to compute the additional correlation increments. Thus, a size-extensivity error, especially for higher-order increments, was also found for this method in a work on Li_n rings.[210] The MR incremental approach has been further applied to get magnetic coupling constants for open-shell high-spin thioferrate(III) complexes by means of MRCI and

ACPF.[84] In other works dissociations of C–C and C–H bonds,[218] the systems Ca and Sr[87] as well as dissociation curves and the metal-insulator transition of Be_n clusters[88] have been studied.

4.4 Ground-state properties of materials

In several works important ground-state properties, more exactly cohesive energies, lattice parameters and bulk moduli, have been computed for bulk materials by means of increments.

The cohesive energy E_{coh} corresponds to the energy, which is necessary to disaggregate a unit cell (u.c.) of a crystal with energy $E_{u.c.}$ into free atoms. Within this definition it is computed by

$$E_{coh} = \sum_i E_i - E_{u.c.} \quad i \in \{\text{free atoms}\} \tag{46}$$

and therefore generally holds a positive sign. If the opposite definition is used, negative values are possible, representing the attractive interaction of the free atoms, when forming the solid. For molecular clusters, "cohesive energy" has also been used as the binding energy between the molecules, *i.e.* it describes the disaggregation into free molecules.

Since the atoms in a crystal or cut-out cluster are packed very dense, an atom is considerably described by the basis functions of its neighbours, even if the chosen basis sets are less diffuse. Thus, one can expect a large basis set superposition error (BSSE) in the cohesive energy, since the free atoms are not described by additional basis functions. Hence, in many applications of the incremental method the cohesive energies have been corrected for the BSSE,[199] *e.g.* by the counterpoise (CP) method.

The calculated results for the cohesive energies of various systems are presented in Fig. 15 with respect to the experimental ones. Furthermore, Gaussian distributions of the absolute errors in the cohesive energy as well as for the percental errors of the correlation contributions are shown in Fig. 16. As the HF results are far below the equivalence line of calculated and experimental values, the cohesive energies are clearly underestimated. For some metallic systems the HF cohesive energies are even negative, representing the repulsive interaction of the metal atoms at this level of theory. However, due to the binding character, the cohesive energies for the metals are significantly smaller than for the ionic and covalent systems, but well reproduced by correlated methods. The accuracy of the CCSD cohesive energies for the ionic systems is slightly below that of CCSD(T), but clearly larger than in the covalent cases. Considering all systems, the average error is about 14% for the correlation contributions, but only 6% for the total cohesive energy. However, due to large values of the cohesive energies, these 6% are equal to nearly 45 kJ mol^{-1}.

Another important issue is to get an optimized geometry of the system on a sophisticated post-HF level, *i.e.* in terms of lattice parameters for a given crystal structure. Since there is still no implementation of an
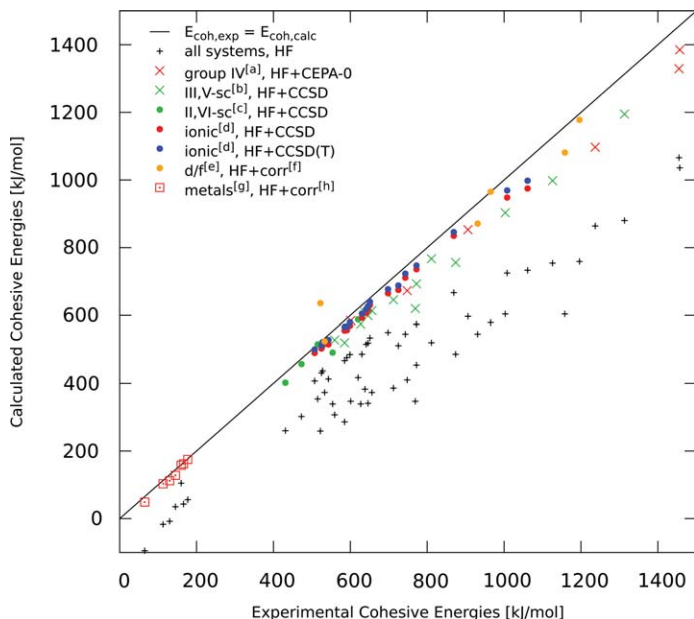
Fig. 15 Calculated cohesive energies with respect to the corresponding experimental ones for various subsets of bulk materials. (a) Diamond, graphene, SiC, Si, Ge, α-Sn.[68,69,73] (b) $X^{3+}Y^{3-}$ semiconductors (X = B, Al, Ga, In; Y = N, P, As, Sb) with covalent cluster models.[70,73] (c) $X^{3+}Y^{3-}$ semiconductors (X = Zn, Cd; Y = S, Se, Te) with ionic cluster models.[71] (d) Alkali halides[77,79] and earth alkali oxides[75,76] with ionic cluster models. (e) Ionic systems with d- and f-valence metals.[74,78,205,206,208,209] (f) CCSD or CCSD(T) (g) metals (Hg, Zn, Cd, Li, Mg, Be, Ca, Sr).[80–83,86,87,212] (h) CCSD(T) (single-reference) or MRACPF (multi-reference).

analytical gradient within the incremental scheme, geometry optimizations have to be done numerically.

There were investigations of *e.g.* tetragonal crystal structures[74,78] (*i.e.* with parameters *a* and *c*), but in most cases cubic crystals with only one lattice parameter *a* have been considered. Therefore, the total energy was calculated for different values of *a*, close to the experimental or HF-optimized one, and the resulting data was fitted by a polynomial function. The analytical minimum of this graph represents approximate values for the optimized lattice parameter and the corresponding total energy. For the fitting quadratic[72] as well as fourth-order functions[71] have been used and a variation of the fit for the same systems yielded nearly the same results.[71]

These results for different systems are summarized in Fig. 17 (left side), which presents the errors of the calculated lattice constants with respect to the experimental ones as Gaussian error distributions. The percental deviations are in general small. However, at the HF level the lattice constant are overestimated by 3% on average with a broad variance. The accuracy is significantly enhanced by the correlation contribution. Considering the covalent and ionic subsets, for the former one a higher accuracy is achieved, whereas for the latter one CCSD(T) provides slightly better results than CCSD. For the covalent subset, the importance of the
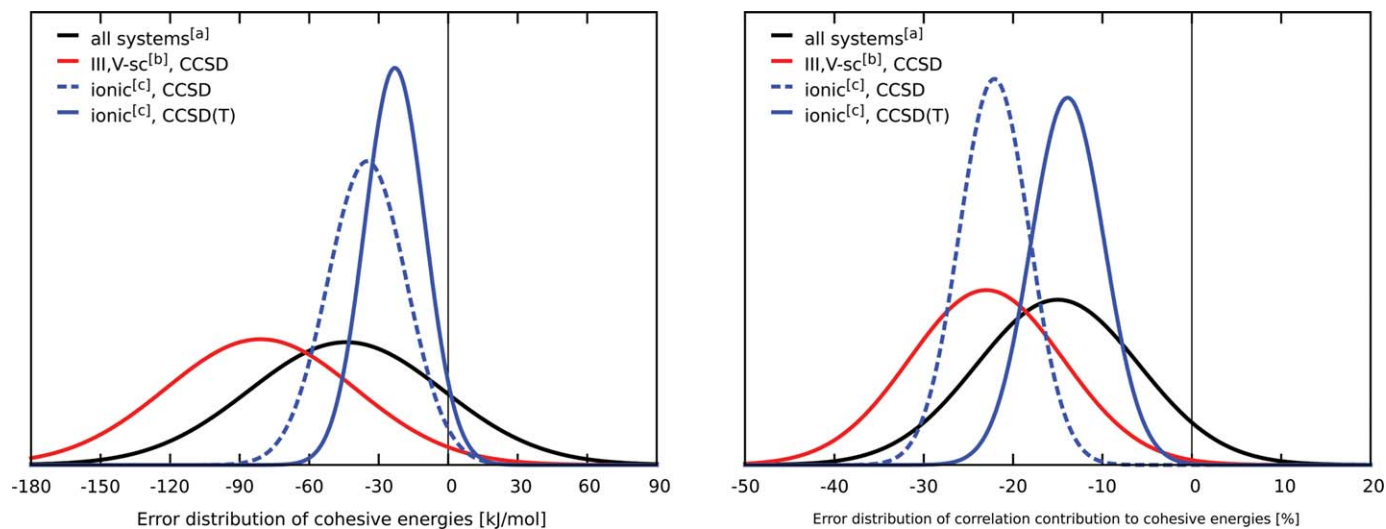
Fig. 16 Errors of the calculated cohesive energies (left) as well as the correlation contributions (right) with respect to the experimental values. (a) Results for all systems with the best available method. (b) $X^{3+}Y^{3-}$ semiconductors (X = B, Al, Ga, In; Y = N, P, As, Sb) with covalent cluster models.[70,73] (c) Alkali halides[77,79] and earth alkali oxides[75,76] with ionic cluster models
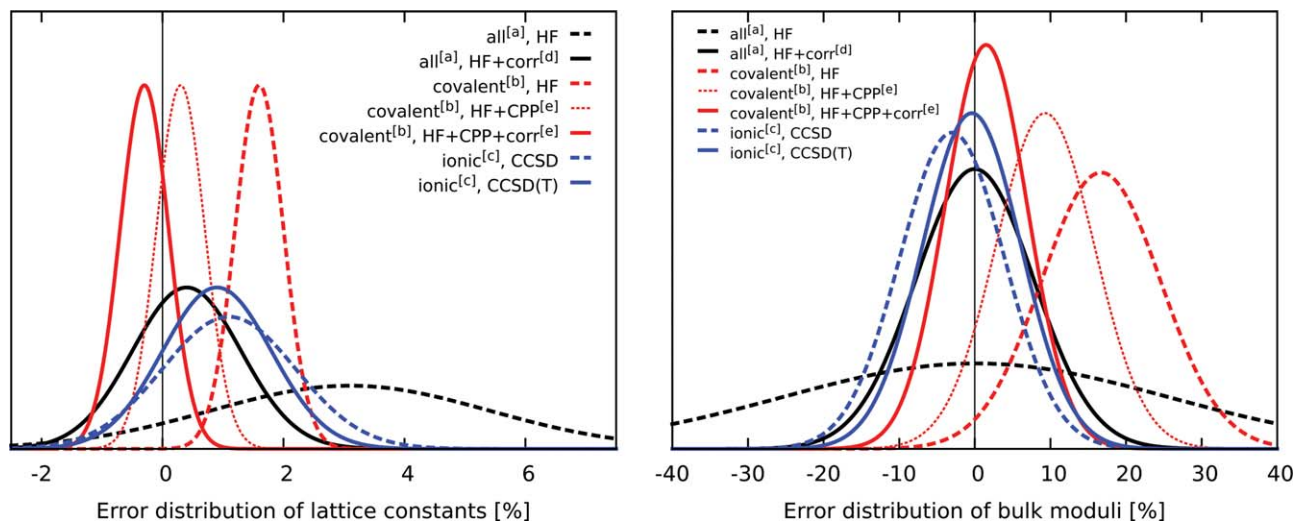
Fig. 17 Errors of the calculated lattice constants (left) as well as bulk moduli (right) with respect to the experimental values. (a) Results for all systems. (b) Systems with covalent cluster models (Si, Ge, α-Sn,[69] III,V-semiconductors[72]). (c) Alkali halides[77,79] and earth alkali oxides[76] (only for lattice constants). (d) Correlation with the best available method. (e) Calence correlation (corr) with CEPA-0 or CCSD, CPP = core polarization potential.

core polarization potentials (CPPs) has been demonstrated, as their introduction in comparison to the pure HF calculations has a larger effect than the correlation contributions for the valence electrons.[72] Due to the valence correlation the experimental lattice constants are even underestimated by nearly the same amount as their overestimated by a HF + CPP treatment.

A third frequently computed ground-state property is the bulk modulus B, which is defined as

$$B = -V\frac{\partial p}{\partial V} \tag{47}$$

and therefore measures the response of a material to uniform compression. Thus, the bulk modulus is a mechanical property with its inverse corresponding to the compressibility. Eqn (47) is equivalent to

$$B = -V\frac{\partial^2 E}{\partial V^2} \tag{48}$$

and can be specified to

$$B = \left(\frac{4}{9a}\frac{\partial^2}{\partial a^2} - \frac{8}{9a^2}\frac{\partial}{\partial a}\right)E(a) \tag{49}$$

for a cubic crystal with $V = a^3$ by change of variables.[69] The calculation of the bulk modulus might be possible at the calculated lattice constant with the benefit of first derivative $\partial E/\partial a$, being equal to 0. However, for an adequate comparison to experimental values, B should rather be evaluated at the experimental lattice constant. Considering the calculation of B by the method of increments, the latter has been done in general. Since the potential curve $E(a)$ has already been fitted within the optimization of a, B can be easily determined by analytical derivation ($\partial E/\partial a$ and $\partial^2 E/\partial a^2$) and inserting the experimental lattice parameter. The calculated results for different bulk systems relative to the experimental values are shown in Fig. 17 (right side) as an error distribution. In general, the percental variance is clearly larger than for the lattice constants, especially at the HF level. Both, HF and HF + corr provide a relatively small mean error, but the correlation contribution lowers the variance significantly. Considering the subsets, the accuracy of the covalent one at the correlated level is now slightly larger than for the ionic subset. Again, CCSD(T) is only a bit more accurate than CCSD for the ionic systems and the CPPs have a decisive influence, as shown for the covalent subset. However, in case of the bulk moduli, valence correlations affect the errors more than for the lattice constants.

4.5 Surface adsorption

Solid surfaces are of special interest, *e.g.* for hetergeneous catalysis. An essential step of such a catalysis is the adsorption of a single molecule on the crystal surface. The method of increments has been used several times to compute such surface energies on a high-accuracy level of

theory. Instead of evaluating the absolute correlation energy, the adsorption correlation energy $E_{\text{ads}}^{\text{corr}}$ can be expanded directly by means of adsorption energy increments η as[90,219]

$$E_{\text{ads}}^{\text{corr}} = \eta_M + \sum_i (\eta_i + \eta_{M,i}) + \sum_{i,j} (\eta_{i,j} + \eta_{M,i,j}) + \cdots \tag{50}$$

with the increments, coming only from the surface,

$$\eta_{i,\ldots} = \Delta\varepsilon_{i,\ldots}(\text{MS}) - \Delta\varepsilon_{i,\ldots}(\text{S}) \tag{51}$$

and those increments, containing the molecule,

$$\eta_M = \Delta\varepsilon_M(\text{MS}) - \Delta\varepsilon_M(\text{M}) \quad \eta_{M,i,\ldots} = \Delta\varepsilon_{M,i,\ldots}(\text{MS}) \tag{52}$$

In the above two equations $i,\ldots$ means, that the equation holds not only for one considered domain i, but also for two (i,j) and more. Furthermore, the letters in parentheses (MS, M, S) denote the system, for which the increment is calculated, *i.e.* M and S refer to the free molecule respectively surface. To avoid a BSSE it is recommended to use a CP correction for the free molecule (surface) by introducing the respective ghost basis functions at the atomic positions of the surface (molecule). Furthermore, the intramonomer increments η_M and $\eta_{i,\ldots}$ are relative ones, as they occur in both the MS and the free structure. The corresponding correlation contributions $\Delta\varepsilon(\text{MS})$ and $\Delta\varepsilon(\text{M})$ respectively $\Delta\varepsilon(\text{S})$ are in general very similar for a physisorption and therefore these adsorption increments are small. The main contribution to the adsorption correlation energy always originated from the attractive two-body intermonomer (*i.e.* mixed adsorbate-surface) increments $\eta_{M,i,\ldots}$. There are no respective correlation contributions in the free structures and these adsorption increments are therefore no relative ones. Furthermore, they describe the correlation part of interactions between surface and molecule, which are obviously the predominant share in a physisorption energy. In the incremental expansion of adsorption energies, domains up to the third order were mostly regarded, but in some cases the three-body increments were even neglected due to their small size.

The same equations to obtain an adsorption energy are also valid for other relative energies, *e.g.* for cohesive energies with atoms as educts.

The surface is modelled *via* similar cluster approaches as described in Section 4.2, *i.e.* the one-body domains may be atoms, ions or even bonds. The method of increments has been applied to various adsorption processes in the last few years, which are presented together with the obtained adsorption energies (HF and correlation parts) in Fig. 19. By varying the distance between the surface adsorption center and the adsorbing atom of the single molecule, potential energy curves with respect to this distance were obtained for some systems, yielding equilibrium distances and adsorption energies. The other geometry parameters were left unrelaxed, *i.e.* the same as in the isolated structures. It has been shown and is also well-known, that a physisorption affects the geometries of surfaces and adsorbate only very slightly. By
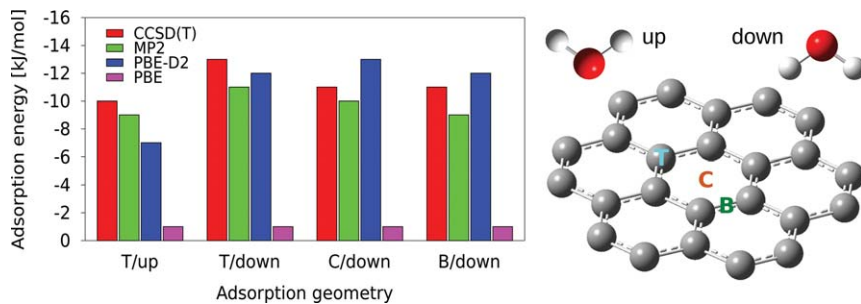
Fig. 18 Adsorption energies for different adsorbate orientations (up, down) and adsorption centers (B, C, T) of the systems H$_2$O@graphene, obtained with various methods.[94]

means of the adsorption energies it can be investigated, which crystallographic surface, molecule orientation or adsorption center is preferred. Examples are the CO adsorption on CeO$_2$ (110)[89] or CeO$_2$ (111),[90] the adsorbing of the N- or O-atom of N$_2$O at CeO$_2$ (111)[92] as well as the adsorption center for a water molecule at a graphene surface,[94] which may be the center of a ring, the center of a bond or a C atom. Results for the latter system are presented in Fig. 18. First, the hydrogen-down orientation of the water molecule and the adsorption at a C atom are preferred. Second, another application of the incremental CCSD(T) adsorption energies is to benchmark other methods. Within the WFT methods, MP2 provides results comparable to CCSD(T). Considering DFT methods, a sufficient accuracy was also obtained, when using a dispersion correction.

The adsorption has also been studied for the open-shell system Cu@ZnO.[85] Here, three approaches to treat the single electron at the Cu atom were considered. In the isolated case, where the singly-occupied orbital forms its own one-body domain, the number of increments is too large. The preferred delocalized scheme is similar to the open-shell incremental method for molecules, used by Friedrich et al., which assignes the singly-occupied orbital to every one-body domain. In another approach the single electron has been localized at the Cu atom, but still the delocalized scheme was preferred.

In some works of de Lara-Castells and co-workers the dispersionless density functional (dlDF)[220–222] was first benchmarked by[96] and later combined with the incremental CCSD(T) method.[97–99] It was found, that the intramonomer contributions (i.e. one-body, two-body surface–surface increments), which are less decisive due to their small size, can be obtained sufficiently accurate from the dlDF adsorption energy $E_{\text{ads}}(\text{dlDF})$[96] by

$$E_{\text{ads}}^{\text{intra–corr}} = E_{\text{ads}}(\text{dlDF}) - E_{\text{ads}}(HF) \tag{53}$$

The dlDF method was applied within a cluster model as well as a periodic approach. In contrast, the large intermonomer (mixed adsorbate-surface) contributions are still obtained via the incremental method at the CCSD(T) level for rather small cluster models. In order to use a larger model, the dlDF + incremental D_{as}^* method has been utilized.[97] In this approach, the intramonomer correlation is described as in eqn (53),

while the intermonomer correlation is obtained from the effective pair-wise interatomic functional

$$D_{as} = - \sum_M \sum_S \sum_n \frac{\sqrt{C_n^M C_n^S}}{R_{M-S}^n} f_n \left(\sqrt{\beta_M \beta_S} R_{M-S}^n \right) \tag{54}$$

Here, M denotes the atoms of the adsorbate and S a sufficiently large number of surface atoms to reach the required accuracy. Furthermore, R_{M-S} refers to the distance of M and S and f_n to the damping functions of Tang and Toennies.[223] The parameters C_n and β are optimized by fitting the intermonomer correlation energies for different atomic distances, obtained with the incremental CCSD(T) method for a smaller cluster. Due to the established transferability of the increments, the fitted D_{as}^* functional can be extended to a larger cluster model. For the system He@TiO$_2$ (110), dlDF + D_{as} has been found to feature an accuracy close to that of CCSD(T) or CCSD(T)-F12b.[96] The incrementally fitted D_{as}^* functional has been further applied to receive adsorption energies for several systems X@graphene (with X = He, Ag$_2$, Ne, Ar, Kr), which are also represented in Fig. 19.[97,99]

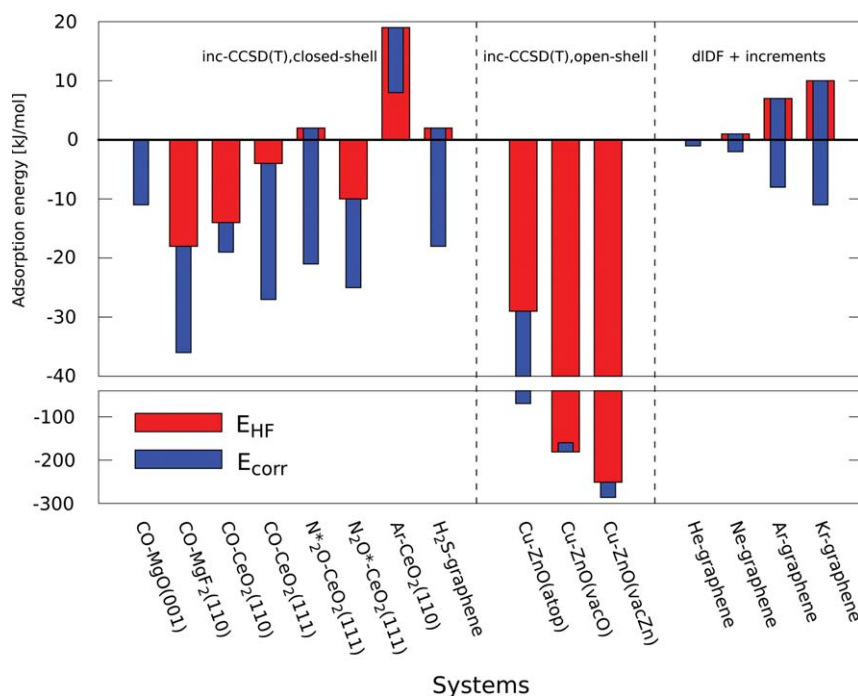**Fig. 19** Adsorption energies (HF and correlation part) for various adsorbate-surface systems. Closed-shell:[89–93,95] * marks the adsorbing atom of N$_2$O Open-shell:[85] adsorption atop an O (atop), at O vacancy site (vacO), at Zn vacancy site (vacZn) dlDF + increments:[97,99] intermonomer correlation with incremental CCSD(T) (He-graphene) or D_{as}^*, based on incremental CCSD(T) (other systems)

Finally, by means of the method of increments not only adsorption, but also surface energies E_{surf} have been computed as[224]

$$E_{\text{surf}} = \frac{1}{2A}\left(E_{\text{slab}} - n \cdot E_{\text{bulk}}\right) \tag{55}$$

Here E_{slab} is the energy of an n-layer slab and E_{bulk} refers to the energy of a unit cell in the bulk. Furthermore, A is the surface of the slab unit. With this approach surface energies for the MgF_2 (110)[224] and the $Mg(0001)$[225] surfaces were studied.

5 Conclusions

Since its "birth" in 1992 the incremental approach[54] has been implemented in different ways and for various systems, as demonstrated in this review. However, the general ansatz of the incremental methods is always same: the total correlation energy is obtained as a sum of non-negligible correlation energy increments for different domains, *i.e.* groups of orbitals to be correlated, according to a many-body expansion equation (see eqn (2), (4), (6) and (7)). During the development processes of the various implementations, many influential parameters have been investigated and improved regarding to accuracy as well as efficiency. The most important are:

- Domain partitioning and sizes:

Different ways to obtain the domains by hand or an automated algorithm were presented in this review. However, using smaller domains (for the same truncation order of the incremental series) provokes larger errors due to the neglect of more correlations. In contrast, the efficiency for large systems, which are the main focus of incremental methods, is increased. Therefore, the choice of well-balanced domain sizes is essential.

- Convergence and truncation of the incremental series:

The pure summation of the correlation energies of the disjoint domains (*i.e.* a first-order incremental expansion) is not enough, as there are non-additive terms, arising from the correlation between electrons from different one-body domains. However, these corrections to the correlation energy decrease with increasing order, *i.e.* number of simultaneously correlated domains. The incremental series was found to converge fast for appropriate domain sizes and usually a third-order expansion is sufficient. A proper convergence of the incremental series was also found for other observables, *i.e.* molecular properties like the dipole moment[110,112] or polarizabilities.[111]

- Screening off negligible increments by distance:

Due to the locality of electron correlation, it decreases fast with respect to the orbital distances. Therefore, not all increments need to be taken into account, especially for the much smaller higher-order increments, and a truncation by a distance criterion can be utilized. In this context,

order-dependent distance truncation functions as well as the nearest-neighbour approach have been established.

- Treatment of the domain surroundings:

Within an incremental calculation only the orbitals in a specific (one-, two-, three-body,...) domain are correlated. However, the domain environment cannot be neglected completely and its treatment may be crucial. For cluster models to simulate solid structures different embeddings were investigated, *e.g.* point charges, pseudopotentials or HF orbitals. The latter "embedding" is also used within the molecular incremental scheme. The main region of the domain should be described with a large target basis set to receive a high accuracy, but a significantly smaller basis set, *i.e.* a truncated virtual space, is sufficient for the environment. Thus, different approaches have been utilized *e.g.* the PAO technique[53] or dual-basis set implementations like the DSBS[127] and the db-B0 approach.[120]

However, today's incremental methods can treat various structures and problems, *e.g.* molecular and periodic systems, closed- and open-shell cases as well as energies and several other properties, all on high levels of correlation methods. Multi-reference treatments have also been performed, but are more critical as there were no implementations with a method, providing strict size-extensivity, which is required within the incremental method.

However, incremental energy calculations have been established at the CCSD(T) level of theory, which is presumed as the "gold standard" in *ab initio* quantum chemistry. This is a unique feature of the method of increments for periodic systems, while the molecular incremental scheme can successfully compete with other established local correlation methods and partially even outperform them. This has been shown in several cases in terms of a very high accuracy together with a very good efficiency, especially for systems with increasing sizes. Thus, the local errors were found to be negligible, when using an appropriate optimized setting of the parameters. Another essential benefit of the incremental methods is the possibility of a complete parallelization of the incremental calculations, which can reduce the wall time to a small percentage of the total CPU time. Therefore, calculations for larger systems with clearly larger basis sets compared to standard CCSD(T) can be performed, while the local errors are kept negligible. That way a higher total accuracy is received, providing useful benchmarks on a very sophisticated level of theory. These may then be applied to improve other quantum chemical methods, like DFT, which is the leading method to model chemical reactions in the field of computational chemistry.

Acknowledgements

References

1   A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, Mineola, New York, 1989.

2   F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Ltd, Chichester, England, 2007.

3   T. Helgaker, P. Jørgensen and J. Olsen, in *Molecular Electronic-Structure Theory*, John Wiley & Sons, Ltd, 2000, ch. The Standard Models, pp. 142–200.

4   T. Helgaker, P. Jørgensen and J. Olsen, in *Molecular Electronic-Structure Theory*, John Wiley & Sons, Ltd, 2000, ch. Calibration of the Electronic-Structure Models, pp. 817–883.

5   K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, *Chem. Phys. Lett.*, 1989, **157**, 479–483.

6   P. Pulay, *Chem. Phys. Lett.*, 1983, **100**, 151–154.

7   S. Saebø and P. Pulay, *Chem. Phys. Lett.*, 1985, **113**, 13–18.

8   P. Pulay and S. Saebø, *Theor. Chim. Acta*, 1986, **69**, 357–368.

9   S. Saebø and P. Pulay, *J. Chem. Phys.*, 1987, **86**, 914–922.

10 S. Saebø and P. Pulay, *Annu. Rev. Phys. Chem.*, 1993, **44**, 213–236.

11 C. Hampel and H.-J. Werner, *J. Chem. Phys.*, 1996, **104**, 6286–6297.

12 M. Schütz, G. Hetzer and H.-J. Werner, *J. Chem. Phys.*, 1999, **111**, 5691–5705.

13 M. Schütz, *J. Chem. Phys.*, 2000, **113**, 9986–10001.

14 M. Schütz and H.-J. Werner, *J. Chem. Phys.*, 2001, **114**, 661–681.

15 H.-J. Werner, *J. Chem. Phys.*, 2008, **129**, 101103.

16 T. B. Adler and H.-J. Werner, *J. Chem. Phys.*, 2009, **130**, 241101.

17 H.-J. Werner and M. Schütz, *J. Chem. Phys.*, 2011, **135**, 144116.

18 T. B. Adler and H.-J. Werner, *J. Chem. Phys.*, 2011, **135**, 144117.

19 F. Neese, F. Wennmohs and A. Hansen, *J. Chem. Phys.*, 2009, **130**, 114108.

20 F. Neese, A. Hansen and D. G. Liakos, *J. Chem. Phys.*, 2009, **131**, 064103.

21 D. G. Liakos, A. Hansen and F. Neese, *J. Chem. Theory Comput.*, 2011, **7**, 76–87.

22 A. Hansen, D. G. Liakos and F. Neese, *J. Chem. Phys.*, 2011, **135**, 214102.

23 D. G. Liakos and F. Neese, *J. Phys. Chem. A*, 2012, **116**, 4801–4816.

24 F. Pavosevic, F. Neese and E. F. Valeev, *J. Chem. Phys.*, 2014, **141**, 054106.

25 C. Riplinger and F. Neese, *J. Chem. Phys.*, 2013, **138**, 034106.

26 C. Riplinger, B. Sandhoefer, A. Hansen and F. Neese, *J. Chem. Phys.*, 2013, **139**, 134101.

27 D. G. Liakos, M. Sparta, M. K. Kesharwani, J. M. L. Martin and F. Neese, *J. Chem. Theory Comput.*, 2015, **11**, 1525–1539.

28 J. Yang, Y. Kurashige, F. R. Manby and G. K. L. Chan, *J. Chem. Phys.*, 2011, **134**, 044123.

29 J. Yang, G. K. L. Chan, F. R. Manby, M. Schütz and H.-J. Werner, *J. Chem. Phys.*, 2012, **136**, 144105.

30 W. Förner, J. Ladik, P. Otto and J. Cizek, *Chem. Phys.*, 1985, **97**, 251–262.

31 S. Li, J. Shen, W. Li and Y. Jiang, *J. Chem. Phys.*, 2006, **125**, 074109.

32 W. Li, P. Piecuch, J. R. Gour and S. Li, *J. Chem. Phys.*, 2009, **131**, 114109.

33 W. Li and P. Piecuch, *J. Phys. Chem. A*, 2010, **114**, 6721–6727.

34 W. Li, Y. Guo and S. Li, *Phys. Chem. Chem. Phys.*, 2012, **14**, 7854–7862.

35 Z. Rolik and M. Kállay, *J. Chem. Phys.*, 2012, **135**, 104111.

36 Z. Rolik, L. Szegedy, I. Ladjánski, B. Ladóczki and M. Kállay, *J. Chem. Phys.*, 2013, **139**, 094105.

37 N. Flocke and R. J. Bartlett, *J. Chem. Phys.*, 2004, **121**, 10935–10944.

38 T. F. Hughes, N. Flocke and R. J. Bartlett, *J. Phys. Chem. A*, 2008, **112**, 5994–6003.

39   T. F. Hughes and R. J. Bartlett, *J. Chem. Phys.*, 2008, **129**, 054105.
40   W. Yang, *Phys. Rev. Lett.*, 1991, **66**, 1438–1441.
41   M. Kobayashi and H. Nakai, *J. Chem. Phys.*, 2009, **131**, 114108.
42   M. Kobayashi and H. Nakai, *Phys. Chem. Chem. Phys.*, 2012, **14**, 7629–7639.
43   D. G. Fedorov and K. Kitaura, *J. Chem. Phys.*, 2004, **121**, 2483–2490.
44   D. G. Fedorov and K. Kitaura, *J. Chem. Phys.*, 2005, **123**, 134103.
45   D. G. Fedorov, T. Nagata and K. Kitaura, *Phys. Chem. Chem. Phys.*, 2012, **14**, 7562–7577.
46   J. E. Subotnik and M. Head-Gordon, *J. Chem. Phys.*, 2005, **123**, 064108.
47   J. E. Subotnik, A. Sodt and M. Head-Gordon, *J. Chem. Phys.*, 2006, **125**, 074116.
48   J. E. Subotnik, A. Sodt and M. Head-Gordon, *J. Chem. Phys.*, 2008, **128**, 034103.
49   M. Ziółkowski, B. Jansk, T. Kjærgaard and P. Jørgensen, *J. Chem. Phys.*, 2010, **133**, 014107.
50   K. Kristensen, M. Ziółkowski, B. Jansík, T. Kjærgaard and P. Jørgensen, *J. Chem. Theory Comput.*, 2011, **7**, 1677–1694.
51   I.-M. Høyvik, K. Kristensen, B. Jansík and P. Jørgensen, *J. Chem. Phys.*, 2012, **136**, 014105.
52   H. W. Qi, H. R. Leverentz and D. G. Truhlar, *J. Phys. Chem. A*, 2013, **117**, 4486–4499.
53   J. Friedrich, M. Hanrath and M. Dolg, *J. Chem. Phys.*, 2007, **126**, 154110.
54   H. Stoll, *Phys. Rev. B*, 1992, **46**, 6700–6704.
55   R. K. Nesbet, *Phys. Rev.*, 1967, **155**, 51–55.
56   R. K. Nesbet, *Phys. Rev.*, 1967, **155**, 56–58.
57   R. K. Nesbet, *Phys. Rev.*, 1968, **175**, 2–9.
58   R. K. Nesbet, *Adv. Chem. Phys.*, 1969, **14**, 1–34.
59   R. K. Nesbet, *Phys. Rev. A*, 1971, **3**, 87–94.
60   H. A. Bethe and J. Goldstone, *Proc. R. Soc. London, Ser. A*, 1957, **238**, 551–567.
61   K. A. Brueckner and C. A. Levinson, *Phys. Rev.*, 1955, **97**, 1344–1352.
62   K. A. Brueckner, *Phys. Rev.*, 1955, **97**, 1353–1366.
63   K. A. Brueckner, *Phys. Rev.*, 1955, **100**, 36–45.
64   K. A. Brueckner and W. Wada, *Phys. Rev.*, 1956, **103**, 1008–1016.
65   L. D. Faddeev, *Sov. Phys. J.Exptl. Theoret. Phys.*, 1961, **12**, 1014–1019.
66   L. D. Faddeev and S. P. Merkuriev, *Quantum Scattering Theory for Several Particle Systems*, Kluwer Academic Publishing, Dordrecht, The Netherlands, 1993.
67   H. Stoll, *Chem. Phys. Lett.*, 1992, **191**, 548–552.
68   H. Stoll, *J. Chem. Phys.*, 1992, **97**, 8449–8454.
69   B. Paulus, P. Fulde and H. Stoll, *Phys. Rev. B*, 1995, **51**, 10572–10578.
70   B. Paulus, P. Fulde and H. Stoll, *Phys. Rev. B*, 1996, **54**, 2556–2560.
71   M. Albrecht, B. Paulus and H. Stoll, *Phys. Rev. B*, 1997, **56**, 7339–7347.
72   S. Kalvoda, B. Paulus, P. Fulde and H. Stoll, *Phys. Rev. B*, 1997, **55**, 4027–4030.
73   B. Paulus, F.-J. Shi and H. Stoll, *J. Phys.: Condens. Matter*, 1997, **9**, 2745–2758.
74   K. Rosciszewski, K. Doll, B. Paulus, P. Fulde and H. Stoll, *Phys. Rev. B*, 1998, **57**, 14667–14672.
75   K. Doll, M. Dolg, P. Fulde and H. Stoll, *Phys. Rev. B*, 1995, **52**, 4842–4848.
76   K. Doll, M. Dolg and H. Stoll, *Phys. Rev. B*, 1996, **54**, 13529–13535.
77   K. Doll and H. Stoll, *Phys. Rev. B*, 1997, **56**, 10121–10127.
78   K. Doll, P. Pyykkö and H. Stoll, *J. Chem. Phys.*, 1998, **109**, 2339–2345.
79   K. Doll and H. Stoll, *Phys. Rev. B*, 1998, **57**, 4327–4331.

80 B. Paulus, K. Rosciszewski, N. Gaston, P. Schwerdtfeger and H. Stoll, *Phys. Rev. B*, 2004, **70**, 165106.

81 E. Voloshina and B. Paulus, *Phys. Rev. B*, 2007, **75**, 245117.

82 N. Gaston and B. Paulus, *Phys. Rev. B*, 2007, **76**, 214116.

83 E. Voloshina, B. Paulus and H. Stoll, *J. Phys.: Conf. Ser.*, 2008, **117**, 012029.

84 M. Mödl, M. Dolg, P. Fulde and H. Stoll, *J. Chem. Phys.*, 1997, **106**, 1836–1846.

85 I. Schmitt, K. Fink and V. Staemmler, *Phys. Chem. Chem. Phys.*, 2009, **11**, 11196–11206.

86 H. Stoll, B. Paulus and P. Fulde, *Chem. Phys. Lett.*, 2009, **469**, 90–93.

87 E. Voloshina and B. Paulus, *J. Chem. Theory Comput.*, 2014, **10**, 1698–1706.

88 E. Fertitta, B. Paulus, G. Barcza and O. Legeza, *J. Chem. Phys.*, 2015, **143**, 114108.

89 C. Müller, B. Herschend, K. Hermansson and B. Paulus, *J. Chem. Phys.*, 2008, **128**, 214701.

90 C. Müller, B. Paulus and K. Hermansson, *Surf. Sci.*, 2009, **603**, 2619–2623.

91 B. Paulus and K. Rosciszewski, *Int. J. Quantum Chem.*, 2009, **109**, 3055–3062.

92 C. Müller, K. Hermansson and B. Paulus, *Chem. Phys.*, 2009, **362**, 91–96.

93 V. Staemmler, *J. Phys. Chem. A*, 2011, **115**, 7153–7160.

94 E. Voloshina, D. Usvyat, M. Schütz, Y. Dedkov and B. Paulus, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12041–12047.

95 L. Hammerschmidt, C. Müller and B. Paulus, *J. Chem. Phys.*, 2012, **136**, 124117.

96 M. P. de Lara-Castells, H. Stoll and A. O. Mitrushchenkov, *J. Phys. Chem. A*, 2014, **118**, 6367–6384.

97 M. P. de Lara-Castells, H. Stoll, B. Civalleri, M. Causa, E. Voloshina, A. O. Mitrushchenkov and M. Pi, *J. Chem. Phys.*, 2014, **141**, 151102.

98 M. P. de Lara-Castells, A. O. Mitrushchenkov and H. Stoll, *J. Chem. Phys.*, 2015, **143**, 102804.

99 M. P. de Lara-Castells, M. Bartolomei, A. O. Mitrushchenkov and H. Stoll, *J. Chem. Phys.*, 2015, **143**, 194701.

100 K. Rosciszewski, B. Paulus, P. Fulde and H. Stoll, *Phys. Rev. B*, 1999, **60**, 7905–7910.

101 C. Buth and B. Paulus, *Chem. Phys. Lett.*, 2004, **398**, 44–49.

102 C. Buth and B. Paulus, *Phys. Rev. B*, 2006, **74**, 045122.

103 C. Müller, D. Usvyat and H. Stoll, *Phys. Rev. B*, 2011, **83**, 245136.

104 C. Müller and D. Usvyat, *J. Chem. Theory Comput.*, 2013, **9**, 5590–5598.

105 K. G. Steenbergen, N. Gaston, C. Müller and B. Paulus, *J. Chem. Phys.*, 2014, **141**, 124707.

106 J. Friedrich, M. Hanrath and M. Dolg, *J. Phys. Chem. A*, 2008, **112**, 8762–8766.

107 T. Anacker, D. P. Tew and J. Friedrich, *J. Chem. Theory Comput.*, 2016, **12**, 65–78.

108 J. Zhang and M. Dolg, *J. Chem. Theory Comput.*, 2015, **11**, 962–968.

109 J. Yang and M. Dolg, *J. Chem. Phys.*, 2007, **127**, 084108.

110 J. Friedrich, S. Coriani, T. Helgaker and M. Dolg, *J. Chem. Phys.*, 2009, **131**, 154102.

111 J. Friedrich, H. R. McAlexander, A. Kumar and T. D. Crawford, *Phys. Chem. Chem. Phys.*, 2015, **17**, 14284–14296.

112 B. Fiedler, S. Coriani and J. Friedrich, *J. Chem. Theory Comput.*, 2016, **12**, 3040–3052.

113 J. Friedrich and M. Dolg, *J. Chem. Theory Comput.*, 2009, **5**, 287–294.

114 J. M. Foster and S. F. Boys, *Rev. Mod. Phys.*, 1960, **32**, 300–302.

115 J. Pipek and P. G. Mezey, *J. Chem. Phys.*, 1989, **90**, 4916–4926.

116 G. Karypis and V. Kumar, *SIAM J. Sci. Comput.*, 1998, **20**, 359–392.

117 J. Friedrich and J. Hänchen, *J. Chem. Theory Comput.*, 2013, **9**, 5381–5394.

118 E. Forgy, *Biometrics*, 1965, **21**, 768–769.

119 J. MacQueen, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, Berkeley, CA, 1967, pp. 281–297.

120 J. Zhang and M. Dolg, *J. Chem. Theory Comput.*, 2013, **9**, 2992–3003.

121 J. Friedrich, M. Hanrath and M. Dolg, *J. Phys. Chem. A*, 2007, **111**, 9830–9837.

122 J. Friedrich, M. Hanrath and M. Dolg, *Chem. Phys.*, 2008, **346**, 266–274.

123 J. Friedrich, K. Walczak and M. Dolg, *Chem. Phys.*, 2009, **356**, 47–53.

124 K. Walczak, J. Friedrich and M. Dolg, *Chem. Phys.*, 2010, **376**, 36–45.

125 R. Mata and H. Stoll, *Chem. Phys. Lett.*, 2008, **465**, 136–141.

126 J. W. Boughton and P. Pulay, *J. Comput. Chem.*, 1993, **14**, 736–740.

127 J. Friedrich and M. Dolg, *J. Chem. Phys.*, 2008, **129**, 244105.

128 J. Friedrich, *J. Chem. Theory Comput.*, 2012, **8**, 1597–1607.

129 J. Friedrich, *J. Chem. Theory Comput.*, 2010, **6**, 1834–1842.

130 T. Anacker, J. G. Hill and J. Friedrich, *J. Phys. Chem. A*, 2016, **120**, 2443–2458.

131 D. G. Liakos and F. Neese, *J. Chem. Theory Comput.*, 2015, **11**, 4054–4063.

132 T. Helgaker, P. Jørgensen and J. Olsen, in *Molecular Electronic-Structure Theory*, John Wiley & Sons, Ltd, 2000, ch. Atomic Basis Functions, pp. 201–255.

133 W. Kutzelnigg, *Theor. Chim. Acta*, 1985, **68**, 445–469.

134 S. J. Zhong, E. C. Barnes and G. A. Petersson, *J. Chem. Phys.*, 2008, **129**, 184116.

135 A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, *Chem. Phys. Lett.*, 1998, **286**, 243–252.

136 W. Klopper and W. Kutzelnigg, *Chem. Phys. Lett.*, 1987, **134**, 17–22.

137 W. Kutzelnigg and W. Klopper, *J. Chem. Phys.*, 1991, **94**, 1985–2001.

138 J. Noga, W. Kutzelnigg and W. Klopper, *Chem. Phys. Lett.*, 1992, **199**, 497–504.

139 J. Noga and W. Kutzelnigg, *J. Chem. Phys.*, 1994, **101**, 7738–7762.

140 S. Ten-no, *Chem. Phys. Lett.*, 2004, **398**, 56–61.

141 D. P. Tew and W. Klopper, *J. Chem. Phys.*, 2005, **123**, 074101.

142 S. Ten-no, *J. Chem. Phys.*, 2004, **121**, 117–129.

143 H. Fliegl, W. M. Klopper and C. Hättig, *J. Chem. Phys.*, 2005, **122**, 084107.

144 D. P. Tew, W. Klopper, C. Neiss and C. Hättig, *Phys. Chem. Chem. Phys.*, 2007, **9**, 1921–1930.

145 C. Hättig, D. P. Tew and A. Köhn, *J. Chem. Phys.*, 2010, **132**, 231102.

146 T. B. Adler, G. Knizia and H.-J. Werner, *J. Chem. Phys.*, 2007, **127**, 221106.

147 G. Knizia, T. B. Adler and H.-J. Werner, *J. Chem. Phys.*, 2009, **130**, 054104.

148 E. F. Valeev, *Chem. Phys. Lett.*, 2004, **395**, 190–195.

149 K. E. Yousaf and K. Peterson, *J. Chem. Phys.*, 2008, **129**, 184108.

150 K. E. Yousaf and K. Peterson, *Chem. Phys. Lett.*, 2009, **476**, 303–307.

151 F. Weigend, A. Köhn and C. Hättig, *J. Chem. Phys.*, 2002, **116**, 3175–3183.

152 C. Hättig, *Phys. Chem. Chem. Phys.*, 2005, **7**, 59–66.

153 W. Klopper, F. R. Manby, S. Ten-No and E. F. Valeev, *Int. Rev. Phys. Chem.*, 2006, **25**, 427–468.

154 L. Kong, F. A. Bischoff and E. F. Valeev, *Chem. Rev.*, 2012, **112**, 75–107.

155 J. Friedrich, D. P. Tew, W. Klopper and M. Dolg, *J. Chem. Phys.*, 2010, **132**, 164114.

156 J. Friedrich, E. Perlt, M. Roatsch, C. Spickermann and B. Kirchner, *J. Chem. Theory Comput.*, 2011, **7**, 843–851.

157 J. Friedrich and K. Walczak, *J. Chem. Theory Comput.*, 2013, **9**, 408–417.

158 J. Zhang and M. Dolg, *J. Chem. Phys.*, 2014, **140**, 044114.

159 K. D. Vogiatzis, W. Klopper and J. Friedrich, *J. Chem. Theory Comput.*, 2015, **11**, 1574–1584.

160 M. Schmittel, M. Strittmatter and S. Kiau, *Tetrahedron Lett.*, 1995, **36**, 4975–4978.

161 G. Knizia and H.-J. Werner, *J. Chem. Phys.*, 2008, **128**, 154103.

162 K. Peterson, M. K. Kesharwani and J. M. L. Martin, *Mol. Phys.*, 2015, **113**, 1551–1558.

163 O. Marchetti and H.-J. Werner, *Phys. Chem. Chem. Phys.*, 2008, **10**, 3400–3409.

164 D. W. Schwenke, *J. Chem. Phys.*, 2005, **122**, 014107.

165 J. G. Hill, K. A. Peterson, G. Knizia and H.-J. Werner, *J. Chem. Phys.*, 2009, **131**, 194105.

166 J. Friedrich and B. Fiedler, *Chem. Phys.*, 2016, **472**, 72–80.

167 S. Tsuzuki, T. Uchimaru, M. Mikami, K. Tanabe, T. Sako and S. Kuwajima, *Chem. Phys. Lett.*, 1996, **255**, 347–349.

168 P. Jurecka, J. Sponer, J. Cerny and P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985–1993.

169 B. Brauer, M. K. Kesharwani and J. M. L. Martin, *J. Chem. Theory Comput.*, 2014, **10**, 3791–3799.

170 J. Sponer and P. Hobza, *J. Am. Chem. Soc.*, 2002, **124**, 11802–11808.

171 E. Papajak and D. G. Truhlar, *J. Chem. Phys.*, 2012, **137**, 064110.

172 M. S. Marshall, L. A. Burns and C. D. Sherrill, *J. Chem. Phys.*, 2011, **135**, 194102.

173 T. H. Dunning Jr. and K. A. Peterson, *J. Chem. Phys.*, 2000, **113**, 7799–7808.

174 K. D. Vogiatzis and W. Klopper, *Mol. Phys.*, 2013, **111**, 2299–2305.

175 K. D. Vogiatzis, R. Haunschild and W. Klopper, *Theor. Chem. Acc.*, 2014, **133**, 1446.

176 S. Ahnen, A.-S. Hehn, K. D. Vogiatzis, M. A. Trachsel, S. Leutwyler and W. Klopper, *Chem. Phys.*, 2014, **441**, 17–22.

177 M. R. Nyden and G. A. Petersson, *J. Chem. Phys.*, 1981, **75**, 1843–1862.

178 K. D. Vogiatzis, E. C. Barnes and W. Klopper, *Chem. Phys. Lett.*, 2011, **503**, 157–161.

179 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 5656–5667.

180 J. Friedrich, M. Hanrath and M. Dolg, *Chem. Phys.*, 2007, **338**, 33–43.

181 I. Polenz, F. G. Schmidt, J. Friedrich, I. Tchernook and S. Spange, *Macromol. Chem. Phys.*, 2013, **214**, 1473–1483.

182 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.

183 K. Banert, M. Hagedorn, T. Pester, N. Siebert, C. Staude, I. Tchernook, K. Rathmann, O. Holloczki and J. Friedrich, *Chem. – Eur. J.*, 2015, **21**, 14911–14923.

184 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

185 C. Spickermann, E. Perlt, M. von Domaros, M. Roatsch, J. Friedrich and B. Kirchner, *J. Chem. Theory Comput.*, 2011, **7**, 868–875.

186 F. Weinhold, *J. Chem. Phys.*, 1998, **109**, 367–372.

187 F. Weinhold, *J. Chem. Phys.*, 1998, **109**, 373–384.

188 M. von Domaros, S. Jähnigen, J. Friedrich and B. Kirchner, *J. Chem. Phys.*, 2016, **144**, 064305.

189 L. Fiedler, H. R. Leverentz, S. Nachimuthu, J. Friedrich and D. G. Truhlar, *J. Chem. Theory Comput.*, 2014, **10**, 3129–3139.

190 J. Friedrich, H. Yu, H. R. Leverentz, P. Bai, J. I. Siepmann and D. G. Truhlar, *J. Phys. Chem. Lett.*, 2014, **5**, 666–670.

191 V. S. Bryantsev, M. S. Diallo, A. C. T. van Duin, I. William and A. Goddard, *J. Chem. Theory Comput.*, 2009, **5**, 1016–1026.

192 T. Anacker and J. Friedrich, *J. Comput. Chem.*, 2014, **35**, 634–643.

193 G. S. Fanourgakis, E. Apra and S. S. Xantheas, *J. Chem. Phys.*, 2004, **121**, 2655–2663.

194 L. Goerigk and S. Grimme, *J. Chem. Theory Comput.*, 2010, **6**, 107–126.

195 J. Zhang, N. Heinz and M. Dolg, *Inorg. Chem.*, 2014, **53**, 7700–7708.

196 N. Heinz, J. Zhang and M. Dolg, *J. Chem. Theory Comput.*, 2014, **10**, 5593–5598.

197 J. Friedrich, M. Dolg, A. Gansäuer, D. Geich-Gimbel and T. Lauterbach, *J. Am. Chem. Soc.*, 2005, **127**, 7071–7077.

198 J. Friedrich, K. Walczak, M. Dolg, F. Piestert, T. Lauterbach, D. Worgull and A. Gansäuer, *J. Am. Chem. Soc.*, 2008, **130**, 1788–1796.

199 B. Paulus, *Phys. Rep.*, 2006, **428**, 1–52.

200 M. Yu, S. Kalvoda and M. Dolg, *Chem. Phys.*, 1997, **224**, 121–131.

201 A. Abdurahman, A. Shukla and M. Dolg, *Chem. Phys.*, 2000, **257**, 301–310.

202 A. Abdurahman, A. Shukla and M. Dolg, *J. Chem. Phys.*, 2000, **112**, 4801–4805.

203 R. Fondermann, M. Dolg, M. Raab and E. Niecke, *Chem. Phys.*, 2006, **325**, 291–298.

204 B. Paulus, *Int. J. Quantum Chem.*, 2004, **100**, 1026–1032.

205 K. Doll, M. Dolg, P. Fulde and H. Stoll, *Phys. Rev. B*, 1997, **55**, 10282–10288.

206 E. Voloshina and B. Paulus, *J. Chem. Phys.*, 2006, **124**, 234711.

207 E. Voloshina and B. Paulus, *Theor. Chem. Acc.*, 2005, **114**, 259–264.

208 E. Voloshina and B. Paulus, *J. Comput. Chem.*, 2008, **29**, 2107–2112.

209 S. Kalvoda, M. Dolg, H.-J. Flad, P. Fulde and H. Stoll, *Phys. Rev. B*, 1998, **57**, 2127–2133.

210 B. Paulus, *Chem. Phys. Lett.*, 2003, **371**, 7–14.

211 B. Paulus and K. Rosciszewski, *Chem. Phys. Lett.*, 2004, **394**, 96–100.

212 E. Voloshina, N. Gaston and B. Paulus, *J. Chem. Phys.*, 2007, **126**, 134115.

213 E. Voloshina and B. Paulus, *Chemical Modelling: Applications and Theory Volume 6*, The Royal Society of Chemistry, 2009, vol. 6, pp. 162–209.

214 N. Gaston, B. Paulus, U. Wedig and M. Jansen, *Phys. Rev. Lett.*, 2008, **100**, 226404.

215 N. Gaston, D. Andrae, B. Paulus, U. Wedig and M. Jansen, *Phys. Chem. Chem. Phys.*, 2010, **12**, 681–687.

216 R. J. Gdanitz and R. Ahlrichs, *Chem. Phys. Lett.*, 1988, **143**, 413–420.

217 H.-J. Werner and P. J. Knowles, *Theor. Chim. Acta*, 1990, **78**, 175–187.

218 P. Fulde, H. Stoll and K. Kladko, *Chem. Phys. Lett.*, 1999, **299**, 481–487.

219 C. Müller and B. Paulus, *Phys. Chem. Chem. Phys.*, 2012, **14**, 7605–7614.

220 K. Pernal, R. Podeszwa, K. Patkowski and K. Szalewicz, *Phys. Rev. Lett.*, 2009, **103**, 263201.

221 R. Podeszwa, K. Pernal, K. Patkowski and K. Szalewicz, *J. Phys. Chem. Lett.*, 2010, **1**, 550–555.

222 R. Podeszwa and K. Szalewicz, *J. Chem. Phys.*, 2012, **136**, 161102.

223 K. T. Tang and J. P. Toennies, *J. Chem. Phys.*, 1984, **80**, 3726–3741.

224 L. Hammerschmidt, L. Maschio, C. Müller and B. Paulus, *J. Chem. Theory Comput.*, 2015, **11**, 252–259.

225 E. Voloshina, *Phys. Rev. B*, 2012, **85**, 045444.

Double hybrid density functional approximations

Tobias Schwabe

1 Introduction

The development of double hybrid density functional approximations (DHDFA) is one of the latest steps in an ongoing process of improving the density functional theory (DFT) method.[1,2] The ultimate goal in this process is to find an exact but feasible energy expression for the electrons of a given system. The basic idea of DHDFAs to approach this goal is the combination of exchange and correlation contributions, both, from wave function theory and DFT.[3–5] This expands the approach of standard hybrid density functional approximations (DFAs) by the inclusion of post-Hartree–Fock contributions and thereby addresses many challenges in DFT.[6]

The seminal publication on DHDFA is the one by Grimme[7] who proposed to mix DFT exchange-correlation methods with Fock exchange and a correlation contribution based on second order perturbation theory (PT2) like in wave function theory. This semi-empirical method is very successful and efficient and therefore attracted a lot of interest also for the more rigorous theoretical foundations of the approach. But it should be mentioned that other approaches existed before which made use of an amount of PT2 contribution in the energy expression either based on the Kohn–Sham DFT orbitals[8] or by combining independent DFT and MP2 computations in so-called multi-component approaches.[9]

This chapter covers the general concepts behind the DHDFA approach and introduces the different variants that can be found in literature. An overview of the improved performance with respect to the previous stage of development and the differences for the various DHDFA methods is provided. Their performance for thermochemistry and other molecular properties are discussed and for the former, some exemplary results have been computed. This broad review should give a general introduction to DHDFAs and demonstrate their achievements within DFT. But first, the general concepts in DFT are revised to provide the necessary foundations for the theory of DHDFAs.

Since Kohn and Sham (KS) reformulated the DFT, it became the most widely applied electronic structure theory in computational chemistry and materials science. The reason for this is quite pragmatic: DFT has a very favourable cost-performance ratio and can easily be applied within the two main technical approaches to solve the electronic structure problem, *i.e.* using molecular orbitals (one electron wave functions)

Center for Bioinformatics, University of Hamburg, Bundesstr. 43, 20146 Hamburg, Germany. E-mail: schwabe@zbh.uni-hamburg.de

either in the basis of atomic orbitals or in the basis of plane waves with periodic boundary conditions. One might also argue that the underlying concepts of DFT provide a different insight to the electronic structure problem which is more familiar to other known concepts in chemistry and this has laid the foundation for its success. But in practise, the analysis of DFT and wave function methods is carried out in a similar fashion and the differences are not always clearly distinguished or even recognized.

Because of its importance, several text books and reviews are available that cover the fundamentals and also specific topics of DFT and here, only the most relevant details for the understanding of the extension given by a double hybrid approach are presented, while the theoretical foundations of DHDFAs are discussed in detail in the following section.

Central to the theory of KS-DFT is the approximation to treat all electrons as independent, *i.e.* not directly interacting fermionic particles. Further, it can be shown that an external potential common to all electrons in the system exists that depends on the all-electron density and yields the exact electron density of the system under consideration when solved self-consistently (for the non-degenerate case). From that exact density knowing the exact potential the true electron energy of the system within the Born–Oppenheimer-approximation can be computed. Alternatively, the one-electron energy contributions can be computed independently and the remaining contributions are computed by using the electron density for their evaluation with an according functional. This separation is necessary, because a general functional for the kinetic energy contribution is not known.

In turn, the potential is defined to be the energy functional variation with the variation of the density. Then, the actual computational task becomes solving the one-electron-wave functions for all electrons of a given model system based on the effective eigenpair problem defined by the KS-operator:

$$\hat{f}^{\mathrm{KS}}|i\rangle = \mathrm{e}_i|i\rangle \tag{1}$$

The KS-operator can be separated into different compounds:

$$\hat{f}^{\mathrm{KS}} = \hat{v}_{\mathrm{kin}} + \hat{v}_{\mathrm{nuc}} + \hat{v}_{\mathrm{Jxc}} \tag{2}$$

where $\hat{v}_{\mathrm{kin}}$ is the operator for the kinetic energy, $\hat{v}_{\mathrm{nuc}}$ the operator for the electron-nuclei interaction and $\hat{v}_{\mathrm{Jxc}}$ the operator for the (classical) Coulomb, exchange, and correlation contribution, *i.e.* for the interaction of a charge cloud and the quantum-mechanical contributions from fermionic exchange and correlated movement.

While the theory is exact in principle, in practise the terms for the quantum-mechanical contributions have to be approximated because they are unknown for the general case. This has led to the whole field of functional development. It would be a formidable task and is out of scope for the present review to give a complete overview of density functional development. Instead, a general classification of functionals is presented.
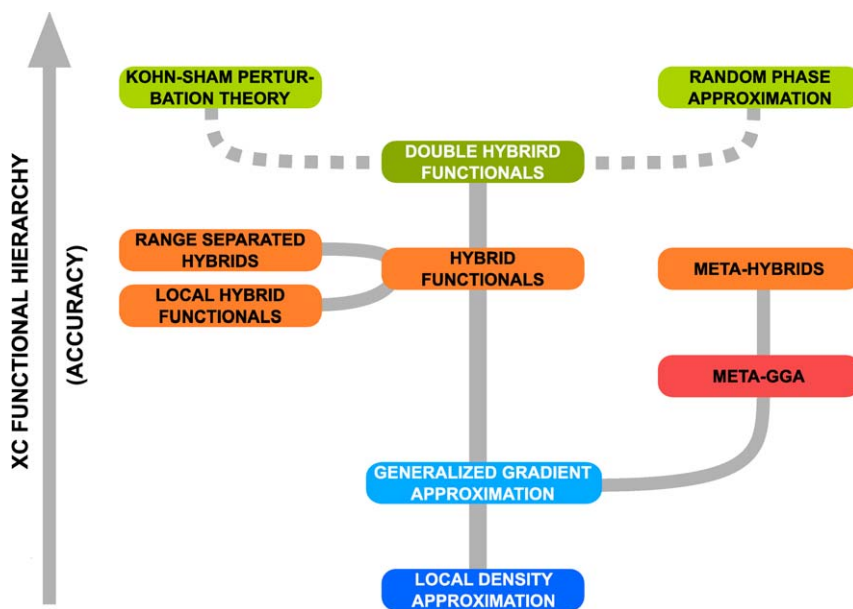
Fig. 1 The functional tree illustrates the hierarchy and improvements in DFA development.

This classification follows the general ideas derived by Perdew who introduced a concept based on the actual information used to evaluate the XC functional. While the original idea is presented as a hierarchical ladder in relation to the biblical Jacob's ladder,[10] the hierarchy depicted in Fig. (1) is given as tree-like structure to emphasize that the development actually evolved in different branches.

The foundation (roots) of the theory are given in the Hartree approach which completely neglects any quantum-mechanical contributions. In a next step, the KS approach is introduced but for the XC potential only the density at a point in space is exploited. This is known as local-density approximation (LDA). A first improvement is the consideration of the variation of the density at a point in space by its gradient at that point. This is called generalized gradient approximation (GGA). Finally, this can be taken a step further to the second derivative of the density at a point which is equivalent to consider the kinetic energy density at that point. This leads to the class of meta-GGAs.

So far, only a local potential and only occupied orbitals are taken into account. For the next class of functionals, this is extended by mixing in a fraction of the non-local Fock exchange evaluated from the KS-orbitals. Because this approach arises from an interpolation between a (meta-)GGA approach and the pure Hartree–Fock method (*via* the adiabatic connection) the according functionals are called hybrid functionals to emphasize their relation to both, DFT and wave function concepts. The exact amount of non-local exchange can be determined by different methods. Most straight forward, a fixed amount between 0 and 100% is assumed. Such functionals are also called global hybrid functionals to

distinguish them from those methods where the fraction of non-local exchange varies. The fraction is then either dependent on the electron-electron distance (which is the case for so-called range-separated hybrid functionals)[11] or it is a function of the local density or some related property at each point in space (which is the case for so-called local hybrid functionals).[12]

Finally, at the top of the functional hierarchy are those functionals which also make use of the information from the unoccupied (virtual) orbitals. By that, the correlation potential also becomes non-local (but could be mapped back to a local potential, if desired). The exact ways how this information is exploited differ. For most cases, a perturbation expansion similar to the wave function equivalent or the random-phase-approximation (RPA) is applied. DHDFAs form a subset of this highest class.

For each class in this hierarchy, several different functionals exist which differ in their mathematical form and their parametrisation. Because no general rules or guidelines exist how to construct the XC functional, this allows to come up with very different approaches but also to vary a given functional only slightly to better suit one's purposes or personal preferences. This can make the variety of available functionals quite incalculable.

To judge the quality and general applicability of each functional approximation, extensively and thoroughly assessing the performance on as many benchmark sets as possible is required. Therefore, many of such benchmarks have been established and are applied to DFT methods regularly.[13–19] There is no guaranteed increase in accuracy for a problem at hand when going to the higher functional approximations. Nevertheless, based on these results, some general trends can be identified and often, an improved performance can be expected for higher functional classes and this improvement is normally also more pronounced than the variations within the same class. This will be demonstrated for a few representing functionals in this chapter.

The better results do not come for free. The computational cost clearly depends on the functional class (and increases with the increase of information to be included). Obviously, for the most advanced functionals it is most important to exploit methodical improvements for the acceleration of computations. Therefore, the most important developments in combination with DHDFAs are also reviewed at the end of this chapter.

2 Theoretical background and variants of double hybrid density functional approximations

The very basic idea of DHDFAs is not only to mix Fock exchange (E_X(Fock)) with the exchange functional contribution (E_X(DFT)) but to do the same with electron correlation contributions based on DFT (E_c(DFT)) and post-HF methods (E_c(pHF)). These energy contributions have to be added to the Hartree energy (E_H), the sum of kinetic energy, nuclei–electron interaction, and (classical) Coulomb energy, which is the same

in DFT and wave function theory. This automatically leads to the following energy expression:

$$E(\text{DHDFA}) = E_H + a_1 E_x(\text{DFT}) + a_2 E_x(\text{Fock}) + b_1 E_c(\text{DFT}) + b_2 E_c(\text{pHF}) \quad (3)$$

Two theoretical derivations exist to introduce the concept of DHDFAs to DFT in a rigorous way. They are presented in the following sections along with several approximations to these derivations. The approaches are discussed with a focus on perturbation theory for the wave function part but in principle, other post-HF methods can be (and have been) applied.

2.1 The Görling–Levy KS perturbation theory

To transfer the ideas of the perturbation theory to KS-DFT in a rigorous way, Görling and Levy defined a reference system based on KS-DFT but with exchange only. Based on that, they could establish a connection of all KS orbitals (including the virtual ones) which shows that the orbitals have a functional dependence on the exact electron density of the system. Then, electron correlation can be introduced as perturbation of the reference system and can be developed in a series expansion of higher order corrections. This defines the Görling–Levy KS perturbation theory (GLKS-PT).[20,21]

As in wave function theory, the second order correction is the lowest order that corrects the energy for correlation contributions. In opposite to the wave function counterpart, all operators are local in GLKS-PT. Therefore, the exchange and correlation potential has to be mapped to a local potential. Further, orbitals contain information of the correlation potential and therefore, their form has to be solved self-consistently. As a consequence, the whole approach is computationally quite demanding and can lead to numerical instabilities.

The GLKS-PT leads to a special case of the energy expression in eqn (3). One can think of it as the case $a_1 = b_1 = 0$ and, consequently, $a_2 = b_2 = 1$ because all functional parts are replaced by orbital-dependent terms which are similar to wave function theory expressions. Of course, they are not identical to the equivalent wave function terms due the difference described above. For example single excitation contributions also enter in the correlation term which can be neglected in wave function theory due to Brillouin's theorem.

The special form of the GLKS-PT energy expression eliminates the dependence on the (unknown) exchange-correlation functional and the requirement to fix the mixing parameters. Despite these not to underrate advantages, GLKS-PT has not seen a widespread use. The main reason for this might be the computational difficulties which have to be compared to only minor improvements to other KS-DFT approaches and related, more efficient wave function approaches.

A much simpler, more heuristic approach for a KS-PT theory was later given by Filippi et al.[22] Here, the KS determinant of the non-interacting system is used a the reference system and higher order energies are obtained in analogy to Møller–Plesset perturbation theory by replacing the Fock exchange potential with the KS exchange-correlation potential

in the fluctuation operator (the perturbation term). The theory has only been applied up to first order corrections.

2.2 The constraint search mechanism with a partitioned Hamiltonian

Another way to combine DFT and WFT can be derived as follows. First, the electron–electron interaction operator has to be partitioned:

$$\hat{v}_{ee} = \hat{\nu}_{ee} + \hat{\bar{\nu}}_{ee} \tag{4}$$

$$\hat{\nu}_{ee} = \frac{f(r_{12})}{r_{12}} \tag{5}$$

$$\hat{\bar{\nu}}_{ee} = \frac{\bar{f}(r_{12})}{r_{12}} \tag{6}$$

$$1 = f(r_{12}) + \bar{f}(r_{12}) \tag{7}$$

The partitioning depends on the electron-electron distance r_{12} *via* any function $f(r_{12})$ for which a complimentary function exists such that the condition in eqn (7) and $0 \leq f(r_{12}) \leq 1 \forall r_{12}$ can be fulfilled.

Now, a Hamiltonian of partly interacting electrons, $\hat{H}(f(r_{12}))$, can be defined by using $\hat{\nu}_{ee}$ instead of $\hat{v}_{ee}$. Likewise, a complimentary density functional, $\bar{E}_{Jxc}[\rho]$, exists which yields the remaining energy contributions missing when the energy is evaluated with $\hat{H}(f(r_{12}))$. In addition, the potential belonging to $\bar{E}_{Jxc}[\rho]$ ensures that by minimising the total energy the exact electron density of the system is found. The energy then is:

$$E = \langle \Psi^{f(r_{12})} | \hat{T} + \hat{v}_{ne} + \hat{\nu}_{ee} | \Psi^{f(r_{12})} \rangle + \bar{E}_{Jxc}[\rho_{\Psi^{f(r_{12})}}] \tag{8}$$

To derive $\Psi^{f(r_{12})}$, it is known from the constraint search mechanism that one has to find the wave function that minimises the energy in the bracket term *and* yields the exact density of the system (The one that minimises the energy of the system). This can be done with the usual wave function machinery, although based on $\hat{H}(f(r_{12}))$ and respecting the external potential from $\bar{E}_{Jxc}[\rho_{\Psi^{f(r_{12})}}]$, thereby introducing post-HF concepts to a density functional dependent energy. Of course, the necessary $\bar{E}_{Jxc}[\rho_{\Psi^{f(r_{12})}}]$ is unknown in general. Nevertheless, this energy expression can be useful to derive a more feasible one and justifies an energy approximation as in eqn (3).

2.3 Empirically fitted DHDFAs (including spin-component scaling)

Based on the general idea of DHDFAs and the results of the previous section, several ways how to fix the parameters and choose the functional forms in eqn (3) have been derived. These are discussed here and in the following sections.

The parametrisation of eqn (3) can been done on a more heuristic basis. And this has actually been put forward in the seminal paper by Grimme which stimulated the whole research about DHDFAs.[7] Grimme was looking for an efficient way to improve DFAs and to include

information from the virtual orbitals. Obviously, the most efficient way known for post-HF methods to do so is a MP2-like correlation contribution. And based on the argument that all exchange and correlation contributions, respectively, should add up to 1 and that the model should be as simple as possible, he fixed parameter a_1 by $a_1 = 1 - a_2$ and b_1 by $b_1 = 1 - b_2$. The actual values of the two remaining free parameters for the amount of Fock exchange and MP2-like correlation had to be fixed by fitting.

Nevertheless, as sketched in Fig. 2, some parameter combinations can be excluded by rational arguments. Especially extreme points have been ruled out as rather hardly meaningful. Likewise, combinations of large portion of MP2 correlation but low amounts of Fock exchange were identified to be possibly plagued by double-counting effects because the exchange part of a DFA also accounts for some correlation effects.

The first of these semi-empirical DHDFAs is called B2PLYP and has 53% Fock exchange and 27% MP2 correlation. Later on, several other variants have been fitted for either very special application cases but also to improve the overall performance. General trends are that, in comparison to standard global hybrids, quite high amounts of Fock exchange ($\geq 50\%$) are made use of. This is especially useful to reduce the self-interaction error of DFT. Further, it often holds that $b_2 \leq (a_2)^2$, a condition which was derived later by Fromager[23] based on more rigorous theoretical arguments in an extension of the work by Shakras *et al.* (see below).[24] Again, this agrees to rationalisations made above for meaningful parameters. For a detailed compilation of various of these DHDFAs, see ref. 5.
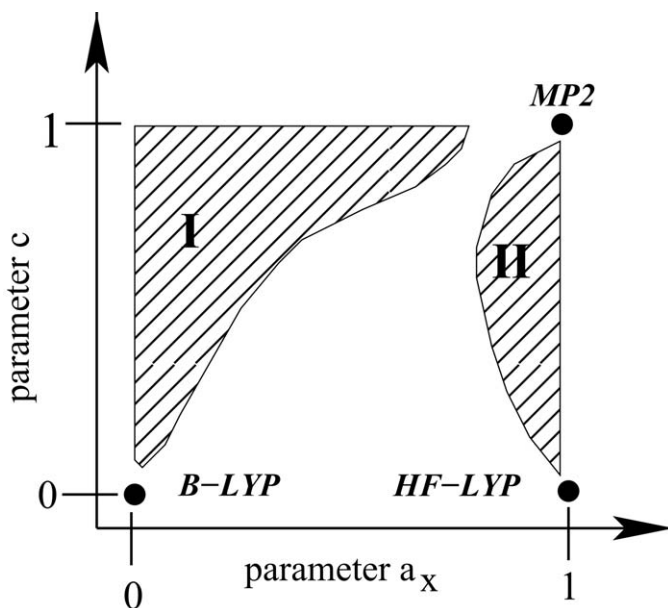

Fig. 2 Schematic description of possible methods arising from different combinations of HF exchange end PT2 mixing parameters. The dashed areas I and II can be excluded by physical reasoning. Reprinted from S. Grimme, *J. Phys. Chem*, 2006, **124**, 034108 with the permission of AIP Publishing.

Two special variants of fitted DHDFAs should be mentioned. First of all, for the MP2-like correlation contribution the spin-component scaling (SCS) approach can be employed.[25] For MP2, it is known that same-spin and opposite-spin electron pairs are not treated equally because some same-spin correlation is already described at the HF level. As a result, the same-spin correlation is overestimated by MP2. This can be remedied by introducing two scaling parameters for same-spin and opposite-spin electron pair contributions which leads to the SCS approach.

This idea has been transferred to DHDFAs independently in two publications.[26,27] The better known version are the so-called dispersion-corrected spin-component scaled double-hybrids (DSD, the dispersion correction is discussed below). The effect is not as good as for SCS-MP2 but some improvements can be found. Of course, now two parameters for the MP2-like correlation part have to be fitted which introduces some more flexibility in the optimisation. The DSD methods also have been revised testing a diverse set of possible DFA components but without refitting their parametrisation.[28]

Another advantage of spin-component scaling comes from the fact that by this approach the same-spin contribution can be removed completely which leads to spin-opposite scaling (SOS). Once only the SOS term enters the general PT2 energy expression, a Laplace transformation can be applied to it. This allows to reduce the computational effort from a formal $\mathcal{O}(N^5)$ to a $\mathcal{O}(N^4)$ scaling where N is a measure for the system size. The technical details will be discussed in more detail below. Examples for fitted SOS-DHDFAs are B-OS3PLP,[29] PTPSS,[19] and PWPB85.[19]

The other variant is the XYG3 method[30] which also introduces additional parameters by splitting the exchange and correlation term in local density and generalised density contributions which are scaled independently, yielding:

$$E(\text{XYG3}) = E_\text{H} + (1 - a_1)E_\text{x}(\text{LDA}) + a_1 E_\text{x}(\text{Fock}) + a_2 E_\text{x}(\text{GGA})$$
$$+ (1 - b_1)E_\text{c}(\text{LDA}) + b_1 E_\text{c}(\text{MP2}) + b_2 E_\text{c}(\text{GGA}) \tag{9}$$

By setting $b_2 = 1 - b_1$ the authors of the method remove one degree of freedom in the parametrisation but still, there is one more parameter for fitting than in other approaches. Another speciality of the XYG3 method is that the corresponding orbitals are not obtained from the underlying global hybrid according to eqn (9) but from a standard B3LYP computation. Therefore, the method cannot be seen as a perturbation series development because the underlying reference system does not correspond to the applied perturbation. Further, but this is inherited from the B3LYP method, XYG3 combines LDA correlation with a GGA contribution from the LYP correlation functional. In contrast to many other GGA approaches, LYP has not been developed as an enhancement factor to LDA and also does not yield the LDA correlation energy for the uniform electron gas for which LDA is exact. This makes the whole approach an ad hoc mixing of energy expressions with even weaker justification as the approaches before. This has to be seen in light of the fact that a clear

improvement in performance with respect to other empirical DHDFAs cannot be found.

The idea was then also exploited with PBE as underlying DFA where the entering energy terms are a little more consistent, yielding the xDH-PBE functional.[31] This was later on combined with a dispersion correction.[32] As can be expected, both variants show satisfying results.

2.4 (Scaled-density) one-parameter DHDFAs

Based on the results from the constraint search mechanism, Sharkas et al. formulated one-parameter DHDFAs.[24]

Consider a constant scaling of the electron-electron interaction like in global (double) hybrid functionals: $f(r_{12}) = \lambda$. In that case,

$$\bar{E}_{\text{Jxc}}[\rho_{\Psi f(r_{12})}] = E_{\text{Jxc}}[\rho_\Psi] - E_{\text{Jxc}}^{f(r_{12})}[\rho_{\Psi f(r_{12})}], \tag{10}$$

where $\bar{E}_{\text{Jxc}}[\rho_\Psi]$ is the electron–electron interaction functional for the full interacting system, and accordingly, $E_{\text{Jxc}}^{f(r_{12})}[\rho_{\Psi f(r_{12})}]$ the one for the scaled interaction. The (classical) Coulomb and exchange interactions just scale linearly in this case. This is not true for the correlation contribution, but this can be expressed by density scaling:

$$E_{\text{c}}^\lambda[\rho_{\Psi(\lambda)}] = \lambda^2 E_{\text{c}}[\rho_{\Psi(\lambda)}/\lambda] \tag{11}$$

Approximating $E_{\text{c}}[\rho_{\Psi(\lambda)}/\lambda] \approx E_{\text{c}}[\rho_{\Psi(\lambda)}]$ and inserting all terms due to constant scaling into eqn (8) leads to:

$$E = \langle \Psi^\lambda | \hat{T} + \hat{v}_{\text{ne}} + \lambda \hat{v}_{\text{ee}} | \Psi^\lambda \rangle + (1-\lambda)E_{\text{Jx}}[\rho] + (1-\lambda^2)E_{\text{c}}[\rho] \tag{12}$$

If now the search for the minimising wave function is restricted to a single determinant and a MP2-like approach for the correlation contribution is exploited, the final energy expression becomes:

$$E = E_{\text{H}} + \lambda E_{\text{x}}(\text{Fock}) + (1-\lambda)E_{\text{x}}[\rho] + (1-\lambda^2)E_{\text{c}}[\rho] + \lambda^2 E_{\text{c}}(\text{PT2}) \tag{13}$$

Here, $E_{\text{c}}(\text{PT2})$ is the MP2-like correlation energy contribution evaluated with the single determinant orbitals which minimises the energy:

$$E = E_{\text{H}} + \lambda E_{\text{x}}(\text{Fock}) + (1-\lambda)E_{\text{x}}[\rho] + (1-\lambda^2)E_{\text{c}}[\rho] \tag{14}$$

Finding this determinant is equivalent to solve a standard global hybrid problem with a fraction of λ Fock-exchange. Comparing the linear-scaling DHDFA energy expression with eqn (3), it is obvious that now all parameters except for one are fixed.

The final parameter λ has been fixed by empirical tuning and optimal values range from 0.55 to 0.80 depending on the actual choice of the exchange-correlation functionals and whether density scaling is considered or not. Sharkas et al. tested DHDFAs which either considered density scaling or not in the correlation functional. Interestingly, for PBE it is beneficial to include the latter while results are worsened for B-LYP.

The approach has also been extended for DHDFAs based on the meta-GAA TPSS.[33]

2.5 Parameter-free DHDFAs

In an attempt to exclude any empiricism from the functional construction, Brémond and Adamo introduce a so-called parameter-free variant of the scaled-density one-parameter DHDFAs.[34] Of course, they do not get rid of the necessity to define the λ parameter but instead of fitting they derive it based on theoretical arguments. First of all, they approximate the density scaling of the correlation functional to be:

$$E_c[\rho/\lambda] \approx \lambda E_c[\rho] \tag{15}$$

which then yields in total:

$$E = E_H + \lambda E_x(\text{Fock}) + (1-\lambda)E_x[\rho] + (1-\lambda^3)E_c[\rho] + \lambda^3 E_c(\text{MP2}) \tag{16}$$

λ is fixed at 0.5 based on arguments for a linearised adiabatic connection. It should be noted that this value is actually justified for an integration of the system with no electron-electron interaction to the full interacting system. In the case here, the integration should have started at λ not Zero, which should have resulted in a different value. Further, the scaling for $E_c(\text{MP2})$ has been derived by arguing that all correlation contributions should add up to 1 but actually, the scaling factor should still be λ^2 because this is a result of the perturbation correction not the additivity of terms.

Brémond and Adamo combined their derivation with the PBE functional, which is considered to be a non-empirical GGA functional, and thereby defined the PBE0-DH method.

Interestingly, another approximation:

$$E_c[\rho/\lambda] \approx (1-\lambda)E_c(\text{PT2}) + \lambda E_c[\rho] \tag{17}$$

leads to the same energy expression but does also not give a rational for the final value of λ.[35] Accordingly, Cortona presented a derivation of four different λ values $\left(\dfrac{1}{2}, \dfrac{1}{3}, \dfrac{1}{4}, \text{and } \dfrac{1}{5}\right)$, all based on equally rigorous theoretical arguments.[36]

2.6 Approximations to the integrand of the adiabatic connection to define DHDFAs

Another way to exploit the adiabatic connection is to approximate the integrand in the (exact) expression:

$$E_{xc} = \int_0^1 E_{xc}^\alpha \, d\alpha \tag{18}$$

E_{xc}^α is the exchange correlation contribution of an electron system with an electron interaction strength scaled by α (To avoid confusion, the notation α instead of λ is used here, although the meaning is similar). The exact form of the integrand is unknown, therefore it needs to be

approximated. On its own, an approximation does not introduce any correlation contribution based on virtual orbitals. But some boundary conditions for the integrand are known. These are:

$$E_{xc}^{\alpha=0} = E_x(\text{HF}) \tag{19}$$

$$\left.\frac{\partial E_{xc}^{\alpha}}{\partial \alpha}\right|_0 = 2E_c(\text{PT2}) \tag{20}$$

$$E_{xc}^{\alpha=\infty} = \text{const.} \tag{21}$$

And further, one normally assumes that standard exchange-correlation DFAs, even LDA, work quite well such that $E_{xc}^{\alpha=1} = E_{xc}(\text{DFT})$. Based on these constraints, approximations to E_{xc}^{α} can be parametrised. Typical choices are either a polynomial expansion:

$$E_{xc}^{\alpha} = \sum_{i=0} c_i \alpha^i \tag{22}$$

or a Padé expansion:

$$E_{xc}^{\alpha} = \frac{\sum_{i=0} c_i \alpha^i}{1 + \sum_{j=1} d_j \alpha^j} \tag{23}$$

The earliest of such models is the [2/2]-Padé method by Ernzerhof.[8] A speciality of this method is that is has been constructed to directly compute atomisation energies and that it contains parameters which have to be determined system-dependently.

Another approach is PBE-ACDH from Su and Xu.[37] Their approximation can be seen as being based on a [2/2]-Padé approximation although the expression is slightly more complicated. By arguments following the construction of the original global hybrid, Becke half-and-half, they set the amount of Fock exchange and DFA contributions to 50% which in turn fixes all other free parameters in the model.

Other variants are based on a polynomial expansion of either second[38,39] or third order.[40] Kim et al.[38] and Brémond et al.[39] worked independently on a quadratic integrand form (QI) and came up with quite similar results:

$$E_{xc}(\text{QI}) = \frac{2+a}{3}E_x(\text{HF}) + \frac{1-a}{3}E_x(\text{DFT}) + \frac{1}{3}E_c(\text{PT2}) + \frac{1+b}{3}E_c \tag{24}$$

The QACF-2 method by Kim et al. based on PBE has $a = 0$ and $b = 0$, while the QIDH approach by Brémond has $a = 3^{2/3} - 2 \approx 0.0801$ and $b = 1$. The differences in the parameters are based on slightly different approximations to $E_{xc}^{\alpha=1}$. The model of Brémond does not fix a and it was set to its value such that $\left(\frac{2+a}{3}\right)^3 = \frac{1}{3}$ based on arguments for the derivation of PBE0-DH.

The cubic integrand approximation by Alipour[40] leads to:

$$E_{xc}(CI) = \frac{1+a}{2}E_x(HF) + \frac{1-a}{2}E_x(DFT) + \frac{1}{6}E_c(PT2) + \frac{5}{6}E_c \quad (25)$$

and a was again chosen such that $\left(\frac{1+a}{2}\right)^3 = \frac{1}{6}$ which yields $a \approx 0.05$

It should also be noted that the QI as well as the CI model integrands diverge for $\lim_{\alpha \to \infty}$ and therefore violate a known constraint on the integrand. Anyhow, as all of the derived parameters fall into a meaningful region, the corresponding DHDFAs work quite well. Nevertheless, a systematic improvement over fitted DHDFAs cannot be found.

Recently, some of the rigorousness of the QIDH and CIDH approaches have been abandoned to find suitable expressions for SOS variants with the aim to exploit the better scaling of such an approach.[41,42] To express the same spin correlation contribution in terms of the opposite spin only, Brémond *et al.* as well as Alipour assume $E_{c,os} = (PT2) = 3E_{c,ss}(PT2)$ based on literature data and their own computations for ratios of the energy contributions. Brémond *et al.* introduce an additional empirical parameter but after they already determined the average ratio of $E_{c,os} = (PT2)$ and $E_{c,ss}(PT2)$ they could have set this parameter directly to 1 to find $E_c(PT2) \approx \frac{4}{3}E_{c,os}(PT2)$. Instead, they found this value by fitting. Alipour skipped this detour and therefore calls his approach SOS0 instead of SOS1 to indicate Zero fitted parameters. In the end, both SOS models are identical and are purely based on an empirical determination of the same-spin to opposite-spin correlation energy ratio.

A more general analysis how DHDFAs behave along the adiabatic connection has been given by Cornaton *et al.*[43] They find discontinuities along the integration path which are not in general problematic but their removal might be taken as guidance to further improve the methods.

2.7 Performance of the DHDFA energy model

To put the performance of the various DHDFA variants into context, some of them along with their predecessor hybrid and (meta-)GGA functionals have been computed here for a challenging test set of isomerisation energies which has been published recently.[44] The computational details have been kept, *i.e.* the same molecular geometries as well as the cc-p(wC)V5Z basis set has been applied. (For more details, see the original work.) Isomerisation energies are a good test case because they directly assess the abilities of electron structure methods but avoid specific pitfalls like for example in atomisation energies. The latter depends to a large extend on the ability to describe the open-shell atomic systems correctly, which normally adds to the complexity of the task.

This specific test set consists of small compounds containing only main group elements. Therefore, the size of the molecules does not play a role and a dispersion correction can safely be neglected. Likewise, virtually no multi-reference character should dominate the electronic structures. The deviation histograms for the DFAs B-LYP, B3-LYP,
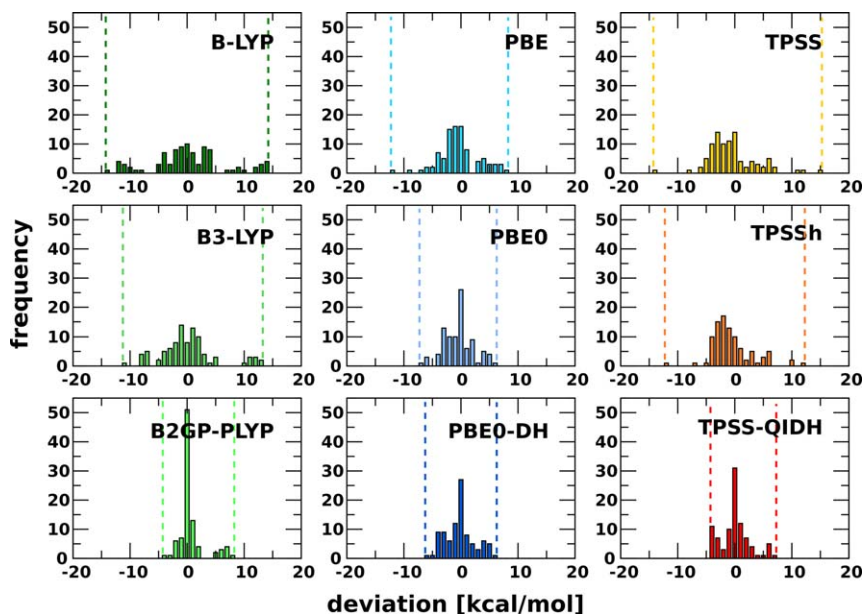
Fig. 3 Distribution of deviations from the reference for various DFAs for a set of isomerisation reactions. For clarity the deviation range are indicated by dashed lines.

B2GP-PLYP, PBE, PBE0, PBE0-DH, TPSS, TPSSh, and TPSS-QIDH are given in Fig. (3). Deviations are determined against CCSD(T)/cc-p(wC)VTZ + basis set correction results which are close to the CCSD(T) basis set limit.

As can be seen, the range of observed deviations (indicated by vertical dashed lines) narrows and the distributions get more systematic around Zero for the progression from (meta-)GGAs (top row) to hybrid DFAs (middle row) to DHDFAs (bottom row). Of the standard DFAs, only PBE0 shows a reasonable performance. On the contrary, there is almost no improvement, when PBE0-DH is considered. The reason for this might be related to the quite small amount of perturbation correction (12.5%) and the somewhat lower amount of Fock exchange (50%) in comparison to the other two DHDFAs. Although TPSS-QIDH (Fock exchange: 74.7%, 33.3% MP2 correction) does not perform significantly better, it clearly outperforms the meta-hybrid DFA TPSSh. The empirically fitted B2GP-PLYP (Fock exchange: 65.0%, 36.0% MP2 correction) shows the best performance for this test case and about $\frac{3}{4}$ of all results are within chemical accuracy (1 kcal mol^{-1}) to the reference results.

Nevertheless, the deviation range is not significantly smaller in comparison to the other DHDFAs. Interestingly, a small cluster of outliers can be identified which form a set of nine similar reactions. The other two DHDFAs work much better in these cases: The deviations for B2GP-PLYP lie in the range of 5.4 to 7.9 kcal mol^{-1}, for PBE0-DH in the range of −0.1 to 4.2 kcal mol^{-1} and for TPSS-QIDH in the range of 0.0 to 2.8 kcal mol^{-1}. This might indicate that not only the amounts of Fock exchange and perturbation correction play a role for the performance but

just as much does the underlying DFA and here seems to be room for improvement.

The better performance of B2PLYP *vs.* PBE0-DH has also been found by Bousquet *et al.*[45] who evaluated both approaches for a large benchmark set for thermochemistry, properties, and kinetics (GMTKN30).[19] Overall, PBE0-DH had a mean absolute error of 3.2 kcal mol^{-1} (PBE0: 3.5 kcal mol^{-1}), while B2PLYP had a mean absolute error of 2.7 kcal mol^{-1} (B3LYP: 4.7 kcal mol^{-1}).

A similar test set of diene isomerisation reactions confirms these results: empirically fitted DHDFAs normally outperform their non-fitted counterparts, but all do show an acceptable performance comparable to the wave function methods counterparts, MP2 and SCS-MP2, respectively.[46] It should be noted that SCS-MP2 works exceptionally well for exact these kinds of problems but its performance is much more sensitive to the underlying HF reference and DHDFAs proved often more robust for complicated electronic structure problems.[19,47]

For example, DHDFAs have been applied to compute reliable heats of formation for large fullerenes.[48] And a recent benchmark on cycloreversion barrier heights also underlined the robust performance of DHDFAs[49] although long-range corrected hybrid DFAs (see below) generally performed better in these cases which might indicate a problem with self-interaction.

Today, it is unclear if a DFA can be constructed from the known design principles which improves results for outliers while not deteriorating those which already work very well. One of the most recent attempts combined the strongly constrained and appropriately normed semilocal DFA[50] (SCAN) with the concepts of fitting-free DHDFAs.[51] Though as good as comparable DHDFAs, no clear improvement to other methods could be found but further testing might be needed for a final conclusion.

Beside the fact that not all problems are solved for DHDFAs, they are still robust and accurate approaches which normally improve on other DFAs and also on related MP2 methods.

3 Extensions of the DHDFA approach

Beside the general variants to derive the energy expressions for a DHDFA, the methods have seen some further general extensions. These are introduced in the following.

3.1 Dispersion correction

Modern DFA are compromised by two major flaws: the problems of self-interaction and the lack of a quantitatively right description of London dispersion effects over all interaction ranges. As mentioned before, the former is somewhat softened for DHDFAs because of the relatively high amount of Fock exchange. The latter could also be addressed by this approach because London dispersion is basically a long-range non-local correlation phenomenon which is accounted for by the PT2 term in DHDFAs.
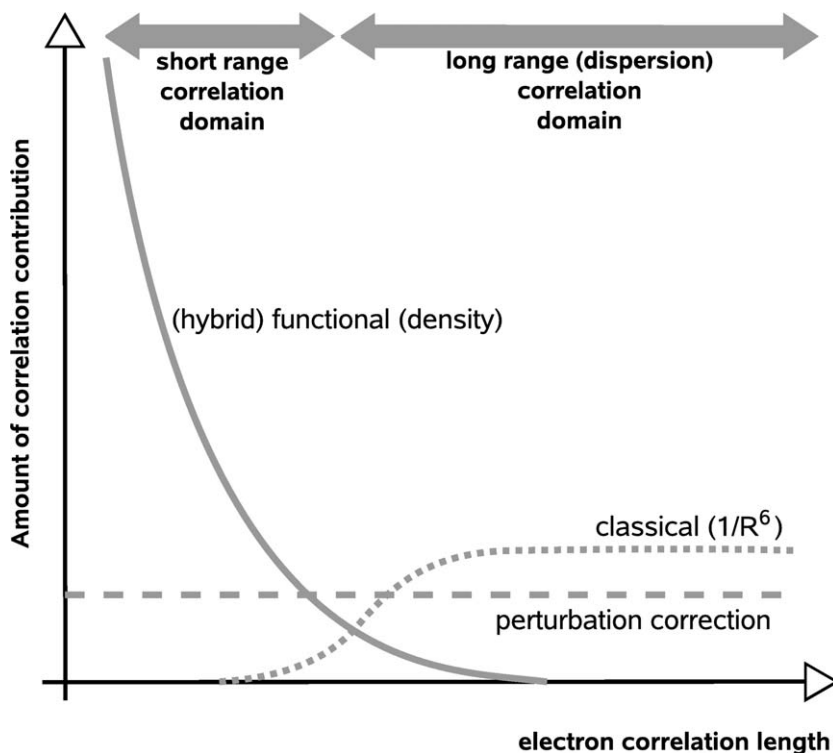
Fig. 4 Scheme for the energy contributions to the London dispersion energy at different interaction ranges. Reprinted with permission from T. Schwabe and S. Grimme, *Acc. Chem. Res.*, 2008, **41**, 569–579. Copyright (2008) American Chemical Society.

In practise, the quite low amount of the PT2 correlation energy (normally <40%) cannot compensate for the deficiency. A remedy can be the addition of a dispersion correction term. Various forms of the latter have been established for DFA development in general.[52,53] Because in short to mid-range regions, DFAs themselves are capable of computing the according correlation energy and because some fraction of PT2 correlation energy is present at all interaction ranges, a suitable damping of the dispersion correction is necessary. This is depicted in Fig. (4).

Efficient formulations for a dispersion correction can be derived from a force-field like term. Such approaches are subsumed under the label DFT-D. The energy correction can be generally given as

$$E_{\text{disp}}(\text{DFT-D}) = \sum_{n = 6,8,\ldots} \sum_{AB} f_{\text{dmp}}(R^{AB}) \frac{C_n^{AB}}{(R^{AB})^n} \quad (26)$$

Here, the sums run over all orders n of the dispersion interaction considered in the actual model (normally truncated at 6 or 8) and all atom pairs AB in the system, which have a distance R^{AB}. The dispersion coefficients C_n^{AB} are atom pair dependent and their actual values depend on the dispersion correction model. Finally, to avoid divergence of the energy expression and suppress double counting effects, a damping function f_{dmp} for short range interactions is introduced.

Already the first attempt to combine DHDFAs with a dispersion correction showed a significant improvement in the accuracy of the functional[54] and newer versions of the dispersion correction[55,56] have also been successfully combined with DHDFAs. Not only in (standard) benchmark tests, but also in application studies the good performance has been confirmed.[57,58]

It should be kept in mind that the dispersion correction also proofed beneficial for the description of intramolecular stabilisation effects thereby directly influencing thermochemical properties so that it is not only important for computing large non-covalent systems. This is the reason why a dispersion correction was included right from the beginning to parametrise the DSD-DHDFAs.[27,28]

The non-empirical DHDFAs have also been combined with an empirical dispersion correction albeit the improvement is a little less distinct than for empirical fitted DHDFAs.[59] The authors of that studied conclude that this might be rooted in the way how the DHDFAs are derived but it is more likely related to the underlying (pure) DFA. It is well known that PBE and TPSS functionals profit less from an empirical dispersion correction because they already overestimate correlation effects to some extend in the short and mid-range interaction region.

Another approach to the dispersion problem is the non-local van-der-Waals correction called DFT-NL. The correction is a density based additive energy term:

$$E_{\text{disp}}(\text{DFT-NL}) = \int dr \rho(r) \left[\frac{1}{32} \left(\frac{3}{b^2} \right)^{3/4} + \frac{1}{2} \int dr' \rho(r') \phi(r, r', b) \right] \qquad (27)$$

where $\phi(r, r', b)$ is an interaction kernel[60] and b a functional dependent parameter. Some of the authors that combined non-empirical DHDFAs with an empirical dispersion correction also did the same with DFT-NL[61] and for empirically fitted DHDFAs with spin-component scaling, this was repeated later.[62] In both studies, only non-covalent interaction energies are assessed. No definite statement can be made whether this strategy is more promising for all kinds of applications than the one with a DFT-D approach. Further, the DFT-NL correction has not been scaled down so that one can actually expect some double-counting effects. Nevertheless, good results for other non-covalently bound systems have been found in a further study although an additional repulsive three-body term originating from DFT-D methods has been shown to improve results[63] albeit DFT-NL already considers many-body effects implicitly *via* the non-linear density dependency. This might indicate that the double-counting problematic should be investigated further.

3.2 Analytical derivatives

For many properties other than the energy, analytical derivatives of the energy expression with respect to a given perturbation are needed for an efficient computation. For DHDFAs, these have been derived for the first[64] and second[65] derivative of the nuclei coordinates. The resulting

expressions are required for (efficient) geometry optimisations and frequency computations, respectively. Very recently, the gradient expression has been extended for a hybrid quantum mechanics/molecular mechanics approach including a fluctuating charge model.[66] But a detailed discussion of such approaches is outside the scope of the present chapter.

A complication for these derivatives result from the fact that the DHDFA energies are not optimised with respect to the parameters which define the corresponding wave function model (*i.e.* the molecular orbital coefficients). This is a known problem from standard MP2 and can be solved by a Lagrangian method (also called Z-vector method). Further, due to presence of the DFA exchange-correlation potential in the KS-operator, the second derivative of the DFA exchange-correlation energy already enters for the geometry gradient. While this does not prevent the implementation of an analytic gradient (and higher derivatives), the cost for the gradient computation increases much more than the energy computation if compared to a standard SCF procedure like a (meta-)GGA or hybrid DFA. The complexity even rises for DHDFAs of the XYG3 type where the KS orbitals stem from another hybrid DFA as the one applied to evaluate the energy. Nevertheless, analytic gradient formulations can be found.[67]

Although B2PLYP turned out as a very accurate and robust method for geometry optimisation, its relatively high computational cost normally does not warrant its application because good geometries can already be obtained at lower level of theory when sufficiently large basis sets of at least triple-ζ quality are applied. Nevertheless, in complicated cases it might be a last resort but should then be combined with a suitable dispersion correction. This combination allowed the computation of several rotation constant for molecules in the gas phase with an accurracy of about 3 MHz.[68] Recent studies on bond length alteration in series of linear molecules with conjugated π-systems[69] and on superhalogene molecular structures confirm these results.[70]

On the other hand, DHDFAs proved to be very competitive in the computation of frequencies and anharmonic vibration corrections in comparison to the even more costly CCSD(T). Kesharwani *et al.* fitted scaling parameters for DHDFA vibrational frequencies to reproduce experimental data, a standard approach for many quantum chemistry methods, and obtained good results.[71] Barone *et al.* tested B2PLYP to compute accurate vibrational and rotational spectra of gas phase molecules and found that it can replace CCSD(T) in according computational protocols.[72,73] B2PLYP has also been used as accurate reference method to train low-cost method for the determination of reliable vibrational wavenumbers.[74] Howard *et al.* on the other hand saw no significant improvement when they applied DHDFAs to study vibrational frequencies of water clusters, irrespectively whether they applied a dispersion correction in addition or not.[75] A reason might be that these complexes are dominated by hydrogen bonds which still prove challenging to be described with dispersion corrections and their description is indeed not significantly improved by DHDFAs.

Alipour exploited the analytic derivative implementations to compute (hyper-)polarisabilites and although he suggests several new DHDFAs to compute these properties, the more general DHDFAs show already good results.[76–79] Similar findings, yet based on numerical differentiation techniques, have been found before for non-linear optical properties by Champagne et al.[80,81]

3.3 Excitation energies

A further property accessible with DHDFAs are excitation energies although their computation requires the extension of the underlying model. DFT is a ground-state theory and therefore, it cannot directly describe excited states. To get access to excitation energies and properties of excited states either response theory (TD-DFT) or real-time propagation methods, respectively, have to be applied. A rigorous derivation of a linear response theory for a post-Hartree–Fock/short-range DFT method has been presented, but the underlying wave function method is a multi-reference SCF approach, not a perturbation approach.[82]

For standard DHDFAs another route has been presented. Based on analogy to the CIS(D) method,[83] a perturbation correction to the excitation energies of the underlying hybrid functional obtained from a TD-DFT computation has been proposed.[84] The correction is a combination of a *direct* ($\Delta E_{\text{ecx}}^{\text{D}}$) and an *indirect* term ($\Delta E_{\text{ecx}}^{\text{I}}$). The former can be understood as an orbital response term due to perturbation by the electric field of the incoming photon and the latter is a perturbation correlation correction to the excited state energy such that

$$E_{\text{exc}}(\text{DHDFA}) = E_{\text{exc}}(\text{TDDFT}) + b_2(\Delta E_{\text{exc}}^{\text{D}} + \Delta E_{\text{exc}}^{\text{I}}) \qquad (28)$$

The model allows for some freedom how to parametrise the energy correction. But Grimme and Neese found that already good results can be obtained when the scaling parameter b_2 is set equal to the scaling factor in the ground state energy. This should also be consistent with a direct derivation of the excitation energy from a linear response approach.

Empirically fitted DHDFAs were then tested for several challenging cases of larger dyes. Especially B2GP-PLYP is competitive to iterative and therefore computationally more demanding second-order wave function methods like ADC(2) and CC2.[85,86] In another study, B2PLYP showed good results for cyanine dyes in comparison to other more sophisticated wave function methods[87] and DHDFAs performed well for the challenging task to describe the L_a and L_b states and their relative order in polycyclic aromatic hydrocarbons.[88,89] Non-empirical DHDFAs have also been assessed for excitation energies and found to be an improvement with respect to other levels of DFAs, but the empirical DHDFAs still performed slightly better.[90] These findings have been confirmed recently.[91] DHDFAs were also the most accurate DFT methods for a benchmark of photosensitizers with a competitive accuracy to more expansive wave function methods.[92] Finally, DHDFAs can also be applied to compute circular dichroism spectra.[93]

Another strategy to compute DHDFA-like excitation energies is followed by Rebolini *et al.*[94] They start from the simplified KS perturbation theory from Filippi *et al.*[22] but make use of a range-separated approach for the electron–electron interaction (see below). They also extend the method up to second order. Although this development improved ground-state energies, not much gain was found for excitation energies.

3.4 Orbital-optimised approaches

It is intrinsic to all perturbation methods that they break down the faster the worse the unperturbed reference system is. To improve on this, an orbital-optimisation (OO) with respect to the total DHDFA energy can be carried out. The idea is borrowed from OO-MP2[95] and was first applied to DHDFAs by Peverati and Head-Gordon[96] and was assessed again recently by Sancho-García *et al.*[97]

As already for the wave function counterparts,[95,98] the OO approach improved especially on open-shell systems like radicals or transition-metal systems and also for stretched bonds (like they occur for dissociation curves and in transition states). Of course, the larger the amount of the MP2 correction the more pronounced is the effect. The OO approach has a higher computational complexity wherefore it should only be used when it is sensible. For example, a large spin contamination in a standard DHDFA computation might indicate that an improvement form the OO approach can be expected.

Alternatively, Thompson and Hratchian addressed the problem of spin contamination by applying a spin projection approach to the KS reference and evaluated the perturbation correction based on the resulting determinant.[99]

3.5 Long-range corrected DHDFAs

As mentioned above, a general derivation of DHDFA can be found based on several splitting schemes of the Hamiltonian or the electron–electron interaction, respectively. One of such schemes is the short-range DFT approach (srDFT),[100] which is not based on a constant scaling but on a distance dependent function for the splitting:

$$f(r_{12}) = \frac{\text{erf}(r_{12})}{r_{12}} \tag{29}$$

$$\bar{f}(r_{12}) = \frac{\text{erfc}(r_{12})}{r_{12}} \tag{30}$$

Here, erf and erfc are the error function and its complementary, respectively. Another choice could be a Yukawa type potential. The basic idea is to apply wave function methods to the long-range part and describe close contacts of electrons with DFT. Lately, this approach has mainly be exploited to treat the exchange part with a long-range (lr) correction from Fock exchange to deal with the self-interaction error[11] but the first approaches combined (multideterminantal) post-HF methods with srDFT.[100,101] Ángyán *et al.*

derived the first srDFT/lrMP2 which proved especially successful when it was applied to van-der-Waals complex systems.[102]

Two different approaches, which also exploit a range-separation dependency, have been presented by Head-Gordon *et al.* In the first version, only the MP2 part of a standard DHDFA (B2PLYP) is complemented with a long-range contribution (in addition to the standard scaled contribution).[29] The motivation was to account for the missing energy contributions to describe London dispersion. The resulting B2-P3LYP and B2-OS3LYP (spin-opposite scaled variant) do not only improve on non-covalent bound systems but also on the thermochemistry of larger systems, which is in accord with the findings for the dispersion corrected versions. An analogue approach was applied to the XYG3 functional to get lrc-XYG3 with similar results.[103]

The other method from the Head-Gordon group, named ωB97X-2, combines a lr-corrected hybrid DFA with an SCS-MP2 part based on the KS orbitals.[104] The motivation not to include a range-separation approach into the perturbation correction was mainly of technical nature to avoid a sr-corrected DFA for the correlation part.

Cornaton *et al.*[105] followed the strategy of Ángyán *et al.* but pointed out that with a splitting treatment of the electron–electron interaction in the MP2 term coupling terms occur between the sr and lr components:

$$E_c(MP2) = \mathbf{t}^{lr}\mathbf{V}_{ee}^{lr} + \mathbf{t}^{sr}\mathbf{V}_{ee}^{sr} + 2\mathbf{t}^{lr}\mathbf{V}_{ee}^{sr} \tag{31}$$

where $\mathbf{t}$ are the MP2 amplitudes and $\mathbf{V}_{ee}$ the two-electron interaction integrals either for the sr or lr variant and to derive the last term on the right-hand side, symmetry relations have been exploited. Normally, in lr-corrected DHDFAs, only the lr/lr terms are treated explicitly while the remaining terms are treated implicitly by the sr DFA. They tested an approach where only the sr/sr contributions are covered by a DFA. The method was later on extended to a Coulomb attenuated method where for short distances the explicit electron–electron interaction operator goes to a finite constant and interpolates to a scaled fraction not necessarily 1 for long distances.[106] Both variants have only been tested on small test systems so that no conclusion can be drawn whether they pose a significant improvement for general applications.

3.6 DHDFA with post-HF methods beyond MP2

As already mentioned, there is no restriction to use an MP2 like approach in combination with a DFA to obtain a DHDFA method. The main reason for this are efficiency arguments although it can be a consequence of how the DHDFA is constructed (*e.g.* see the DHDFAs derived from an approximation to the integrand in the adiabatic connection). It should be emphasised that a full DHDFA model based on a post-HF method beyond MP2 formally requires to evaluate the DFA contributions based on the final post-HF wave function (see eqn (8)) which, in turn, leads to a new potential from the complementary functional and the determination of the wave function (and the density) has to be solved self-consistently including the post-HF contribution. In most of the approaches introduced in the

following, this has been neglected for efficiency reasons. Such an approximation might be justified for energy computations because the change in the density due to a correlation treatment are normally small.[107] For property determinations, the approximation might not be valid but experience with that is scarce and needs further investigations.

Combinations of DFAs with post-HF methods are often based on the srDFT approach. For example, it has been combined with a coupled cluster (CC) treatment.[108–110] Two improvements could be found: the basis sets dependency of the CC part has been softened and the requirement for higher order corrections in the CC expansion, *i.e.* triple excitations, has also been reduced. The explanation for both observations are related: The high angular momenta of basis sets are needed to describe the cusp for the short-range electron–electron interaction, the triple excitations contribute mostly to situations where all three corresponding electrons are close to each other. Both situations are now implicitly described by the srDFT part. srDFT/lrCC improves the description of dipole moments[111] and likewise the computation of atomisation energies with respect to its individual methods (DFT and CC). Nevertheless, the best empirical DHDFAs are still better and have a lower computational cost.

Another post-HF method which can be combined with DFT is the (direct) random-phase approximation (RPA), which can be equally formulated as adiabatic-connection fluctuation–dissipation theorem or ring coupled cluster, an approximation to CCD. For the exact connection between these methods and the different variants of RPA, the reader might see ref. 112. A standard approach is just to apply GGA KS molecular orbitals to compute pure Fock exchange and RPA correlation energies in a post-SCF way.[113] This can be seen as a DHDFA in the flavour of XYG3 but the parameters have been set by rational arguments, not fitting.

The concept has been extended by a combination of srDFT with lrRPA by Scuseria *et al.*[114–117] and Toulouse *et al.*[118–121] Their exact RPA terms differ but the concepts are closely related and show similar results. The approaches have been tested mainly on non-covalent interactions where they show promising results.

Scuseria *et al.* extended their approach by using srDFT and pair coupled cluster in the long-range part.[122] The latter is another approximation to CCSD which can account for static correlation effects. To improve the methods for long-range dynamic correlation effects, long-range direct RPA has also been added to the model.[123] The approaches show good results for dissociation curves even of pathological systems like N_2 and for small non-covalent bond dimers. A systematic test for general applications in computational chemistry is missing though.

Recently, attempts to derive DHDFAs with higher order post-HF methods and scaled global hybrid DFAs have also been presented. Chan *et al.* investigated to what extend an improvement from higher correlation corrections can be gained by using either MP3 and MP4 like correction terms or CCSD- and CCSD(T)-based corrections.[124] The basic functional was BLYP including a large fraction of Fock exchange and a SCS-MP2 correction. The effect of going to higher correlation corrections

was insignificant especially if the much higher computational cost is considered. But it should be noted that the parameters entering the hybrid part of the energy (and thereby determining the MO basis for the correlation contributions) have not been refitted and the DFA contributions have only been computed from the SCF density, not from the one of the post-HF wave functions. Whether changing any of these decisions would change the outcome of the study, is still an open question.

Recently, RPA was also combined with global hybrids in a DHDFA like approach. The method by Grimme and Steinmetz (PWRB95)[125] and Mezei *et al.* (dRPA75)[126] differ slightly. The former holds on to the post-SCF idea of standard RPA approaches but mixes in DFA exchange and correlation as well as some amount of DFT-NL dispersion correction to compensate the downscaled RPA correlation energy for long-range interaction. The post-SCF approach allows to use only the meta-GGA part of the functional to obtain the molecular orbitals which yields a considerable acceleration of the total computation time. dRPA75 just adds the direct RPA correlation to a PBE hybrid with 75% Fock exchange.

Both methods perform very well and are competitive to the best DHDFAs. For example, for an extensive benchmark set, the weighted mean absolute deviation is 2.2 kcal mol^{-1} for PWRB95, 1.6 kcal mol^{-1} for the best DHDFA (PWPB95) and 7.9 kcal mol^{-1} for dRPA@PBE.[125] Comparing the two RPA based DHDFAs, PWRB95 proves more robust for reaction energies when electron pairs get broken (for example, to compute atomisation energies). dRPA75 fails badly for these cases which is in common with standard RPA approaches. For a set of atomisation energies, PWRB95 yields a mean absolute deviation of 4.7 kcal mol^{-1} while dRPA75 yields 52.9 kcal mol^{-1} which is still almost 10 kcal mol^{-1} better than dRPA@PBE (62.2 kcal mol^{-1}). But for other types of reaction energies, both methods show equal performance.[125]

4 Important technical improvements to extend the applicability of DHDFAs

Despite the undeniable improvements by DHDFAs in terms of accuracy, they clearly come along with a drawback: the considerably increased computational cost due to the MP2 term which scales with $\mathcal{O}(N^5)$ in comparison of the $\mathcal{O}(N^3)$ to $\mathcal{O}(N^4)$ scaling of the underlying DFAs. But actually, this poses only a problem when none of the many methodical improvements for speeding up theses computations can be exploited. Here, DHDFAs especially benefit from the wealth of research dedicated to the MP2 method.

A very simple, yet quite effective approach for larger systems is the use of the frozen-core approximation for the computation of the correlation contribution. Inner shell electrons are assumed not to contribute significantly to the forming and breaking of bonds. Therefore, they are just left out of the correlation treatment. This approximation is not fully compatible with DFT where the whole electron density has to be considered. On the other hand, the frozen-core approximation cannot only

reduce computational costs but likewise the basis-set superposition error due to core-electron correlation.[127] Standard basis sets are not developed to describe the correlation of the inner electrons which requires additional core polarisation functions. When the latter are not included it is better to exclude the core-electrons from the correlation treatment.

One of the most efficient improvements which can be applied straight forwardly is the use of the resolution-of-the-identity (RI) method to replace the costly computation of two-electron-four-centre integrals by approximate three centre integrals only.[128] This is also known as density fitting. Although the general scaling of the method is *not* reduced, the computation time is scaled down considerably. The method is very robust and reliable and therefore available in many electronic structure codes. Whenever possible, it should be made use of.

It should be noted that an efficient and accurate approximation of the integrals requires reliable auxiliary basis sets but those have been fitted for many AO basis sets.[129–131] Further, to gain the most of the RI approximation, sufficient main memory should be provided but for modern computer architectures, this is only a limiting factor for the largest systems/basis sets. Recently, the RI approach has been successfully combined with the related Cholesky-decomposition method which reduced the scaling of the MP2 part to $\mathcal{O}(N^2)$ only, but the approach is not yet available in standard computational chemistry programs.[132]

Another strategy is based on the SOS scaling of the MP2 energy and use of the RI approximation. This allows for a Laplace transformation of the corresponding energy expression. The resulting integral can be evaluated by numerical quadrature techniques. This leads to a formal $\mathcal{O}(N^4)$ scaling but with a prefactor equal to the number of integration points, typically 10 to 15.[133] As a rough estimate, the method will be faster if the number of occupied orbitals (excluding frozen orbitals) exceeds six times the number of quadrature grid points.[134]

The efficiency of DHDFAs is also slightly affected when the wave function correlation part is evaluated based on a scaled electron–electron interaction (in opposite to a range-separation approach) because this introduces a larger basis set dependency. Higher angular momenta in the basis set are needed to model the electron cusp correctly. The effect is the same as for wave function methods where it is even more pronounced. It is rooted in the approximation of the wave function itself based on Slater determinants (or even Hartree products) because there is no explicit dependency on the electron–electron distance in such a wave function. Chan and Radom addressed the problem by suggesting an approach to estimate the DHDFA energy at the quadruple-ζ level based on smaller basis sets.[135]

Care for the basis set dependency can also be taken by making the wave function explicitly depend on the inter-electronic distance or on a function of that. The latter is now the standard approach in modern variants and led to the name F12 or explicitly correlated wave function methods.[136,137] All efficient implementations rely on a RI approach in addition. Although the many variants found in the literature all differ in detail, the general idea is to compute a basis set incompleteness

correction which is added to the MP2 energy correction. Of course, this has to be scaled accordingly to the parametrisation of the chosen DHDFA when applied in this context. The MP2-F12 method was evaluated for DHDFAs by Karton and Martin and they found that it yields energies with the same accuracy as those computed in a standard fashion with an one order of size larger basis set (so triple-ζ instead of double-ζ and so on).[138] Nevertheless, the MP2-F12 approach requires more computational resources than the standard approach which limits the actual gain in speed.

Karton and Martin have also noted that in a MP2-F12 approach, now the accuracy of the underlying global hybrid can be a problem whose energy is often not converged at a double-ζ basis set level. In principle, the F12 approach can be used also to account for the HF basis set incompleteness[136] but so far, the effect has not been investigated in detail for DHDFAs.

For really large molecular systems, none of the afore mentioned approaches (alone) is sufficient to compute them with DHDFAs. In these cases, one has to exploit the locality of the electron–electron correlation and thereby reduce the actual size of the mathematical problem. The ultimate goal are linear-scaling methods but also a reduced scaling in comparison to other methods would help. Examples are the density-fitted local MP2 variant (DF-LMP2),[139] the pair-natural orbital with orbital-specific virtuals MP2 (PNO-OSV-MP2),[140] the domain-local pair natural orbital MP2 (DLPNO-MP2),[141] and the divide–expand–consolidate RI-MP2 (DEC-RI-MP2).[142] All of these methods introduce some overhead so that the turnover point is normally quite late for increasing system sizes and they also introduce some further approximations to the actual MP2 energy thereby affecting the accuracy but this can be controlled to some extend *via* input parameters. Of course, the cruder the approximation the larger the speed up. Nevertheless, impressive applications have been presented. And when the methods are available, they can be applied straight forwardly for DHDFAs.

Another important development for MP2 methods were extensions of the approach to systems with periodic boundary conditions.[143,144] The extensions facilitate a very efficient treatment of condensed phase systems. A successful implementation requires a local variant of MP2 but the approximations are not more severe as for standard molecule applications. Accordingly, DHDFAs have also been assessed for the treatment of molecular crystals[144,145] and recently, a srDFT/lrMP2 variant has been developed and tested for this application case.[146] Liquid systems have been computed with RPA and a DHDFA to determine electrochemical energy levels.[147] In the latter case, much improvement over standard hybrid DFAs has been observed with RPA and the DHDFA, presumably because of the higher amount of Fock exchange which leads to better band gaps. For molecular crystals, DHDFAs perform very satisfying but a distinct improvement with respect to pure MP2 or other DFT approaches has not been found so far.

Finally, it should be noted that DHDFAs are suitable for molecular crystals and solute/solvent systems but should not be applied to

semiconductors or metallic systems. In the latter, the small or vanishing band gaps lead to a break down of the perturbation approach. In those cases, RPA or double hybrids derived from RPA (dRPA75,[126] PWRB95[125]) are better suited. Another alternative might be the renormalized finite-temperature second order many-body perturbation theory[148] but more experience with that method is lacking and it has not been tried with DHDFAs so far.

5 Outlook and conclusion

Since Grimme presented his concept of a DHDFA a decade ago, the whole field has seen a lot of attention. And although it is debatable if there has been any significant improvement in the method itself after it was combined with a decent dispersion correction, the theoretical foundations are now much better understood. Further, important extensions of the DHDFA approach have been made. Analytic gradients and second derivatives are now available and electronic excitation energies can be computed.

Of course, this does not mean that there is no more room for improvement. The general problems of DFT, all above the self-interaction error, are mitigated to a good extent but not solved. So, addressing the remaining problematic cases without deteriorating the general performance should be the strategy for further developments.

It is unlikely that this can be accomplished by just finding another pair of mixing parameters or trying an untested combination of existing exchange and correlation functionals together with the DHDFA approach. (Almost) all tricks known from MP2 have now also been applied to DHDFAs. And just bluntly adding even higher correlation corrections from WFT methods seems also not to be very promising. Probably, considerable improvements require a new design of the underlying functionals, not only a re-parametrisation of existing approaches but complete new strategies.

Another way to improve the method is its extension to further properties like, for example, excited state gradients or magnetic properties. There are still some worthwhile improvements to be made.

The many studies on DHDFAs also helped to better judge the approach. And after many assessments and several studies for vary diverse application cases, the general findings about DHDFAs still remain. They are among the most accurate and widely applicable DFT methods. They are always compatible and often better than related pure wave function methods of similar computational cost. If combined with standard techniques to accelerate the computations, they are insignificantly more expensive than standard hybrid DFAs. In short: the method has matured and should now be incorporated into the standard toolbox for computational chemists.

Abbreviations

DFA	density functional approximation
DFT	density functional theory
DF	density fitting

DHDFA	double-hybrid density functional approximation
GGA	general gradient approximation
GLKS-PT	Görling–Levy KS perturbation theory
HF	Hartree–Fock
KS	Kohn–Sham
LDA	local-density approximation
MP2	2nd order Møller-Plesset perturbation theory
OO	orbital optimised
PT2	2nd order perturbation theory
RI	resolution-of-the-identity
RPA	random-phase-approximation
RS	range-separated
SCF	self-consistent field
SCS	spin-component scaling
SOS	spin-opposite scaling
TD	time-dependent
WFT	wave function theory
XC	exchange-correlation

References

1 R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules,* Oxford University Press, Oxford, 1989.

2 W. Koch and C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, 2000.

3 I. Y. Zhang and X. Xu, *Int. Rev. Phys. Chem.*, 2011, **30**, 115–160.

4 J. C. Sancho-García and C. Adamo, *Phys. Chem. Chem. Phys.*, 2013, **15**, 14581–14594.

5 L. Goerigk and S. Grimme, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 576–600.

6 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Chem. Rev.*, 2012, **112**, 289–320.

7 S. Grimme, *J. Chem. Phys.*, 2006, **124**, 034108.

8 M. Ernzerhof, *Chem. Phys. Lett.*, 1996, **263**, 499–506.

9 Y. Zhao, B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 4786–4791.

10 J. P. Perdew, A. Ruzsinszky, L. A. Constantin, J. Sun and G. I. Csonka, *J. Chem. Theory Comput.*, 2009, **5**, 902–908.

11 T. Tsuneda and K. Hirao, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 375–390.

12 J. Jaramillo, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.*, 2003, **118**, 1068–1073.

13 L. A. Curtiss, P. C. Redfern and K. Raghavachari, *J. Chem. Phys.*, 2005, **123**, 124107.

14 P. Jurečka, J. Šponer, J. Černý and P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985–1993.

15 A. Karton, A. Tarnopolsky, J.-F. Lamère, G. C. Schatz and J. M. L. Martin, *J. Phys. Chem. A*, 2008, **112**, 12868–12886.

16 M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer and W. Thiel, *J. Chem. Phys.*, 2008, **129**, 104103.

17 D. Jacquemin, V. Wathelet, E. A. Perpète and C. Adamo, *J. Chem. Theory Comput.*, 2009, **5**, 2420–2435.

18 L. Goerigk and S. Grimme, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6670–6688.

19 L. Goerigk and S. Grimme, *J. Chem. Theory Comput.*, 2011, **7**, 291–309.

20 A. Görling and M. Levy, *Phys. Rev. B*, 1993, **47**, 13105–13113.

21 A. Görling and M. Levy, *Phys. Rev. A*, 1994, **50**, 196–204.

22 C. Filippi, C. J. Umrigar and X. Gonze, *J. Chem. Phys.*, 1997, **107**, 9994–10002.

23 E. Fromager, *J. Chem. Phys.*, 2011, **135**, 244106.

24 K. Sharkas, J. Toulouse and A. Savin, *J. Chem. Phys.*, 2011, **134**, 064113.

25 S. Grimme, L. Goerigk and R. F. Fink, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 886–906.

26 T. Schwabe, *Ph.D. thesis*, Department of Chemistry and Pharmacy, University of Münster, 2009.

27 S. Kozuch and J. M. L. Martin, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20104–20107.

28 S. Kozuch and J. M. L. Martin, *J. Comput. Chem.*, 2013, **34**, 2327–2344.

29 T. Benighaus, R. A. DiStasio, R. C. Lochan, J.-D. Chai and M. Head-Gordon, *J. Phys. Chem. A*, 2008, **112**, 2702–2712.

30 Y. Zhang, X. Xu and W. A. Goddard, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 4963–4968.

31 I. Y. Zhang, N. Q. Su, É. A. G. BrÉmond, C. Adamo and X. Xu, *J. Chem. Phys.*, 2012, **136**, 174103.

32 N. Q. Su and X. Xu, *Mol. Phys.*, 2015, **114**, 1207–1217.

33 S. M. O. Souvi, K. Sharkas and J. Toulouse, *J. Chem. Phys.*, 2014, **140**, 084107.

34 E. Brémond and C. Adamo, *J. Chem. Phys.*, 2011, **135**, 024106.

35 J. Toulouse, K. Sharkas, E. Brémond and C. Adamo, *J. Chem. Phys.*, 2011, **135**, 101102.

36 P. Cortona, *J. Chem. Phys.*, 2012, **136**, 086101.

37 N. Q. Su and X. Xu, *J. Chem. Phys.*, 2014, **140**, 18A512.

38 J. Kim and Y. Jung, *J. Chem. Theory Comput.*, 2015, **11**, 45–54.

39 É. Brémond, J. C. Sancho-García, Á. J. Pérez-Jiménez and C. Adamo, *J. Chem. Phys.*, 2014, **141**, 031101.

40 M. Alipour, *Theor. Chem. Acc.*, 2015, **134**, 1–8.

41 É. Brémond, M. Savarese, J. C. Sancho-García, Á. J. Pérez-Jiménez and C. Adamo, *J. Chem. Phys.*, 2016, **144**, 124104.

42 M. Alipour, *J. Phys. Chem. A*, 2016, **120**, 3726–3730.

43 Y. Cornaton, O. Franck, A. M. Teale and E. Fromager, *Mol. Phys.*, 2013, **111**, 1275–1294.

44 T. Schwabe, *Phys. Chem. Chem. Phys.*, 2014, **16**, 14559–14567.

45 D. Bousquet, E. Brémond, J. C. Sancho-García, I. Ciofini and C. Adamo, *J. Chem. Theory Comput.*, 2013, **9**, 3444–3452.

46 M. Wykes, A. J. Pérez-Jiménez, C. Adamo and J. C. Sancho-García, *J. Chem. Phys.*, 2015, **142**, 224105.

47 T. Schwabe and S. Grimme, *Acc. Chem. Res.*, 2008, **41**, 569–579.

48 B. Chan, Y. Kawashima, M. Katouda, T. Nakajima and K. Hirao, *J. Am. Chem. Soc.*, 2016, **138**, 1420–1429.

49 L.-J. Yu, F. Sarrami, R. J. O'Reilly and A. Karton, *Chem. Phys.*, 2015, **458**, 1–8.

50 J. Sun, A. Ruzsinszky and J. P. Perdew, *Phys. Rev. Lett.*, 2015, **115**, 036402.

51 K. Hui and J.-D. Chai, *J. Chem. Phys.*, 2016, **144**, 044114.

52 S. Grimme, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 211–228.

53 S. Grimme, A. Hansen, J. G. Brandenburg and C. Bannwarth, *Chem. Rev.*, 2016, **116**, 5105–5154.

54 T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.*, 2007, **9**, 3397–3406.

55 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

56   H. Schröder, A. Creon and T. Schwabe, *J. Chem. Theory Comput.*, 2015, **11**, 3163–3170.

57   S. Pakhira, C. Sahu, K. Sen and A. K. Das, *Struct. Chem.*, 2012, **24**, 549–558.

58   S. Pakhira, K. Sen, C. Sahu and A. K. Das, *J. Chem. Phys.*, 2013, **138**, 164319.

59   D. Bousquet, E. Brémond, J. C. Sancho-García, I. Ciofini and C. Adamo, *Theor. Chem. Acc.*, 2014, **134**, 1–13.

60   O. A. Vydrov and T. V. Voorhis, *J. Chem. Phys.*, 2010, **133**, 244103.

61   J. Aragó, E. Ortí and J. C. Sancho-García, *J. Chem. Theory Comput.*, 2013, **9**, 3437–3443.

62   F. Yu, *J. Chem. Theory Comput.*, 2014, **10**, 4400–4407.

63   J. Calbo, E. Ortí, J. C. Sancho-García and J. Aragó, *J. Chem. Theory Comput.*, 2015, **11**, 932–939.

64   F. Neese, T. Schwabe and S. Grimme, *J. Chem. Phys.*, 2007, **126**, 124115.

65   M. Biczysko, P. Panek, G. Scalmani, J. Bloino and V. Barone, *J. Chem. Theory Comput.*, 2010, **6**, 2115–2125.

66   I. Carnimeo, C. Cappelli and V. Barone, *J. Comput. Chem.*, 2015, **36**, 2271–2290.

67   H. Ji, Y. Shao, W. A. Goddard and Y. Jung, *J. Chem. Theory Comput.*, 2013, **9**, 1971–1976.

68   S. Grimme and M. Steinmetz, *Phys. Chem. Chem. Phys.*, 2013, **15**, 16031–16042.

69   M. Wykes, N. Q. Su, X. Xu, C. Adamo and J.-C. Sancho-García, *J. Chem. Theory Comput.*, 2015, **11**, 832–838.

70   Y.-Y. Sun, J.-F. Li, M.-M. Li, F.-Q. Zhou, J.-L. Li and B. Yin, *J. Chem. Phys.*, 2016, **144**, 054303.

71   M. K. Kesharwani, B. Brauer and J. M. L. Martin, *J. Phys. Chem. A*, 2015, **119**, 1701–1714.

72   V. Barone, M. Biczysko, J. Bloino, P. Cimino, E. Penocchio and C. Puzzarini, *J. Chem. Theory Comput.*, 2015, **11**, 4342–4363.

73   F. Vazart, C. Latouche, P. Cimino and V. Barone, *J. Chem. Theory Comput.*, 2015, **11**, 4364–4369.

74   T. Fornaro, M. Biczysko, J. Bloino and V. Barone, *Phys. Chem. Chem. Phys.*, 2016, **18**, 8479–8490.

75   J. C. Howard, J. D. Enyard and G. S. Tschumper, *J. Chem. Phys.*, 2015, **143**, 214103.

76   M. Alipour, *J. Phys. Chem. A*, 2013, **117**, 2884–2890.

77   M. Alipour, *J. Phys. Chem. A*, 2013, **117**, 4506–4513.

78   M. Alipour, *J. Phys. Chem. A*, 2014, **118**, 5333–5342.

79   M. Alipour, *Chem. Phys. Lett.*, 2016, **644**, 163–166.

80   M. de Wergifosse and B. Champagne, *J. Chem. Phys.*, 2011, **134**, 074113.

81   F. Castet and B. Champagne, *J. Chem. Theory Comput.*, 2012, **8**, 2044–2052.

82   E. Fromager, S. Knecht and H. J. A. Jensen, *J. Chem. Phys.*, 2013, **138**, 084101.

83   M. Head-Gordon, R. J. Rico, M. Oumi and T. J. Lee, *Chem. Phys. Lett.*, 1994, **219**, 21–29.

84   S. Grimme and F. Neese, *J. Chem. Phys.*, 2007, **127**, 154116.

85   L. Goerigk, J. Moellmann and S. Grimme, *Phys. Chem. Chem. Phys.*, 2009, **11**, 4611–4620.

86   L. Goerigk and S. Grimme, *J. Chem. Phys.*, 2010, **132**, 184103.

87   R. Send, O. Valsson and C. Filippi, *J. Chem. Theory Comput.*, 2011, **7**, 444–455.

88   L. Goerigk and S. Grimme, *J. Chem. Theory Comput.*, 2011, **7**, 3272–3277.

89   A. Prlj, M. E. Sandoval-Salinas, D. Casanova, D. Jacquemin and C. Corminboeuf, *J. Chem. Theory Comput.*, 2016, **12**, 2652–2660.

90   F. D. Meo, P. Trouillas, C. Adamo and J. C. Sancho-García, *J. Chem. Phys.*, 2013, **139**, 164104.

91   M. Alipour, *Theor. Chem. Acc.*, 2016, **135**, 1–8.

92   M. R. Momeni and A. Brown, *J. Phys. Chem. A*, 2016, **120**, 2550–2560.

93   L. Goerigk and S. Grimme, *J. Phys. Chem. A*, 2009, **113**, 767–776.

94   E. Rebolini, J. Toulouse, A. M. Teale, T. Helgaker and A. Savin, *Mol. Phys.*, 2015, **113**, 1740–1749.

95   R. C. Lochan and M. Head-Gordon, *J. Chem. Phys.*, 2007, **126**, 164101.

96   R. Peverati and M. Head-Gordon, *J. Chem. Phys.*, 2013, **139**, 024110.

97   J. C. Sancho-García, A. J. Pérez-Jiménez, M. Savarese, E. Brémond and C. Adamo, *J. Phys. Chem. A*, 2016, **120**, 1756–1762.

98   F. Neese, T. Schwabe, S. Kossmann, B. Schirmer and S. Grimme, *J. Chem. Theory Comput.*, 2009, **5**, 3060–3073.

99   L. M. Thompson and H. P. Hratchian, *J. Chem. Phys.*, 2014, **141**, 034108.

100 A. Savin and H.-J. Flad, *Int. J. Quantum Chem.*, 1995, **56**, 327–332.

101 B. Miehlich, H. Stoll and A. Savin, *Mol. Phys.*, 1997, **91**, 527–536.

102 J. G. Ángyán, I. C. Gerber, A. Savin and J. Toulouse, *Phys. Rev. A*, 2005, **72**, 012510.

103 I. Y. Zhang and X. Xu, *J. Phys. Chem. Lett.*, 2013, **4**, 1669–1675.

104 J.-D. Chai and M. Head-Gordon, *J. Chem. Phys.*, 2009, **131**, 174105.

105 Y. Cornaton, A. Stoyanova, H. J. A. Jensen and E. Fromager, *Phys. Rev. A*, 2013, **88**, 022516.

106 Y. Cornaton and E. Fromager, *Int. J. Quantum Chem.*, 2014, **114**, 1199–1211.

107 E. Fromager and H. J. A. Jensen, *J. Chem. Phys.*, 2011, **135**, 034116.

108 E. Goll, H.-J. Werner and H. Stoll, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3917–3923.

109 E. Goll, M. Ernst, F. Moegle-Hofacker and H. Stoll, *J. Chem. Phys.*, 2009, **130**, 234112.

110 E. Goll, H.-J. Werner and H. Stoll, *Z. Physik. Chem.*, 2010, **224**, 481–491.

111 E. Goll, H. Stoll, C. Thierfelder and P. Schwerdtfeger, *Phys. Rev. A*, 2007, **76**, 032507.

112 H. Eshuis, J. E. Bates and F. Furche, *Theor. Chem. Acc.*, 2012, **131**, 1–18.

113 H. Eshuis, J. Yarkony and F. Furche, *J. Chem. Phys.*, 2010, **132**, 234114.

114 B. G. Janesko, T. M. Henderson and G. E. Scuseria, *J. Chem. Phys.*, 2009, **130**, 081105.

115 B. G. Janesko, T. M. Henderson and G. E. Scuseria, *J. Chem. Phys.*, 2009, **131**, 034110.

116 B. G. Janesko and G. E. Scuseria, *J. Chem. Phys.*, 2009, **131**, 154106.

117 R. M. Irelan, T. M. Henderson and G. E. Scuseria, *J. Chem. Phys.*, 2011, **135**, 094105.

118 J. Toulouse, I. C. Gerber, G. Jansen, A. Savin and J. G. Ángyán, *Phys. Rev. Lett.*, 2009, **102**, 096404.

119 W. Zhu, J. Toulouse, A. Savin and J. G. Ángyán, *J. Chem. Phys.*, 2010, **132**, 244108.

120 J. Toulouse, W. Zhu, J. G. Ángyán and A. Savin, *Phys. Rev. A*, 2010, **82**, 032502.

121 J. Toulouse, W. Zhu, A. Savin, G. Jansen and J. G. Ángyán, *J. Chem. Phys.*, 2011, **135**, 084119.

122 A. J. Garza, I. W. Bulik, T. M. Henderson and G. E. Scuseria, *Phys. Chem. Chem. Phys.*, 2015, **17**, 22412–22422.

123 A. J. Garza, I. W. Bulik, A. G. S. Alencar, J. Sun, J. P. Perdew and G. E. Scuseria, *Mol. Phys.*, 2015, **114**, 997–1018.

124 B. Chan, L. Goerigk and L. Radom, *J. Comput. Chem.*, 2015, **37**, 183–193.

125 S. Grimme and M. Steinmetz, *Phys. Chem. Chem. Phys.*, 2016, DOI: 10.1039/ c5cp06600j.

126 P. D. Mezei, G. I. Csonka, A. Ruzsinszky and M. Kállay, *J. Chem. Theory Comput.*, 2015, **11**, 4615–4626.

127 T. Schwabe, *J. Phys. Chem. A*, 2013, **117**, 2879–2883.

128 M. Feyereisen, G. Fitzgerald and A. Komornicki, *Chem. Phys. Lett.*, 1993, **208**, 359–363.

129 F. Weigend, A. Köhn and C. Hättig, *J. Chem. Phys.*, 2002, **116**, 3175–3183.

130 C. Hättig, *Phys. Chem. Chem. Phys.*, 2005, **7**, 59–66.

131 A. Hellweg, C. Hättig, S. Höfener and W. Klopper, *Theor. Chem. Acc.*, 2007, **117**, 587–597.

132 S. A. Maurer, L. Clin and C. Ochsenfeld, *J. Chem. Phys.*, 2014, **140**, 224112.

133 Y. Jung, R. C. Lochan, A. D. Dutoi and M. Head-Gordon, *J. Chem. Phys.*, 2004, **121**, 9793–9802.

134 N. O. C. Winter and C. Hättig, *J. Chem. Phys.*, 2011, **134**, 184101.

135 B. Chan and L. Radom, *Theor. Chem. Acc.*, 2013, **133**, 1–10.

136 C. Hättig, W. Klopper, A. Köhn and D. P. Tew, *Chem. Rev.*, 2012, **112**, 4–74.

137 L. Kong, F. A. Bischoff and E. F. Valeev, *Chem. Rev.*, 2012, **112**, 75–107.

138 A. Karton and J. M. L. Martin, *J. Chem. Phys.*, 2011, **135**, 144119.

139 H.-J. Werner, F. R. Manby and P. J. Knowles, *J. Chem. Phys.*, 2003, **118**, 8149–8160.

140 G. Schmitz, B. Helmich and C. Hättig, *Mol. Phys.*, 2013, **111**, 2463–2476.

141 P. Pinski, C. Riplinger, E. F. Valeev and F. Neese, *J. Chem. Phys.*, 2015, **143**, 034108.

142 P. Baudin, P. Ettenhuber, S. Reine, K. Kristensen and T. Kjærgaard, *J. Chem. Phys.*, 2016, **144**, 054102.

143 C. Pisani, L. Maschio, S. Casassa, M. Halo, M. Schütz and D. Usvyat, *J. Comput. Chem.*, 2008, **29**, 2113–2124.

144 M. D. Ben, J. Hutter and J. VandeVondele, *J. Chem. Theory Comput.*, 2012, **8**, 4177–4188.

145 K. Sharkas, J. Toulouse, L. Maschio and B. Civalleri, *J. Chem. Phys.*, 2014, **141**, 044105.

146 G. Sansone, B. Civalleri, D. Usvyat, J. Toulouse, K. Sharkas and L. Maschio, *J. Chem. Phys*, 2015, **143**, 102811.

147 J. Cheng and J. VandeVondele, *Phys. Rev. Lett.*, 2016, **116**, 086402.

148 S. Hirata, X. He, M. R. Hermes and S. Y. Willow, *J. Phys. Chem. A*, 2014, **118**, 655–672.

DFT calculations of atoms and molecules in Cartesian grids

Abhisek Ghosal and Amlan K. Roy*

1 Introduction

The past several decades have witnessed a proliferation of *ab initio* quantum mechanical methods for elucidation of electronic *structure, properties and dynamics* of many-electron systems such as atoms, molecules, clusters, solids. With rapid advances in mathematical formalisms and smart numerical algorithms, as well as computer architecture and resources, lately, some of these have been extended to the case of materials science, nano-science and biological science, *etc.*, with unprecedented success. The *electronic* Schrödinger equation (SE) of such an *N*-electron system is essentially a many-body problem involving space, spin and time coordinates. This usually causes enormous complexity and hence, *exact* analytical solutions are unavailable in vast majority cases, leaving aside a few countable idealized model problems which offer such solutions. Thus almost invariably recourse must be taken to approximate methods. So a well-defined theoretical framework of approximation is indeed necessary to explore the electronic problem in a systematic way as accurate as possible. Nowadays, it is not a mere contemplation to obtain an optimal solution for a system containing 100 or more electrons, with physically and chemically meaningful accuracy. However, as the number of electrons increases, computational complexity grows up exponentially, which requires new paradigms of development regarding basic theory, methodology, mathematical algorithm, sophisticated computer code and practical computational implementation. The ultimate objective is to probe bigger and bigger systems with as much accuracy and efficiency, as possible. Significant strides have been made in last few decades in these aforementioned areas, and in the coming years, a lot of progress will be continued in order to reach the eventual goal of theory predicting experimental results in an explanatory level relevant to practical applications.

Besides ground-state electronic structure, a lot of important aspects are involved in atomic physics to explain the experimental observations in a satisfactory level. Some of these, *e.g.*, regard excited states of discrete and continuum spectra, ionization dynamics of Rydberg states, spectroscopic levels, oscillator strengths for electronic transitions, fine-structures due to spin–orbit coupling, external perturbations due to static probes, multi-photon ionizations, above-threshold ionizations, electron scattering, relativistic corrections to the Hamiltonian and so on. But in case of

Department of Chemical Sciences, Indian Institute of Science Education and Research (IISER) Kolkata, Mohanpur-741246, Nadia, West Bengal, India.
E-mail: akroy@iiserkol.ac.in; akroy6k@gmail.com

molecules, calculation of physical and chemical properties are somehow less straightforward and clouded, simply because of the presence of more than one nucleus. At a first glance, it requires a rigorous understanding of the molecular potential energy surface through construction of the Hessian matrix, and all other molecular properties are inherently dependent on it. Such a surface consists of a large number of maxima and minima at unknown locations, and the global minimum corresponds to the ground-state structure, whereas the paths between the minima are the primary focus of chemical reactions. It is worthwhile to mention a few relevant molecular properties, such as dipole and multipole moments, vibrational frequencies, molecular dissociation energies, linear and non-linear responses to external static probes, molecular scattering through electron beams, molecular reaction dynamics, nonlinear reaction dynamics, photochemistry, *etc.* A significant amount of interest concerns with modern spectroscopic techniques, namely, electronic spectroscopy, photo-electron spectroscopy, ro-vibrational spectroscopy, nuclear magnetic resonance, nuclear quadrupole resonance, electron spin resonance, Mössbauer spectroscopy *etc.*

An integral consideration in modern quantum chemistry is the obligatory computational work, which has profound impact on future theoretical development. In order to perform the complex mathematical operations for larger systems within a suitable framework, it is imperative to have a systematic way of computation, and sufficient intricacy to capture the characteristic subtle interactions. Sometimes the computational output may be very crude (overestimating or underestimating), and other times it could be within the bound accuracy, but that is mostly dictated by the nature of the practical problem at hand. Now, in order for a given computational method to be successful and viable, the two deciding factors are its *accuracy and cost effectiveness*. In today's computational chemistry, it is desirable to achieve the energy of a chemical reaction within the bounds of chemical accuracy, which is less than 1 kcal mole^{-1}. And the main challenge is to lower computational cost without much compromise on accuracy for increasingly bigger and complicated systems. Starting from relatively less accurate and dependable empirical or semi-empirical schemes to modern *ab initio* methods, all fall under the purview of current computational quantum chemistry repertoire. The latter covers a large array of methods including Hartree–Fock (HF), post-HF, multi-reference, quantum Monte-Carlo, density functional theory (DFT), *etc.*, each providing a different computational cost depending on the particular problem. In order to broaden the domain of feasibility to larger systems, significant developments have been made towards the so-called linear-scaling approaches, which slash computational scaling very swiftly with molecular size. In this context, it is appropriate to mention some of the most popular and prominent electronic structure packages, which are routinely used by the scientific community, such as Gaussian, GAMESS, ACES, CADPAC, Dalton, Hyperchem, Jaguar, Molcas, Molpro, NWChem, PQS, Psi4, Q-Chem, Spartan, TURBOMOLE, UTChem, *etc.* On the other hand, one could take help of molecular mechanics which uses classical mechanics to model a

molecular, as well as large biological system or material cluster, having much lower computational cost, but is also less sophisticated and reliable. These are usually applied in the field of molecular dynamics (in the context of thermodynamic and kinetic properties of a system) and potential energy minimization (*e.g.*, to calculate binding constants, protein folding kinetics, protonation equilibria, active site coordinates). In recent years, computational chemistry has also been used vigorously in medicinal chemistry, leading to the discovery of new life-saving drugs. Essentially this helps to model the synthetic compounds in a systematic way, providing some critical checks on huge labor and chemical cost. A large number of beautiful and in-depth books, reviews and articles are available on the subject. Some of them are referred here.[1–17]

In order to solve the electronic SE for an interacting system, two distinct well-established theoretical formalisms have gained popularity and credence in the last few decades, with contributions from physicists and chemists alike. We begin our discussion with a comparatively simple HF theory, which in a sense, is founded on a mean-field approximation. In essence, the instantaneous inter-electronic interaction is accounted for, in an average way, where each electron sees an average field because of the presence of the other electrons, leading to the optimization of a single Slater determinantal wave function consisting of individual spin orbitals. Though it offers a crucial first step towards the problem, its application is limited due to a complete neglect of electron correlation effects. Post HF formalisms are mainly concerned on how this subtle effects could be incorporated in an efficient manner. This family includes a variety of versatile formalisms, such as Møller-Plesset perturbation theory (MPn),[18] configuration interaction (CI),[2] coupled-cluster ansatz (they come in many flavours like CCSD, CCSD(T), CCSDT, CCSDTQ, *etc.*),[19] multi-reference perturbation theory (such as CASPT2),[20] *etc.* These methods offer potentially authentic and reliable results, but are quite difficult to be implemented computationally for large N, primarily because of their unfavorable scaling. They differ in scaling, cost, efficiency and accuracy; naturally a major research effort for such methods is directed towards betterment of these issues. They have their own space of applicability depending upon a given problem and often provide a benchmark of accuracy for the system under investigation.

The second approach relies on utilizing the limited information coming from electron density, density matrix or Green's function in an optimized way with the help of the variational principle. Amongst them, however, DFT has appeared as the most versatile and outstanding method in electronic structure calculation. Many beautiful books and reviews are available on the subject; some recent ones are ref. 7, 21–35. The basic variable, the single-particle electron density, $\rho(\mathbf{r})$, is a real-valued and experimentally realizable quantity, in contrast to the traditional complex-valued N-electron wave function. Furthermore, the most attractive part is that it incorporates the complicated electron correlation effects in a transparent and elegant way. Being a single determinant approach, it is computationally more tractable; the burden remains at same level as HF. However, this apparent simplicity and novelty is

somehow tainted and compensated by the hitherto unavailable, all-important, universal, exchange-correlation (XC) density functional. Although uncovering its exact form constitutes one of the most active and fertile area of research, unfortunately it still remains elusive; hence it must be approximated. Success of DFT lies heavily on the choice of these approximate XC functional. Historically, the journey started with the so-called local density approximation (LDA) in 1930 by Dirac[36] using the form of exchange assuming a uniform electron gas model. Later on, a working LDA correlation part was developed.[37,38] However, the LDA XC functional suffers from serious problems in atoms, molecules and solids, as the density of such systems are far apart from the uniform electron gas. Since it is essential to have a good knowledge of the gradient of the electron density, a whole family of so-called generalized gradient approximation (GGA)-based functionals were proposed over the years. A few prominent ones are: B88,[39] LYP,[40] PBE[41] *etc.* In the next step, higher-order derivatives of the density including the kinetic energy density lead to the development of meta-GGA kind of functionals, such as B88C,[42] Becke-Rousse,[43] TPSS[44] *etc.* Next crucial step lead to hybrid functionals, which bring exact HF exchange into the picture. Depending on the co-factor of different kinds of XC functionals like LDA, GGA with latter, a variety of hybrid functionals have been published, which include one of the most popular and versatile candidates in quantum chemistry, namely, B3LYP.[39,40] The LDA functional typically does not have an impressive energetic performance, but interestingly it executes better performance than GGA functional regarding geometry optimization. Generally it is true that hybrid functionals work better than LDA due to incorporation of HF exchange. But in case of transition-metal chemistry and response properties, hybrid functionals show rather quite poor performance. Also, a majority of these approximate functionals has some sort of deficiency regarding activation energies of chemical reactions, which is somehow circumvented in the recently developed range-separated, density fitting, adiabatic-connection, local hybrid functionals or those involving unoccupied orbitals and eigenvalues, *etc.* These are found to be largely good for dispersion and van der Waals interactions. A vast amount of literature exists on the topic; the interested reader may look at the lucid reviews[31,34,35] and references therein.

Triumphs of DFT is now well-established in quantum chemistry and condensed matter physics; so much so that nowadays it is the most dominant and visible workforce. For large scale computation of materials, two distinct numerical approaches are engaged to solve the relevant many-electron equation. The condensed matter community chooses plane-wave basis sets, where periodic boundary conditions are appropriate. On the other hand, chemists prefer a localized atom-centred Gaussian basis set, because of being fondly attached with molecules (non-periodic systems), where the problematic multi-center integrals can be evaluated analytically. With advances in computational facilities, modern DFT can provide results with benchmark accuracy accompanied with experimental results at least for small molecules in gaseous phase. Throughout the past several decades, a wide variety of noteworthy

applications were made regarding molecular properties (including structures, thermo-chemistry, various spectroscopic quantities, responses to external perturbations), bulk and surface properties of solids, band-gaps and optical properties of solids, interactions of small molecules with surfaces with focus on structures, binding energies and catalytic chemistry, modelling of photochemical reactions, *etc.* Of late, extensive applications were reported in some other areas, such as nano-technology, semiconductor quantum dots, quantum confinement imposed by various kinds of nano structures. Practical applications from biological fields include modelling of enzymatic catalysis and active-site chemistry, cooperative activity in backbone of hydrogen bonding, modelling of beta-sheet formation, enzyme functioning *etc.*

The present chapter gives an account of the work done in our laboratory for static, non-relativistic ground states in many-electron atoms/molecules. Within the Born–Oppenheimer approximation and the Hohenberg–Kohn–Sham framework, an implementation of DFT in a Cartesian Coordinate grid (CCG) is offered. By using a linear combination of Gaussian functions, molecular orbitals (MO) and quantities like basis functions, electron density as well as classical Hartree and non-classical XC potentials are constructed on a 3D real CCG directly.[45–49] No additional auxiliary basis is used for the charge density. A Fourier convolution method, involving a combination of Fast Fourier transform (FFT) and inverse FFT [50,51] is used to obtain the Coulomb potential quite accurately. Analytical one-electron Hay-Wadt-type effective core potentials,[52,53] made of sums of Gaussian type functions, are used to represent the inner core electrons whereas energy-optimized truncated Gaussian bases are used for valence electrons. Viability and suitability of this simpler grid is compared and contrasted with a routinely used atom-centered grid (ACG) for a decent number of atoms, molecules. This is demonstrated by presenting the total energy, energy components, orbital energy, potential energy curve, atomization energy for a bunch of local and non-local XC functionals.

2 The methodology

2.1 Electron density as basic variable

In the usual wave function-based approach, the difficulty to solve the many-electron SE increases rapidly with the number of electrons present in the system. In order to reduce the dimensionality of the problem, one can introduce the concept of reduced density matrices.[54] One- and two-particle reduced density matrices, which are adequate for the evaluation of expectation values of one- and two-particle operators, are defined as,

$$\rho_1(\mathbf{x}_1' \,|\, \mathbf{x}_1) = N \int \cdots \int \psi(\mathbf{x}_1', \mathbf{x}_2, \ldots, \mathbf{x}_N) \psi^\star(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \mathrm{d}\mathbf{x}_2 \cdots \mathrm{d}\mathbf{x}_N \quad (1)$$

$$\tau_2(\mathbf{x}_1', \mathbf{x}_2' \,|\, \mathbf{x}_1, \mathbf{x}_2) = \frac{N(N-1)}{2} \int \cdots \int \psi(\mathbf{x}_1', \mathbf{x}_2', \mathbf{x}_3, \ldots, \mathbf{x}_N)$$
$$\times \psi^\star(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_N) \mathrm{d}\mathbf{x}_3 \cdots \mathrm{d}\mathbf{x}_N \quad (2)$$

where each of ($\{x_i\}$; $i = 1, 2, 3, \ldots, N$) consists of both spatial coordinates $\{\mathbf{r}_i\}$ as well as spin coordinates (σ_i). Now, the spin-less, single-particle density $\rho(\mathbf{r})$, which is the diagonal element of one-particle density matrix, is given as,

$$\rho(\mathbf{r}) = N \int \cdots \int \psi(\mathbf{r}\sigma, \mathbf{x}_2, \ldots, \mathbf{x}_N) \psi^*(\mathbf{r}\sigma, \mathbf{x}_2, \ldots, \mathbf{x}_N) \, d\sigma d\mathbf{x}_2 \cdots d\mathbf{x}_N. \quad (3)$$

In case of fermions, it actually describes a three-dimensional distribution of electrons in a system. In sharp contrast to the complex-valued wave function, $\rho(\mathbf{r})$ is a function of only three coordinates, irrespective of the number of electrons present in it. As it is a fundamental physical observable and can be determined experimentally, this enables one to test directly the veracity of chemical calculations and approximations. The beauty of DFT lies in formulating a many-particle problem within a single-particle framework with the aid of density being considered as a central quantity.

The first seminal work in this direction was provided independently by Thomas[55] and Fermi.[56,57] They proposed a model for calculating atomic properties based on $\rho(\mathbf{r})$ alone. For the kinetic energy, a local density approximation of a uniform, homogeneous electron gas model was suggested. On the other hand, the electron–electron repulsion energy was obtained from the classical Coulomb potential. This leads to the familiar Thomas–Fermi (TF) Euler–Lagrange equation for the density as,

$$\mu_{\text{TF}} = \frac{\delta E_{\text{TF}}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} = \frac{5}{3} C_F \rho(\mathbf{r})^{\frac{2}{3}} + \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + v_{\text{ext}}(\mathbf{r}), \quad (4)$$

where $C_F = \frac{3}{10}(3\pi^2)^{\frac{2}{3}}$, $v_{\text{ext}}(\mathbf{r})$ signifies the external potential at point $\mathbf{r}$ consisting of nuclear-electron attraction and μ_{TF} is, as usual, the Lagrange multiplier, to be identified as chemical potential. Thus we have the first genuine density functional for the energy of an interacting system. Though TF theory approximately describes charge density, electrostatic potential and total energy (at least for atoms and periodic solids), there are serious deficiencies due to a lack of proper description of the charge density at outer regions of atoms. Further, it is unable to produce the periodic variation of many properties with changing atomic number Z, due to the lack of a shell structure in the atoms. Besides, it does not bind atoms (neutral or ionic) to form molecules and solids. In a way, its utility is very limited to atomic systems.

Within this approximate framework, Dirac[36] first made a correction to the TF energy, incorporating exchange effects into it. He derived an exchange term from exchange energy of a homogeneous electron gas of density $\rho(\mathbf{r})$ by reforming the HF theory solely in terms of *density function*. This gives the *Thomas–Fermi–Dirac* (TFD) equation as,

$$E_{\text{TFD}}[\rho(\mathbf{r})] = C_F \int \rho(\mathbf{r})^{\frac{5}{3}} d\mathbf{r} + \int \rho(\mathbf{r}) v_{\text{ext}}(\mathbf{r}) d\mathbf{r} - C_x \int \rho(\mathbf{r})^{\frac{4}{3}} d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (5)$$

where $C_x = \dfrac{3}{4}\left(\dfrac{3}{\pi}\right)^{\frac{1}{3}}$. The second term corresponds to the external potential energy containing the nuclear-electron attraction, and the last expression refers to the classical electrostatic Hartree repulsion energy. While the sheer simplicity of replacing the troublesome many-electron SE by a *single* equation in terms of $\rho(\mathbf{r})$ *alone* is manifestly appealing, underlying approximations are crude and grossly inadequate to be of any practical use in quantum chemistry.

As overwhelming complications are met to advance beyond the primitive levels of the TFD model, this line of thought went into oblivion until 1964, when the future hopes were rekindled in a pioneering work by Hohenberg and Kohn (HK).[58] It laid a rigorous theoretical foundation of modern-day DFT by asking a plain obvious question as, whether the information contained in $\rho(\mathbf{r})$ is sufficient for elucidating a many-electron system completely. Following an amazingly simple proof, they answered the question in an affirmative way. In a landmark paper,[58] they first proved that a non-degenerate ground state of an N-particle system, moving under the influence of their mutual Coulomb repulsion and an external interacting potential $v_{\text{ext}}(\mathbf{r})$, is solely characterized by single-particle density $\rho(\mathbf{r})$, *i.e.*, $v_{\text{ext}}(\mathbf{r})$, $\psi(\mathbf{r})$ and hence all other ground-state properties are uniquely determined by $\rho(\mathbf{r})$ alone (or as unique functionals of $\rho(\mathbf{r})$). They also showed that, for a given $v_{\text{ext}}(\mathbf{r})$, the energy functional is minimum corresponding to the true density. Further, the total energy and density of a given system can be determined variationally by minimizing the functional $E[\rho(\mathbf{r})]$ subject to the normalization condition, $\int \rho(\mathbf{r})\,d\mathbf{r} = N$, as a constant, through the following equation,

$$\delta\left\{E[\rho(\mathbf{r})] - \mu\left[\int \rho(\mathbf{r})\,d\mathbf{r} - N\right]\right\} = 0, \tag{6}$$

where N is a measure of the total number of electrons present. This is the central equation of DFT providing a deterministic route for $\rho(\mathbf{r})$. Later, it has been established that the HK theorem is equally valid for degenerate ground states and has been extended to excited states.

Thus the one-to-one mapping between density and energy, the basic preamble of the HK theorem, can be succinctly expressed by conveniently writing the energy functional as follows,

$$E[\rho] = T[\rho(\mathbf{r})] + V_{\text{ee}}[\rho(\mathbf{r})] + V_{\text{ne}}[\rho(\mathbf{r})] = F[\rho] + \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})\,d\mathbf{r},$$

$$F[\rho] = T[\rho] + V_{\text{ee}}[\rho]. \tag{7}$$

Here $F[\rho]$ signifies the so-called *universal* energy density functional, independent of N and $v_{\text{ext}}(\mathbf{r})$, while T, V_{ne}, V_{ee} refer to the usual kinetic, electron-nuclear attraction and electron–electron repulsion energy operators in the Hamiltonian respectively. The original HK theorem implies that ground-state properties are functionals of v-representable densities only and not of any other trial densities. Later this restriction was lifted;

thus widening its domain of applicability from v-representable to N-representable densities (one that can be obtained from an anti-symmetric wave function).

In order to make sure that a given density is indeed the true ground-state density, now second HK theorem gives the crucial variational theorem,

$$E[\rho(\mathbf{r})] \geq E[\rho_0(\mathbf{r})], \tag{8}$$

where $E[\rho_0]$ corresponds to the ground-state energy of a Hamiltonian with $v_{\text{ext}}(\mathbf{r})$ as external potential, and ρ_0 is the true ground-state density. In other words, the functional attains its minimum value with respect to all allowed densities if and only if our input density corresponds to the true ground-state density. These two theorems, although, validate previous efforts of TFD to utilize $\rho(\mathbf{r})$ as a basic fundamental variable to describe many-electron systems, emphasis must be given to the fact that it is merely a *proof of existence* only. It confirms the much sought-after unique mapping between density and energy, in principle; however it remains absolutely quiet on any guidelines for construction of such a functional. Even though $\rho(\mathbf{r})$ is sufficient, the relation is delicate, intricate, and engenders tremendous complexity. Another major discomfiture lies in the fact that minimization of $E[\rho]$ is, in general, an unusually challenging numerical task. Nature of actual calculations before and after HK theorem, does not change in any noticeable way; for they are as hard as before. No clear-cut clue is available to approximate the unknown functional. Since this does not provide any simplification over MO theory, the ultimate step still remains to be the solution of an excessively clumsy SE equation. It is of very little comfort for chemistry and physics, leading to virtually no actual progress for realistic calculation. Another unpleasant predicament is that the density variation principle holds true only for *exact* functionals. This implies that, in contrast to standard wave-function-based approaches (such as HF or CI), where wave functions are rigorously variational, within the realm of HK DFT, energy offered by a trial functional has no physical meaning whatsoever.

2.2 The single-equation approach

At this stage, we will briefly mention about the so-called *single-equation* approach in DFT, which functions exclusively in terms of $\rho(\mathbf{r})$ alone precluding the consideration of orbitals. Thus this preserves the true spirit of DFT without sacrificing its associated physics. A real crux of problem for development of such methods lies exclusively on the unavailability of accurate (not to speak of exact) kinetic and XC functionals. Nevertheless two major attempts are noteworthy. The first one invokes a link between quantum fluid dynamics and DFT through $\rho(\mathbf{r})$. The former was proposed long times ago by Madelung, de Broglie, and Bohm based on a hydrodynamic formulation of quantum mechanics. It is more complicated than the Schrödinger picture, in a sense that the former requires solving a set of nonlinear partial differential equations, in contrast to linear equations in the latter. However the charm is that

this route has a conceptual appeal, for the electron cloud is treated as a "classical fluid" moving under the influence of classical Coulomb forces and an additional quantum potential. Moreover, amplitude and phase of the wave function (treated explicitly in the fluid formulation as independent variables) are more slowly varying in time than the wave function itself, which leads to computational advantages and offers additional physical insights of quantum dynamics or the quantum trajectory. Thus all the electrons, assumed to be distributed over a 3D space-like continuous fluid, being governed by two basic equations in terms of local variables, $\rho(\mathbf{r}, t)$ and the current density $\mathbf{j}(\mathbf{r}, t)$, can be eventually described by a set of equations, *viz.*, (atomic units used unless otherwise mentioned):

(i) an equation of continuity:

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0, \tag{9}$$

and (ii) an Euler-type equation of motion:

$$\frac{\partial \chi(\mathbf{r}, t)}{\partial t} + \frac{1}{2}(\nabla\chi)^2 + \frac{\delta G[\rho]}{\delta\rho} + \frac{\delta E_{\text{el--el}}[\rho]}{\delta\rho} + v_{\text{ext}}(\mathbf{r}, t) = 0, \tag{10}$$

where $\mathbf{j}(\mathbf{r}, t) = \rho\nabla\chi(\mathbf{r}, t)$, with $\chi(\mathbf{r}, t)$ being the velocity potential. $E_{\text{el--el}}$ represents the inter-electronic Coulomb repulsion energy, $v_{\text{ext}}(\mathbf{r}, t)$ is the external potential containing electron-nuclear attraction and any time-dependent (TD) interaction present in the system, while $G[\rho]$ is a universal functional comprising of kinetic and XC energy contributions. Eqn (9) and (10) can be combined into a single equation by defining a complex-valued TD hydrodynamical wave function for the entire time-evolving system as (in polar form):

$$\Psi(\mathbf{r}, t) = \rho(\mathbf{r}, t)^{1/2} e^{i\chi(\mathbf{r}, t)}, \tag{11}$$

and eliminating $\chi(\mathbf{r}, t)$ from them. The result is a generalized nonlinear SE:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}([\rho]; \mathbf{r}, t) \right]\Psi(\mathbf{r}, t) = i\frac{\partial \Psi(\mathbf{r}, t)}{\partial t}, \tag{12}$$

where $\rho(\mathbf{r}, t) = |\Psi(\mathbf{r}, t)|^2$, and the effective potential $v_{\text{eff}}([\rho]; \mathbf{r}, t)$ contains both classical and quantum potentials. This procedure has been successfully applied to a variety of interesting static and dynamic many-electron situations. Some of these are electronic structure calculations,[59] ion-atom collisions,[60,61] atoms, molecules in presence of strong TD electric and/or magnetic fields[62–67] including high-harmonic generation and multi-photon ionization, *etc.* Many more details could be found in the above references and therein.

The other one is the so-called orbital-free DFT,[68] whose realization most decisively depends on the availability of correct, good-quality

kinetic energy density functionals. Reportedly, it is possible to compute all components of energy functionals effectively in momentum space within a linear scaling process, which is certainly advantageous in terms of simplicity and computational cost. During the past several decades, an immense amount of work has been attempted towards designing suitable kinetic energy functionals[69] including a density dependent kernel[70] and non-local functionals,[71] to mention a few; mostly on periodic system. Efforts were also made to supplement the density-dependent kernel to non-periodic systems, with the aid of finite element modelling and coarse graining.[72] Later, its scope was enlarged to covalent systems and semiconductors[73] as well. Up to now, applications are predominantly limited to metallic system.

2.3 The Kohn–Sham method

While the single-equation route of the previous subsection provides an elegant and attractive direction to deal with an interacting system and has witnessed numerous successes, insurmountable difficulties (as delineated in Section 2.1) need to be overcome for its routine practical application. Unsurprisingly, a massive amount of today's work actually follows an alternative formalism, originally suggested by Kohn–Sham (KS). The prevailing unyielding state of affairs after publication of the HK theorem dramatically changed a year later, in a path-breaking contribution,[74] that suggested a manageable way to approach the hitherto unknown universal functional. In order to mitigate the problem, they introduced the clever concept of a *fictitious*, non-interacting system built from a set of orbitals (KS orbitals) such that the major part of the kinetic energy can be computed exactly. The remaining fairly small portion is absorbed in our non-classical contribution of the electron–electron repulsion, which is also unknown. However, the advantage is that electrons now move in an effective KS single-particle potential, while the mapped auxiliary system yields the same ground-state density as our real interacting system, but this simplifies the actual calculation tremendously.

This is accomplished by setting up a non-interacting reference system (designated with suffix 's') with the Hamiltonian containing an effective local potential $v_s(\mathbf{r})$ expressed as a sum of one-electron operators,

$$H_s = -\frac{1}{2}\sum_i^N \nabla_i^2 + \sum_i^N v_s(\mathbf{r}_i). \tag{13}$$

Note that there is no electron–electron interaction in the Hamiltonian. It is well known that the exact wave function of a system of non-interacting fermions can be represented as Slater determinants of individual one-electron eigenfunctions, and energies are sums of one-electron eigenvalues. In complete analogy to the HF equation, thus, one can write an eigenvalue equation for KS spin–orbitals as,

$$\hat{f}^{KS}\psi_i = \epsilon_i \psi_i, \tag{14}$$

where the one-electron KS operator $\hat{f}^{KS}$ is defined by,

$$\hat{f}^{KS} = -\frac{1}{2}\nabla^2 + v_s(\mathbf{r}). \tag{15}$$

Hence the crucial step is to map our artificial non-interacting system to the interacting, real one through an effective potential $v_s(\mathbf{r})$ which plays a pivotal role. This is ensured by choosing $v_s(\mathbf{r})$ such that the density resulting from a summation of moduli of square of orbitals ($\{\psi_i\}$), is exactly the same as the ground-state density of our real system (σ signifies spin), *i.e.*,

$$\rho_s(\mathbf{r}) = \sum_i^N \sum_\sigma |\psi_i(\mathbf{r},\sigma)|^2 = \rho_0(\mathbf{r}). \tag{16}$$

So the leading part of the kinetic energy of the real system can be restored adequately as a sum of individual kinetic energies of the reference system with same density as real system,

$$T_s = -\frac{1}{2}\sum_i^N \nabla_i^2. \tag{17}$$

However, since the non-interacting kinetic energy is not equal to the true kinetic energy, they suggested that the residual, kinetic energy (T_c), often a small contribution, is submerged to the unknown, non-classical component of electron–electron repulsion, namely,

$$F[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho],$$
$$E_{xc}[\rho] = (T[\rho] - T_s[\rho]) + (E_{ee}[\rho] - J[\rho]) = T_c[\rho] + E_{nc}[\rho]. \tag{18}$$

Associated terms have following meanings: $J[\rho]$ is the known classical part of $E_{ee}[\rho]$, whereas $E_{xc}[\rho]$ contains everything that is unknown, *i.e.*, non-classical electrostatic effects of $E_{ee}[\rho]$ as well as the difference between the true kinetic energy $T[\rho]$ and $T_s[\rho]$. Then the expression of total energy $E[\rho]$ of our real system can be cast in the following way,

$$E[\rho] = \int v_{ext}(\mathbf{r})\rho(\mathbf{r})\,d\mathbf{r} + J[\rho] + T_s[\rho] + E_{xc}[\rho] \tag{19}$$

This allows us to write the celebrated KS orbital equation in canonical form (henceforth atomic units employed unless otherwise stated),

$$\left[-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \tag{20}$$

where the "effective" potential contains following terms,

$$v_{eff}(\mathbf{r}) = v_{ext}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' + v_{xc}(\mathbf{r}). \tag{21}$$

In the above equation, $v_{xc}[\rho(\mathbf{r})]$ signifies the functional derivative, $\dfrac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}$, with respect to charge density. So the KS equation, although *exact* in principle, is structurally similar to the HF equation, but less complicated than the former. Both equations must be solved iteratively by self-consistent field. It has a profound effect on the quantum community because of its ability to account for XC effects in a rigorous, quantitative and transparent manner with a cost level of HF theory. However, it may be noted that, although being local in nature, the KS potential may be formally innocuous and less complicated than its counterpart in HF approximation, it may have a rather not-so-straight-forward and non-local dependence on $\rho(\mathbf{r})$.

Thus, as explained above, for practical purposes, minimization of our explicit energy functional, although possible, *in principle*, is hardly ever recommended. A far more viable and feasible route invokes solution of the KS eqn (20), which, however, interestingly brings back an orbital picture, and strictly speaking, does not function solely in terms of density. But, the good thing is that we have a mechanism to incorporate the intractable many-body effects formally *exactly*, within a single-particle theory.

2.4 Grid consideration

For realistic solution of the KS equation, it is necessary to deal with mathematically non-trivial integrals that cannot be evaluated analytically and pose certain amount of challenge. Even in finite basis-set expansion methods, there does not exist any explicit analytical formulas for requisite XC integrals. In absence of such procedures, one is left with no option but for numerical calculation. It is well acknowledged that such a discrete procedure for multi-center integrals in 3D space is not straightforward. The task becomes all the more formidable from a consideration of the fact that, in order to achieve a satisfactory level of chemical accuracy, a generous grid size and a longer computation time is often mandatory, as a prohibitively huge number of operations is needed. A vital problem for constructing such integration route arises due to cusps in the density and the singular nature of the Coulombic potential. Naturally, in order to pursue high-accuracy calculations within a reasonable number of quadrature grids, one needs efficient and sophisticated numerical integrators which can capture the forms of density at a satisfactory level. This paves the way for a considerable number of integrators having varied degrees of performance. Amongst them, two distinct, well recognized partitioning schemes have shown significant promise. The Voronoi cellular approach was originally proposed by te Velde and Baerends,[75] in which the space is divided into non-overlapping regions of simple geometry. On the other hand, fuzzy cells avenue, or commonly known as ACG, was initially recommended by Becke;[42] later championed by several others groups.[76–79] Though the basic idea was implemented for completely numerical, non-basis set calculations of polyatomic molecules, now it has reached the status of standard for

linear combination of atomic orbitals (LCAO)-MO DFT approach, which is convenient while using spherical grids.

The basic strategy is to divide the space into fuzzy (atom-centered overlapping analytically continuous regions) cells. In order to integrate over each cell, a spherical grid (corresponding to radial and angular quadrature grid) centered on each atom is the natural choice. The integral (I) of an arbitrary integrand $F(\mathbf{r})$ over whole region is then converted to discrete numerical summations over individual atomic regions j:

$$I = \int F(\mathbf{r})\,\mathrm{d}\mathbf{r} = \sum_j I_j = \sum_j \int F_j(\mathbf{r})\,\mathrm{d}\mathbf{r}, \tag{22}$$

where

$$F_j(\mathbf{r}) = w_j(\mathbf{r})F(\mathbf{r}) \tag{23}$$

i.e., a multi-center integration therefore translates to a sum of single-center integrations over each atomic nucleus, whereas the w_j's correspond to relative weight functions over each atomic center respectively. In order to make the discrete cell fuzzy, certain restrictions are imposed on w_j, so that,

$$\sum_j w_j(\mathbf{r}) = 1 \quad \forall \mathbf{r}, \tag{24}$$

and each $w_j(\mathbf{r})$ takes value unity in the neighborhood of its own nucleus, but vanishes near any other nucleus in a well-behaved manner.

Traditional cellular separation of molecular space is best realized by using confocal elliptical coordinates (λ, μ, ϕ). Taking center i as reference, one considers other centers $j \neq i$ and establishes λ_{ij}, μ_{ij}, ϕ_{ij} on the foci i, j. Of particular interest is,

$$\mu_{ij} = \frac{\mathbf{r}_i - \mathbf{r}_j}{R_{ij}} \tag{25}$$

where r_i, r_j identify distances to nuclei i, j respectively, and R_{ij} the inter-nuclear distance. In terms of an auxiliary function, $s(\mu_{ij})$, normalized weight functions, $w_i(\mathbf{r})$, are defined as:

$$w_i(\mathbf{r}) = \frac{W_i(\mathbf{r})}{\sum_j W_j(\mathbf{r})} ; \quad W_i(\mathbf{r}) = \prod_{j \neq i} s(\mu_{ij}). \tag{26}$$

The most recommended form of $s(\mu_{ij})$ is a cut-off function as given below,

$$s_k(\mu) = \frac{1}{2}[1 - f_k(\mu)], \tag{27}$$

where $f_k(\mu)$ indicates a k-th order iterative polynomial. The forms of polynomial, $f_k(\mu)$ and k are chosen such that $w_i(\mathbf{r})$ satisfies the restriction imposed by eqn (24). There is no hard and fast rule to evaluate the value

of k; it is chiefly decided through experience ($k = 3$ is used in some earlier works[42]). Instead of polynomial, if one chooses $s(\mu_{ij})$ as a step function, the fuzzy scheme transforms into Voronoi scheme. Moreover, a size adjustment procedure has been advocated for hetero-nuclear molecules. The above prescription is most suited for molecules; an alternative approach towards periodic system is nicely presented in ref. 80.

The single-center integral, I_j in eqn (22) is expressed in a spherical coordinate system as:

$$I_j = \int_o^\infty \int_0^\pi \int_0^{2\pi} F_j(r_j, \theta_j, \phi_j)\, r_j^2 \sin\theta_j\, \mathrm{d}r_j \mathrm{d}\theta_j \mathrm{d}\phi_j. \qquad (28)$$

Now we can distribute the 3D integration into 2D integration of variables $\Omega(\theta, \phi)$ and 1D integration of radial variable r in the following fashion, by applying some appropriate n-point quadrature formula having weight factor ω for both angular and radial parts,

$$F_j(r_j) = \int_0^\pi \int_0^{2\pi} F_j(r_j, \theta_j, \phi_j)\, \sin\theta_j \mathrm{d}\theta_j \mathrm{d}\phi_j \approx \sum_{i=1}^{n_j^\Omega} \omega_i^\Omega F_j(r_j, \Omega_i), \qquad (29)$$

and

$$I_j = \int_0^\infty F_j(r_j) r_j^2 \mathrm{d}r_j \approx \sum_{i=1}^{n_j^r} \omega_i^r F_j(r_i). \qquad (30)$$

Angular integration is eventually carried out on numerous sizes of a Lebedev grid[81,82] (also called octahedral grid), which is distributed on the surface of a unit sphere in such a way that the resulting point distribution remains invariant with respect to inversion. This grid is a consequence of various quadrature formulas[83] of remarkably high order, where all the spherical harmonics ($Y_{l,m}$) and its squares are integrated accurately. For optimal performance, it is best that the total number of grid points satisfies a certain condition as:

$$\text{No. of points} \approx \frac{(L+1)^2}{3}, \qquad (31)$$

where L is the maximum degree of the spherical harmonics ($0 \le l \le L$). Open-ended quadrature schemes are generally advisable; this is achieved via product type of grids quite easily in polar coordinates (Ω).[84] Efficiency can be further improved using a Lobatto grid, which uses a different criterion for the total number of points $\left(\approx \dfrac{L(L+1)^2}{2} \right)$, for optimal realization.[77] However, the majority of modern KS-DFT programs (within basis-set approach) like Gaussian, NWChem, Q-Chem, GAMESS *etc.*, exploit the Lebedev grid.

For radial integration, a fairly decent number of implementations are conceivable. These are typically carried out with the aid of some suitable

quadrature rules. Typically, a transformation is applied to map radial grid onto the quadrature grid varying the number of sampling points on the grid. Generally, a grid-pruning procedure is also administered in conjunction, to curtail the number of sampling points.[85,86] Some commonly used quadrature grids are: Gauss–Chebyshev integration of second kind, Euler–Maclaurin summation (extended trapezoidal) formula, log-squared quadrature, Gauss–Legendre, Gauss–Laguerre quadrature *etc.*[42,77,84,87–89] The Gauss–Chebyshev integration formula is given as,

$$\int_{-1}^{1} (1-x^2)^{\frac{1}{2}} f(x)\mathrm{d}x \approx \frac{\pi}{n+1} \sum_{i=1}^{n} \sin^2\left(\frac{i\pi}{n+1}\right) f(x_i); \quad x_i = \cos\left(\frac{i\pi}{n+1}\right),$$
(32)

whereas the Euler–Maclaurin scheme is executed through,

$$\int_{0}^{1} f(x)\mathrm{d}x \approx \frac{1}{n+1} \sum_{i=1}^{n} f(x_i); \quad x_i = \frac{i}{n+1}.$$
(33)

Subsequently, a double exponential formula has been proposed, and later refined by several other groups[90–92] to generate a radial grid. For large N, it becomes beneficial in the way of providing fast convergence rate. This is achieved by adopting an equally meshed trapezoidal rule for numerical integration over an arbitrary interval that transforms into an infinite integral by variable transformation. It has several variants depending on the mapping transformation and was invoked for finite, semi-finite, and infinite integrals. Another promising approach involves construction of an adaptive grid, pioneered by ref. 93–95 which is automatically generated for a given accuracy corresponding to a chosen basis function. Some of the prominent DFT programs like NRLMOL, deMon2k routinely use such grids.

It is well-known that the current enviable status of DFT that it enjoys largely depends on basis-set calculations. While such studies are heavily dominated by ACG, real-space grid has been invoked for fully numerical, basis-set free DFT methods. Apart from ACG, some scattered works exist for other grids in literature, *e.g.*, an adaptive Cartesian grid with a hierarchical cubature method,[96] a transformed sparse-tensor product grid,[97] a Fourier Transform Coulomb[98,99] method interpolating density from ACG to a more regular grid that helps to enhance its efficiency. In the latter, a difficulty arises due to computation of XC matrix elements, which is circumvented successfully with help of a multi-resolution method.[100] Here, in this work, we report a simple fruitful DFT implementation[45–49] within the LCAO-MO framework which solely uses CCG. So far, our results have been reported in a *uniform* grid. All relevant quantities such as basis functions, electron densities, MOs as well as various two-electron potentials (Hartree and XC) are directly set up on a real 3D Cartesian grid simulating a cubic box as,

$$r_i = r_0 + (i-1)h_r, \quad i = 1, 2, 3,\ldots,N_r; \quad \text{for } r \in \{x, y, z\},$$
(34)

where h_r, N_r signify grid spacing and total number of grid points respectively $\left(r_0 = -\dfrac{N_r h_r}{2}\right)$.

A major concern in the grid-based approach constitutes an accurate estimation of the classical electrostatic repulsion potential. For finite systems, the simplest and crudest way to calculate $v_h(\mathbf{r})$ is through direct numerical integration on the grid. For smaller systems, this is a feasible option; in all other cases, it is generally tedious and cumbersome. However, the most rewarding and widespread approach is through a solution of the corresponding Poisson equation,

$$\nabla^2 v_h(\mathbf{r}) = -4\pi\rho(\mathbf{r}). \tag{35}$$

The usual way to solve this is by conjugate gradients[101] or through multi-grid solvers.[102] As an alternative, the current work exploits a conventional Fourier convolution method originally suggested by ref. 50, 51 and 103 and adapted in the context of molecular modelling.[104,105] The basic principle can be formulated as:

$$v_h(\mathbf{r}) = \text{FFT}^{-1}\{v_h^c(\mathbf{k})\rho(\mathbf{k})\} \quad \text{and} \quad \rho(\mathbf{k}) = \text{FFT}\{\rho(\mathbf{r})\}, \tag{36}$$

where $v_h^c(\mathbf{k})$ and $\rho(\mathbf{k})$ stand for Fourier integrals of the Coulomb interaction kernel and density respectively, in the cubic box. The quantity $\rho(\mathbf{k})$ can be easily computed by using a discrete Fourier transformation of its real space value. Thus our primary concern here lies in calculation of the Coulomb interaction kernel which has a singularity in real space. For this, we utilize an Ewald summation-type approach,[105] expanding the Hartree kernel into long-range and short-range components,

$$v_h^c(\mathbf{r}) = \frac{\text{erf}(\alpha r)}{r} + \frac{\text{erfc}(\alpha r)}{r} \equiv v_{h_{\text{long}}}^c(\mathbf{r}) + v_{h_{\text{short}}}^c(\mathbf{r}), \tag{37}$$

where erf(x) and erfc(x) signify the error function and its complementary function respectively. The Fourier transform of the short-range part can be treated analytically, whereas the long-range portion needs to be computed directly from FFT of real-space values. A convergence parameter α is used to adjust the range of $v_{h_{\text{short}}}^c(\mathbf{r})$, such that the error is minimized. Following the conjecture of ref. 50, here we employ $\alpha \times L = 7$ (L denotes length of cubic box), which produces quite accurate results. For a non-uniform grid, L is chosen as the smallest side of the simulating box. Some other routes which calculate the long-range part efficiently are fast multipole,[106] multi-level summation,[107] fast Fourier–Poisson[108] method, *etc.*

2.5 Basis-set free methods

This sub-section gives an overview of fully numerical, real-space methods, used for solving the pertinent KS equation without needing any explicit basis set. Typically, the wave function is directly sampled on a discrete grid through a variety of representations; each having its own pros and cons. Most of these attempts also invoke pseudopotentials to

freeze the inner core. One central problem is the formation of an appropriate finer mesh, which normally calls for an enormous number of grid points, making calculations of the electronic structure (especially for large systems) very expensive in terms of both time and computer memory. Apparently, there are ways to minimize discretization error systematically by controlling size, shape and spacing of the 3D mesh. It is worth mentioning two notable works, *viz.*, a higher-order real-space pseudopotential method in uniform CCG,[109,110] as well as real-time propagation of KS orbitals.[111,112] Often, they are easily amenable to linear scaling schemes. Recently, extensions have been made to Graphical Processing Unit interface[113] and domain-decomposition techniques[114] for parallel computation. Some excellent reviews are available on the subject (see, for example, ref. 115–118).

Broadly speaking, real-space techniques can be categorized into three major classes, namely, finite difference (FD), finite element (FE) and wavelets.[119,120] In all cases, the target discretized differential equation produces structured and highly banded matrices, which can be solved readily using efficient multi-scale techniques. The potential operator is diagonal in coordinate space and the Laplacian operator is nearly non-local; that makes them particularly suited for parallel computing. In a uniform orthogonal 3D CCG of spacing h, *e.g.*, the Laplacian operator, within an FD approximation, can be expressed approximately in 1D (x direction) as a part of 3D:

$$\frac{d^2 \phi(x_i)}{dx^2} \approx \frac{1}{h^2} [\phi(x_{i-1}) - 2\phi(x_i) + \phi(x_{i+1})] - \frac{1}{12} \frac{d^4}{dx^4} \phi(x_i) h^2 + \mathcal{O}(h^4). \quad (38)$$

The last two terms account for truncation errors; the first contribution being second order in h with a pre-factor involving the fourth derivative of $\phi(x)$. Here grid spacing, h, is one of the key parameters that controls convergence. This approximation is valid in the limit of $h \to 0$, but for practical purposes, it is unfeasible to go below a certain limit, which leads to the evolution of higher-order FD schemes.[109] Most commonly, regular CCG is engaged and FD methods result from a Taylor series expansion of our desired function around the grid points. However, it is non-variational, because the error can be of either sign depending on the derivatives and value of $\phi(x)$. It is often quite difficult to achieve convergence due to this lack of variationality. While its chief advantage lies in a simplicity of representation and easy implementation, sometimes it is rather troublesome to construct flexible meshes which can reproduce physical geometry satisfactorily. There are reports of a Mehrstellen discretization technique as an alternative route to high-order FD methods.[121,122] Several workers,[123–126] have suggested use of FD techniques with in amalgamation of projector augmented wave methods.[127] Applications of the FD method is quite impressive; mention may be made about some of the notable ones for clusters and other finite systems.[109,110,128,129]

On the other hand, FE-based methods are well known from the engineering field for a long time, and have been implemented on

applications of electronic structure theory.[130] Customarily, one uses some suitable non-orthogonal basis functions, such as piecewise polynomials which are non-zero over a local region of space, and also divide the unit cell into elements. For solution of Poisson equation, one can expand the potential, $v(x)$, in a basis:

$$v(x) = \sum_i u_i \chi_i(x), \tag{39}$$

where u_i's are expansion coefficients corresponding to actual function values on mesh points and χ_i are basis functions. A similar kind of treatment can be performed on the charge density, and the resulting matrix equation will be similar to a one-dimensional Poisson equation. At least for uniform meshes, there is a close correspondence between FD and FE real-space representation. Unlike the FD method however, it has a variational foundation, and provides greater flexibility regarding the unit cell construction. A multitude of efficient basis sets have been suggested for 3D electronic structure calculations for diverse problems,[131] like cubic-polynomial basis,[132] tetrahedral discretization with orders $p = 1–5$,[133] cubic functions,[134] Lobatto–Gauss basis set with orders $p = 5–7$,[135] B-spline basis[136] *etc.*

Standard iterative processes suffer from so-called *critical slow down*;[115] a common complication that occurs when more grid points are used to obtain an enhanced accuracy on a fixed domain and thus makes it less efficient on finer meshes. In order to facilitate better convergence of self-consistent iteration procedure, a few advancements have been made. Most prominent of them is the family of multi-grid iteration techniques,[115,121,137] which ideally utilize information from multiple lengths scales to alleviate this problem. Another promising route involves engaging an adaptive grid in conjunction with multi-grid.[128,138] This has found applications within FD-type representations, but maximum effort was devoted to develop efficient solvers for FE representations.[139,140]

Another useful source of real-space techniques is based on the wavelet basis method having a more complicated matrix structure than either the FD or FE representation.[141] Such basis sets are semi-cardinal, local, and not only that, they are conceived with multi-resolution properties. One crucial advantage is that calculations can be performed with different kinds of boundary conditions. Daubechies wavelet[120] is one of the most favored and frequently used, which is local in both real and Fourier space. These are easily adaptable to parallel programming and are employed in connection with linear scaling methods[119] for electronic structure calculations. Several DFT codes like ABINIT, ONETEP, CONQUEST, CP2K, and SIESTA *etc.*, successfully implement wavelet schemes.

2.6 Basis set and LCAO-MO ansatz

As emphasized earlier, the basis-set approach remains by far the most convenient and pragmatic line of approach towards the solution of KS DFT. It dominates in every field of science whether being chemical, physical or biological applications employed to atoms, molecules or solid

etc. The inspiration mainly comes from the success of basis-set related methodologies in traditional wave function theory, such as HF and post-HF. This is a coupled integro-differential equation where kinetic and potential energies are defined by a differential and integral operator. It must be solved numerically iteratively leading to a self-consistent set of orbitals $\{\psi_i(\mathbf{r})\}$. Essentially, the unknown KS MOs are expanded in terms of K suitably chosen, known basis functions $\{\chi_\mu(\mathbf{r}); \ \mu = 1, \ 2, \ 3,\ldots,K\}$, conventionally called atomic orbitals, in a manner analogous to that in the Roothaan-HF method, such as,

$$\psi_i(\mathbf{r}) = \sum_{\mu=i}^{K} C_{\mu i}\chi_\mu(\mathbf{r}), \quad i = 1, 2, 3, \ldots K. \tag{40}$$

The electron density then takes the following expression in this basis,

$$\rho(\mathbf{r}) = \sum_{i=1}^{N}\sum_{\mu=1}^{K}\sum_{\nu=1}^{K} C_{\mu i}C_{\nu i}\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r}). \tag{41}$$

In principle, one requires a *complete* basis set ($K=\infty$) in order to get an exact expansion of MOs, but in reality it is not feasible. Therefore, suitable truncation is required for realistic computational purposes; it suffices to work with a finite basis set usually.

It is of paramount importance to choose appropriate functions which reproduce KS MOs as precisely as possible. The numerical accuracy of KS DFT is quite sensitive to the choice and design of the basis set for a particular problem, as an incomplete basis set inducts certain constraints on the relaxation of the density through KS orbitals. Considerable developments have taken place towards their construction as well as effects on assorted physico-chemical properties. There is no restriction, as such, for the effective implementation of a specific basis set, but it is required to fulfill certain conditions which may serve as some guidelines: (a) it should have proper behavior according to the physics of the problem (b) it should be friendly in terms of ease and efficiency of computation of requisite integrals (c) number of functions in basis-set expansion should be scaled down as much as possible without sacrificing any accuracy. A sizeable number of elegant, flexible, versatile basis sets has been proposed over the past several decades, from different perspectives such as: exponential,[142] Gaussian,[143] polynomial, cube functions, wavelets,[144] numerical functions,[78] plane waves[145] and many others. However, the problem still now remains open, as there is no universal basis set applicable for all methods or chemical systems. The interested reader may consult many nice reviews and books available on the topic.[2,5,10,13]

For a periodic system (*e.g.*, solid), the preferred choice is a plane wave basis whereas for non-periodic systems (*e.g.*, molecules, clusters) atom-centred localized basis functions, such as Slater Type Orbitals (STO) or Gaussian Type Orbitals (GTO) stand out to be two most attractive options. STOs assume the following functional form:

$$\chi_{\zeta,n,l,m}(r, \ \theta, \ \varphi) = N_c Y_{l,m}(\theta, \ \varphi)r^{n-1}e^{-\zeta r}, \tag{42}$$

where N_c is the normalization constant and $Y_{l,m}$ defines spherical harmonics. It is a well acknowledged fact that multi-center (3c or 4c) two-electron integrals can not be evaluated analytically with STOs, making them somehow less favorable for molecular system. Thus their application principally remains confined to atomic or diatomic systems or in semi-empirical approximations, where two- or three-center integrals are generally neglected.

On the other hand, GTOs are expressed in Cartesian coordinates as:

$$\chi_{\zeta,n_x,n_y,n_z}(x,y,z) = N_c x^{n_x} y^{n_y} z^{n_z} e^{-\zeta r^2}, \tag{43}$$

where n_x, n_y and n_z define the angular momentum such that $\lambda = n_x + n_y + n_z = 0, 1, 2,\ldots$, (s, p, d,... functions), while the orbital exponent ζ governs its compactness (large ζ) or diffuseness (small ζ). In spite of the fact that GTOs face problems to describe the proper behavior near and far from the nucleus than STOs, they are almost universally accepted for quantum chemistry calculations. To a great measure, this is due to their ability to provide closed, analytic formulas for multi-center integrals through Gaussian product theorem. In principle, both STOs and GTOs qualify to form a complete basis, but to reach a certain accuracy, ordinarily more GTOs are needed (roughly at least three times) than STOs. In practical purpose, instead of using individual GTOs as basis function, it is customary to use fixed linear combination of GTOs, called contracted GTOs, defined as:

$$\chi_\mu^{\text{contracted}}(\mathbf{r} - \mathbf{R_A}) = \sum_{p=1}^{L} d_{p\mu} \chi_p(\zeta_{p\mu}, \mathbf{r} - \mathbf{R_A}). \tag{44}$$

Here $d_{p\mu}$ are called contraction coefficients, L signifies length of contraction, whereas "p" stands for individual primitive Gaussians from which contracted Gaussian functions are formed. The normal procedure to find out primitives is to optimize exponents and contraction coefficients so as to obtain variational lowest-energy states through an atomic calculation (within HF, CI or so). However there are instances in literature where they are explicitly optimized through KS procedure using LDA XC functionals. Historically, basis sets were designed for applications in wave-function based methodology, and fortunately it turns out that, for most common molecular property calculations (like equilibrium geometry, energy), results are pleasantly insusceptible towards the manner through which these exponents and coefficients have been determined. That offers a welcome possibility to take advantage of such basis sets, initially constructed for traditional methods, to be directly usable in DFT calculations with confidence.

Now inserting eqn (40) in eqn (20), multiplying the left side of the resulting equation with $\chi_\mu^*(\mathbf{r})$, then integrating over the whole space, followed by some algebraic manipulation generates the following KS matrix equation, in parallel to the HF case,

$$\mathbf{F}^{\text{KS}}\mathbf{C} = \mathbf{S}\mathbf{C}\epsilon, \tag{45}$$

where **F** and **S** imply the $K \times K$ real, symmetric total KS and overlap matrices respectively. The eigenvector matrix **C** contains the basis-set expansion coefficients $\mathbf{C}_{\mu i}$ and the diagonal matrix ϵ holds the orbital energies ϵ_i. It could be readily solved by standard numerical techniques of linear algebra. The individual elements of KS matrix can be written as:

$$F^{\mathrm{KS}}_{\mu\nu} = \int \chi_\mu(\mathbf{r})[h^{\mathrm{core}} + v_{\mathrm{hxc}}(\mathbf{r})]\chi_\nu(\mathbf{r})d\mathbf{r} = H^{\mathrm{core}}_{\mu\nu} + \langle \chi_\mu(\mathbf{r}) | v_{\mathrm{hxc}} | \chi_\nu(\mathbf{r}) \rangle$$
$$= H^{\mathrm{core}}_{\mu\nu} + J_{\mu\nu} + V^{\mathrm{XC}}_{\mu\nu}, \qquad (46)$$

where $H^{\mathrm{core}}_{\mu\nu}$ represents the core bare-nucleus Hamiltonian matrix element consisting of the kinetic energy and nuclear-electron attraction, thus accounting for one-electron energies. The one-electron matrix elements can be evaluated analytically with the help of well-established recursion relations[146] for Gaussian bases. Second term $v_{\mathrm{hxc}}(\mathbf{r})$ contains all two-electron interactions involving classical Coulomb repulsion and the non-classical XC potential. $J_{\mu\nu}$ denotes the matrix element of the classical Hartree potential defined in eqn (21), while the remaining term, $V^{\mathrm{XC}}_{\mu\nu}$ supplies the XC contribution into two-body matrix element, whose construction remains one of the most vital steps in the whole KS DFT process. In absence of any analytical method, two-body matrix elements can be either calculated numerically or fitted by an auxiliary set of Gaussian functions, as suggested by.[147–149] In our work, we employed direct numerical integration on the CCG to obtain these matrix elements:

$$\langle \chi_\mu(\mathbf{r}) | v_{\mathrm{hxc}}(\mathbf{r}) | \chi_\nu(\mathbf{r}) \rangle = h_x h_y h_z \sum_{\mathrm{grid}} \chi_\mu(\mathbf{r}) v_{\mathrm{hxc}}(\mathbf{r}) \chi_\nu(\mathbf{r}). \qquad (47)$$

For gradient-corrected functionals, the non-local XC contribution of the KS matrix is implemented by a finite-orbital basis expansion, without requiring to evaluate the density Hessians. Thus, in such cases, the XC contribution is written in a convenient working form, as suggested in ref. 150,

$$F^{\mathrm{XC}\alpha}_{\mu\nu} = \int \left[\frac{\partial f}{\partial \rho_\alpha} \chi_\mu \chi_\nu + \left(2 \frac{\partial f}{\partial \gamma_{\alpha\alpha}} \nabla \rho_\alpha + \frac{\partial f}{\partial \gamma_{\alpha\beta}} \nabla \rho_\beta \right) \cdot \nabla(\chi_\mu \chi_\nu) \right] d\mathbf{r}, \qquad (48)$$

where $\gamma_{\alpha\alpha} = |\nabla \rho_\alpha|^2$, $\gamma_{\alpha\beta} = \nabla \rho_\alpha \cdot \nabla \rho_\beta$, $\gamma_{\beta\beta} = |\nabla \rho_\beta|^2$, and f is a function only of local quantities ρ_α, ρ_β and their gradients.

2.7 Pseudopotential

In today's repertoire of computational chemistry/physics electronic structure theory, a pseudopotential approximation is a very useful and powerful concept which is extensively used for atoms (especially with heavy nuclei), molecules (containing one/more transition metal/ lanthanide/actinide), solid state, *etc.*[151–153] The idea was originally introduced a long time ago by Hellmann, Fermi and Gombas independently; thereafter applied and popularized by several other groups.[154,155]

It simplifies by dividing the space of many electrons into core and valence categories. To a very good approximation, because of their strong binding with the nucleus, inner-shell electrons form a passive "inert" core (retaining an atomic-like configuration) which supposedly plays a less significant role in the understanding of chemical binding. Consequently, a surprising majority of chemical properties can be followed very satisfactorily by considering only valence electrons. In effective terms, this implies to replace the strong Coulomb potential between nucleus and tightly bound core electrons (as a part of the exact Hamiltonian) by a non-local, smoother potential acting on the valence electrons. That means, the spectrum of the resulting pseudo-Hamiltonian is as close to the exact one and this is also true for the pseudo eigenfunction. It diminishes computational cost in larger molecular systems dramatically, simply by cutting down the basis set, and thus avoiding any integral arising from core electrons. A lot of developments have been recorded regarding the suitability and effectiveness of this approach, compared to the more complete and rigorous, all-electron calculations.[156–158] Moreover, relativistic effects can be brought into a many-electron picture efficiently through pseudo-potential schemes.[155,159]

Thus a molecular pseudo-Hamiltonian H_v can be factored as:

$$H_v = -\frac{1}{2}\sum_i^{n_v} \nabla_i^2 + \sum_{i<j}^{n_v} \frac{1}{\mathbf{r}_{ij}} + \sum_i^{n_v}\sum_a^{N_c}\left[v_a^{\mathrm{PP}}(\mathbf{r}_{ai}) - \frac{Q_a}{\mathbf{r}_{ai}}\right] + \sum_{a<b}^{N_c} \frac{Q_a Q_b}{\mathbf{r}_{ab}}, \quad (49)$$

where n_v denotes number of valence electrons, N_c the number of core (nuclei) ones and Q_a signifies residual charge of core a. Indices (i, j) and (a, b) run over all valence electrons and all cores respectively; while v^{PP}, the pseudopotential operator takes account of the fact that valence space is repulsive in short range and attractive in long range. Therefore, the basic challenge is to find out a decent and computer-proficient approximation to v^{PP} comparable with all-electron calculations. It is to be noted that the distinction between core and valence shells is essential for the success of the pseudopotential approximation and hence, it is beneficial to work within an orbital-based theory. Under this situation, our desired KS equation could be rewritten as:

$$\left[-\frac{1}{2}\nabla^2 + v_{\mathrm{ion}}^{\mathrm{PP}}(\mathbf{r}) + v_{\mathrm{h}}[\rho(\mathbf{r})] + v_{\mathrm{xc}}[\rho(\mathbf{r})]\right]\psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \quad (50)$$

where $v_{\mathrm{ion}}^{\mathrm{PP}}$ designates ionic pseudopotential for the system,

$$v_{\mathrm{ion}}^{\mathrm{PP}} = \sum_{R_a} v_{\mathrm{ion},a}^{\mathrm{PP}}(\mathbf{r} - \mathbf{R}_a). \quad (51)$$

In above equation, $v_{\mathrm{ion},a}^{\mathrm{PP}}$ represents the ion-core pseudopotential associated with atom a, situated at $\mathbf{R}_a$; whereas $v_{\mathrm{h}}[\rho(\mathbf{r})]$ denotes the classical electrostatic potential among the valence electrons and $v_{\mathrm{xc}}[\rho(\mathbf{r})]$, as usual defines the non-classical part of the many-electron Hamiltonian.

At present, two possible approaches are predominant for pseudopotential approximations in molecular applications, namely (i) model core potential[160] and (ii) semi-local approximation. In the former, v^{pp} is formulated as,

$$v^{\text{pp}}(\mathbf{r}) = \sum_k A_k \mathbf{r}^{n_k} e^{-\alpha_k \mathbf{r}^2} + \sum_c B_c |\phi(\mathbf{r})\rangle\langle\phi(\mathbf{r})| \tag{52}$$

where A_k, n_k and α_k are all adjustable parameters and B_c is a chosen parameter such that $B_c = -2\epsilon_c$, the core energies (k runs over Gaussian expansion). It is possible to get HF orbital energies and corresponding radial functions[161] by adjusting above equation. An advantage of this representation is that the inner-shell structure of the pseudo-valence orbitals is conserved and closely resembles the all-electron valence orbitals; as a result, relativistic and spin-orbit effects can be efficiently included into it.[162] Incorporation of non-local core-valence exchange of the all-electron Fock operator into this approximation leads to *ab initio* model potential approximation,[163] which has been profitably implemented into molecular program packages such as Molcas, GAMESS-US *etc.*

On the other hand, within a semi-local approximation,[164–169] v^{pp} is expressed as,

$$v^{\text{pp}}(\mathbf{r}) = v_{\text{local}}(\mathbf{r}) + \sum_{l=0}^{l_{\text{max}}} v_l(\mathbf{r}) P_l$$
$$v^{\text{pp}}(\mathbf{r}) = \sum_k A_k \mathbf{r}^{n_k} e^{-\alpha_k \mathbf{r}^2} + \sum_{l=0}^{l_{\text{max}}} \sum_k B_{lk} \mathbf{r}^{n_{lk}} e^{-\beta_{lk} \mathbf{r}^2} \sum_{m=-l}^{l} |lm\rangle\langle lm|, \tag{53}$$

where A_k, B_{lk}, n_k, n_{lk}, α_{lk} and β_{lk} are all adjustable parameters. The last term in above equation contains a projection operator P_l which projects onto the Hilbert sub-space with angular momentum l. Unlike eqn (52), it does not contain any core function and hence, uses orbitals that are smooth and node-less in short-range of the radial function. This approximation has found profound applications, especially in relativistic and spin–orbit effects[170–172] and, has been implemented in several molecular program packages like Gaussian90, Molpro, Turbomole and also in solid-state program CRYSTAL. Depending on the manner in which adjustable parameters in eqn (53) are ascertained, there are two leading categories in this case, namely, (i) energy-consistent[173] and (ii) shape-consistent.[174] In the former scenario, these parameters are settled down through a fitting procedure (often least square)[175] to a large number of pre-calculated transitions in valence space. In latter plan, this is done in such way that the valence orbitals of different symmetries resemble the shape of the corresponding all-electron orbitals very closely after a certain cut-off radius and also match the orbital energies with high accuracy. This scheme was formulated and implemented by a number of research groups.[52,165,166,169,174,176] Additionally it is usually faster and more efficient than the former, though it requires a large number of Gaussian fitting functions for inversion of the Fock equation involved. Moreover, the first procedure requires a vast number of valence spectra to fulfill the

shape-consistent requirement. On the other hand, there is a family of non-conserving pseudopotentials[177] (slightly different from shape-consistent), which has been found to be most suitable for plane-wave codes within DFT. In our current development, we adopt the latter form as proposed by ref. 52 and 174.

3 Results and discussion

Before we proceed for a discussion of results, mention may be made about some of the technical aspects. Our desired KS equation is solved following the usual self-consistent procedure by imposing three convergence criteria namely, (i) electronic energy differences between two successive iterations lies below 10^{-6} a.u. (ii) maximum absolute deviation in the potential is less than 10^{-5} a.u. and (iii) standard deviation in the density matrix remains below 10^{-5} a.u. The generalized matrix-eigen value problem is solved using standard LAPACK routine[178] efficiently and accurately. Computation of FFT is done by means of the standard FFTW3 package.[179]

We begin by scrutinizing the stability and convergence for a representative homo-nuclear diatomic molecule Cl_2, at an internuclear distance of 4.2 a.u. For this, the non-relativistic ground-state total energy as well as various energy components are reported in Table 1 for eight chosen grid Sets A–H, varying in N_r and h_r ($r \in \{x, y, z\}$). The LDA XC potential, corresponding to homogeneous electron-gas correlation of Vosko-Wilk-Nusair,[37] is used in this table, which is quoted from ref. 45. All results in this and the following tables compare these with corresponding reference values obtained from the standard GAMESS suite of quantum chemistry program[180] incorporating the same XC functional, basis set and effective core potential. All pseudopotential calculations in this work, utilize the Hay-Wadt-type[53] valence basis set in which the orbital is split into inner and outer components (derived by two and one primitive Gaussian functions respectively). Reference values are quoted for both ''grid'' and ''grid-free'' calculations; former has a default ''army'' grade grid with Euler–MacLaurin quadratures for radial integration and Gauss–Legendre quadrature for angular integrations. The latter[181] follows an auxiliary basis-set expansion procedure for molecular integrals instead of a quadrature grid; it has the ability to bypass any error associated with a finite grid, though it suffers from an apparently inherent weakness of incompleteness. Following expectation values of energy components are reported for comparison, *viz.*, kinetic energy $\langle T \rangle$, total nucleus-electron potential energy $\langle V_t^{ne} \rangle$, total two-electron potential energy $\langle V_t^{ee} \rangle$, total potential energy $\langle V \rangle = (\langle V_t^{ne} \rangle + \langle V_t^{ee} \rangle + \langle E_{nu} \rangle)$ where $\langle E_{nu} \rangle$ is nuclear repulsion energy, plus total integrated electron density N. Individual components of $\langle E_{hxc} \rangle$ are not produced as they can not be directly compared (unavailable from output). Clearly, Set A energies deviate maximum from reference values, presumably because our box is too small to account for all important interactions; this is also reflected in a poor N value. As grid spacing is increased from 0.3 of Set A to 0.4 in Set B keeping N_r fixed at 32, with a corresponding enlargement of box

Table 1 Energy components and total number of electrons, N, in various grids, for Cl_2, at $R = 4.2$ a.u., along with reference results. All quantities are in a.u. It is adopted from ref. 45.

Set	A	B	C	D	E	F	G	H	Ref. 180
h_r	0.3	0.4	0.2	0.3	0.4	0.1	0.2	0.1	
N_r	32	32	64	64	64	128	128	256	
$\langle T \rangle$	11.00750	11.17919	11.18733	11.07195	11.06448	11.18701	11.07244	11.07244	11.07320
$\langle V_t^{ne} \rangle$	−83.43381	−83.68501	−83.70054	−83.45722	−83.44290	−83.69988	−83.45810	−83.45810	−83.45964
$\langle V_t^{ee} \rangle$	32.34338	31.22265	31.22885	31.00981	31.00306	31.22832	31.01000	31.01000	31.01078
$\langle E_{nu} \rangle$	11.66667	11.66667	11.66667	11.66667	11.66667	11.66667	11.66667	11.66667	11.66667
$\langle V \rangle$	−39.42376	−40.79570	−40.80503	−40.78074	−40.77317	−40.80489	−40.78144	−40.78144	−40.78219
$\langle E_{el} \rangle$	−40.08293	−41.28318	−41.28437	−41.37545	−41.37535	−41.28455	−41.37566	−41.37566	−41.37566
$\langle E \rangle$	−28.41626	−29.61651	−29.61770	−29.70878	−29.70868	−29.61789	−29.70900	−29.70900	−29.70899[a]
N	13.89834	13.99939	13.99865	14.00002	14.00003	13.99864	14.00000	13.99999	13.99998

[a] Refers to the *grid*-DFT calculation; corresponding *grid-free* DFT value is −29.71530 a.u.

size, all quantities improve considerably. Sets C, F also furnish similar kind of accuracy as the box length is same; however both are inadequate to reach the correct value. It is interesting to note that Sets B, C, F all correspond to the same box size; but the quality of results is slightly better in case of Set F, as it offers a finer grid structure $(h_r = 0.1)$. A similar conclusion can be drawn for Sets E, G too. Results from Sets D, E are quite comparable to reference values. In order to prove stability, some further calculations are performed by elongating the box in Sets F–H in columns 7–9. This discussion evidently establishes that either of four Sets D, E, G, H offer the desired accuracy; while D, E are capable of providing sufficient accuracy for all practical purposes. A similar analysis was done for HCl[45] (at $R = 2.4$ a.u.) in various grids, giving analogous pattern. Set B seems to be quite reasonable for reproducing literature results up to third decimal place, although it was inadequate for Cl_2. All the sets produce results of similar accuracy; Set B amongst them does it slightly poorly in terms of component energies and N. As one passes from B–C–D, a gradual improvement in results is recorded, as expected. Sets C and F indicate agreement with each other as in Cl_2. For Sets D and E, N remains practically unchanged and energy increases by ≈ 0.0002 a.u. Additional calculations with $N_r = 128$, $h_r = 0.3$ suggests complete agreement with Set G. As in Cl_2, the best three sets for HCl are again D, E and G; with the first two being fairly accurate for most calculations.

Up to this point, we focused on results with LDA functionals only. As mentioned in the Introduction, such calculations are fraught with some severe limitations. Hence in order to make practical applications, more refined and accurate functionals are essential. In this context, let us recall that E_{xc} consists of two components, each arising as a difference: (a) between classical and quantum mechanical inter-electronic repulsion (b) between kinetic energy of the hypothetical, non-interacting and the real, interacting system concerned. Usually, the second portion is not explicitly incorporated in the current family of functionals. In this sub-section, we present a cross-section of CCG results within the familiar non-local Becke exchange[39] and LYP[40] correlation, who converted the Colle–Salvetti formula for the correlation energy from HF second order density matrix into a density functional form. This will extend the range and scope of applicability of the present scheme in the context of more realistic non-local functionals which will be crucial for future chemical applications. The latter is implemented using an alternate equivalent form containing only the first derivative through a procedure due to ref. 182. Table 2 presents a comparison of our BLYP energy components and N at some selective grid sets at same R as in Table 1. The same quantities as in the LDA case are reported. Several test calculations were performed in several grids to ensure convergence of results; which offered conclusions quite similar to those found in LDA. From this, two grid sets are given for each of them. Once again, our CCG results exhibit excellent agreement with literature values, as observed in LDA scenario. In Set B, results are slightly better than Set A for obvious reasons of finer grid and an extended length of box. In Cl_2, this effect is more pronounced than in case of HCl. Absolute deviations in Set B energies are 0.00002 and 0.00000 a.u.

Table 2 Variation of the energy components and N with respect to the grid parameters for Cl_2 and HCl with reference values. BLYP results in a.u.

| Set | Cl_2 ($R = 4.2$ a.u.) | | | HCl ($R = 2.4$ a.u.) | | |
	A	B	Ref. 180	A	B	Ref. 180
N_r	64	128		64	128	
h_r	0.3	0.2		0.3	0.2	
$\langle T \rangle$	11.21504	11.21577	11.21570	6.25431	6.25464	6.25458
$\langle V_t^{ne} \rangle$	−83.72582	−83.72695	−83.72685	−37.29933	−37.29987	−37.29979
$\langle V_t^{ee} \rangle$	31.07572	31.07594	31.07594	12.63884	12.63903	12.63901
$\langle E_{nu} \rangle$	11.66667	11.66667	11.66667	2.91667	2.91667	2.91667
$\langle V \rangle$	−40.98344	−40.98434	−40.98424	−21.74382	−21.74417	−21.74411
$\langle E_{el} \rangle$	−41.43506	−41.43524	−41.43522	−18.40618	−18.40620	−18.40620
$\langle E \rangle$	−29.76840	−29.76857	−29.76855[a]	−15.48951	−15.48953	−15.48953[b]
N	14.00006	14.00000	13.99998	8.00002	8.00000	8.00000

[a] Corresponding *grid-free* DFT value is −29.74755 a.u.[180]
[b] Corresponding *grid-free* DFT value is −15.48083 a.u.[180]

Table 3 Comparison of calculated negative eigenvalues of Cl_2 and HCl with reference values. BLYP results are given in a.u. See text for more details.

| MO | Cl_2 ($R = 4.2$ a.u.) | | | MO | HCl ($R = 2.4$ a.u.) | | |
Set	A	B	Ref. 180		A	B	Ref. 180
N_r	64	128			64	128	
h_r	0.3	0.2			0.3	0.2	
$2\sigma_g$	0.8143	0.8143	0.8143	2σ	0.7707	0.7707	0.7707
$2\sigma_u$	0.7094	0.7094	0.7094	3σ	0.4168	0.4167	0.4167
$3\sigma_g$	0.4170	0.4171	0.4171	$1\pi_x$	0.2786	0.2786	0.2786
$1\pi_{xu}$	0.3405	0.3405	0.3405	$1\pi_y$	0.2786	0.2786	0.2786
$1\pi_{yu}$	0.3405	0.3405	0.3405				
$1\pi_{xg}$	0.2778	0.2778	0.2778				
$1\pi_{yg}$	0.2778	0.2778	0.2778				

respectively for Cl_2, HCl. Grid-free and grid DFT results are substantially different for both molecules. More details could be found in ref. 46.

In order to justify the dependability and usefulness of CCG, Table 3 now samples calculated negative eigenvalues for Cl_2 and HCl using BLYP XC functional at same geometries of previous tables. These are provided in two selected grid sets. Comparison with reference values again indicates excellent agreement for both molecules for *all* orbital energies except the lone case of $3\sigma_g$ for Cl_2 (Set A) and 3σ for HCl (Set B); even then the absolute deviation remains well within 0.0001 a.u. More details could be found in ref. 45.

Now, Table 4 offers a comparative chart of $\langle T \rangle$, $\langle V \rangle$, total (E) energies and N for few chosen atoms and molecules using both LDA and XC functionals. LDA results employ Set E grid ($N_r = 64$, $h_r = 0.4$), while for BLYP calculations, except P and H_2S (using Set E), grid Set I ($N_r = 128$, $h_r = 0.3$) was used. Overall, the present calculated results agree with literature values more or less up to two decimal places and N completely matches with the reference values with few exceptions. Results for more atoms and molecules can be found in ref. 45 and 46.

Table 4 Comparison of various energies and *N* for several atoms and molecules for different XC functionals. For more details, see text.

System	$\langle T \rangle$				$-\langle V \rangle$			
	LDA	Ref. 180	BLYP	Ref. 180	LDA	Ref. 180	BLYP	Ref. 180
P	2.35430	2.35334	2.38891	2.38890	8.73501	8.73404	8.78249	8.78248
H$_2$S	4.90204	4.90197	4.98071	4.98066	16.10707	16.10698	16.21919	16.21913
MgCl$_2$	11.62114	11.62208	11.75947	11.75999	42.34513	42.34621	42.54049	42.54103
SiH$_2$Cl$_2$	13.95036	13.94989	14.14948	14.14945	48.78729	48.78685	49.04463	49.04461

System	$-\langle E \rangle$				N			
	LDA	Ref. 180	BLYP	Ref. 180	LDA	Ref. 180	BLYP	Ref. 180
P	6.38070	6.38071	6.39358	6.39358	5.00000	4.99999	4.99999	4.99999
H$_2$S	11.20503	11.20501	11.23848	11.23846	8.00000	7.99999	8.00000	7.99999
MgCl$_2$	30.72399	30.72413	30.78102	30.78104	16.00004	15.99957	16.00004	15.99999
SiH$_2$Cl$_2$	34.83693	34.83696	34.89515	34.89516	19.99999	20.00015	19.99999	20.00000
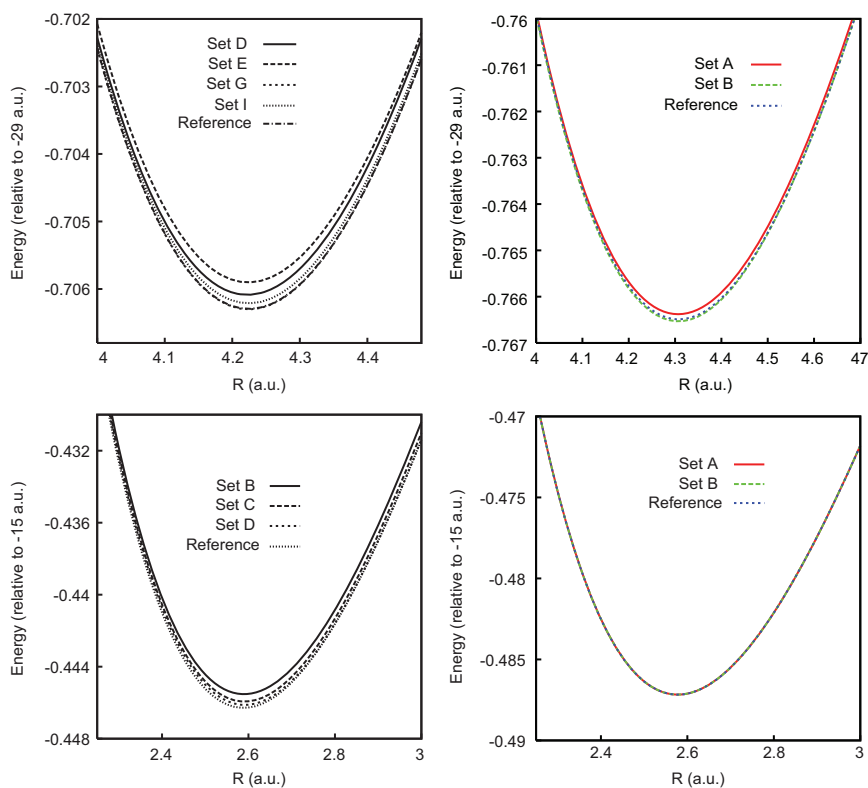
Fig. 1 Potential energy curves for Cl$_2$ (top), HCl (bottom) for different grid sets. Left, right panels correspond to LDA, LYP functionals. Reference grid-DFT results are also given for comparison. Grid sets in left, right panels refer to corresponding (N_r, h_r) values of Tables 1 and 2 respectively.

For further examination, Fig. 1 displays the changes in above LDA energies in Cl$_2$ (relative to -29 a.u.) for four selected Sets D, E, G, I and HCl (relative to -15 a.u.) for Sets B, C, D respectively. They cover a span of 3.5–5.0 and 1.6–3.1 a.u., for Cl$_2$ and HCl. Respective reference plots are

also given for easy comparison. It is quite gratifying that all four sets replicate the shape of the potential energy curve quite well for the entire range of R, in Cl_2. A thorough analysis of results (not given here and the reader may be referred to ref. 45) reveals that, until $R = 4$ a.u., Set D energies are quite close to the reference (higher by only 0.0001 a.u.), while implying a gradual tendency to deviate thereafter. However, the maximum deviation still remains quite small overall (only 0.0007 a.u.) occurring for $R = 5$ a.u. Sets G, I computed energies either completely match with reference or provide a maximum difference of 0.0001 a.u. Further in all sets, our CCG energies have always been above reference except in two occasions ($R = 4$ and 4.3 for Set G). Now we turn to the bottom panel in the left side for HCl. Once again, we observe excellent qualitative agreement for whole range of R, as evident from their nearly identical shapes. Careful examination admits that maximum absolute deviation for three sets B, C, D remains rather small, *viz.*, 0.0014, 0.0007, 0.0002 for $R = 3$, 2.3, 2.3 respectively. It is possible to fine-tune these results further by enforcing more rigorous convergence criteria. For more details, see ref. 45. Similar plots for BLYP functionals are given in right side for Cl_2, HCl in top, bottom panels, for two grids A, B. In both cases, again respective grid-DFT plots are provided for comparison, which practically coincide. A careful analysis[46] shows that maximum disagreements are 0.0002 a.u. (only for 2 instances) and 0.0001 a.u. for Sets A, B respectively in case of Cl_2. However, for HCl, the same for both sets are 0.0001 a.u. (only for 2 instances). It is a matter of fact that the two sets confirm to each other more for HCl than Cl_2.

Now results are given for four selected species in Table 5, for two other functionals which are used in applications of many-electron systems with considerable success, namely, PBE[41] and Filatov–Thiel (FT97).[183,184] These non-local functionals are implemented with the help of Density Functional Repository program.[185] For sake of completeness, respective LDA and BLYP values are also provided. Experimental geometries are taken from NIST computational chemistry database.[186] Set E grid was used for this table. Since correctness of our results has already been

Table 5 Comparison of kinetic ($\langle T \rangle$), potential ($\langle V \rangle$), total (E) energies and N for several atoms and molecules for different XC functionals. See text for details.

System	$\langle T \rangle$				$-\langle V \rangle$			
	LDA	BLYP	PBE	FT97	LDA	BLYP	PBE	FT97
Br	4.22038	4.27022	4.17153	4.22828	17.28157	17.36122	17.29168	17.37615
Al_2	1.35409	1.37642	1.32292	1.34896	5.21624	5.24844	5.20904	5.25881
SiH_4	3.65832	3.76989	3.53302	3.60707	9.82333	9.97905	9.75741	9.86882
S_2	7.58306	7.68138	7.51586	7.59269	27.62272	27.77409	27.64063	27.76097

System	$-\langle E \rangle$				N			
	LDA	BLYP	PBE	FT97	LDA	BLYP	PBE	FT97
Br	13.06119	13.09100	13.12015	13.14787	7.00000	7.00000	6.99999	6.99999
Al_2	3.86215	3.87203	3.88612	3.90985	5.99999	5.99999	5.99999	5.99999
SiH_4	6.16500	6.20916	6.22438	6.26175	7.99999	8.00000	7.99999	7.99999
S_2	20.03966	20.09271	20.12477	20.16828	12.00000	12.00000	12.00000	12.00000

Table 6 Comparison of $-\epsilon_{HOMO}$ (in a.u.) and atomization energies (in kcal mol^{-1}) for several molecules for different XC functionals with literature experimental data.

System	$-\epsilon_{HOMO}$ (a.u.)						Atomization energies (kcal mol^{-1})		
	LDA	BLYP	PBE	FT97	LBVWN	Expt.[187]	LDA	BLYP	Expt.[187]
SiH$_4$	0.3188	0.3156	0.2919	0.2702	0.4624	0.4042	339.43	312.02	302.6
PH$_2$	0.2170	0.2111	0.1970	0.1741	0.3504	0.3610	152.77	139.92	149.2
P$_4$	0.2712	0.2575	0.2525	0.2369	0.3964	0.3432	200.77	142.99	285.9
Al$_2$	0.1407	0.1400	0.1318	0.1161	0.2371	0.1984	22.92	21.42	37.0

established in previous tables, here we omit the reference values; all entries are very close to them. Total energies decrease as one passes from LDA to non-local functionals. Within the latter category, there is some agreement amongst the three, with FT97 offering the lowest. N maintains same accuracy for all functionals.

Finally Table 6, provides a comparison of calculated $-\epsilon_{HOMO}$ and atomization energies with experimental results[187] for 4 selected molecules at their experimental geometries,[186] with LDA and BLYP XC functionals. For both quantities, reference theoretical results[180] are practically identical to those found from CCG, and so omitted here to avoid crowding. It is observed that calculated ionization energies generally signal large deviations (typically off by 30–50%) from experimental results for both functionals. Estimated ionization energies from PBE and FT97 are also reported, which also deliver similar performances. This leads to a general trend, which some of the frequently used functionals suffer from, *i.e.*, even though energies are reproduced properly, ionization energies are described rather poorly. Now, it is a crucial requirement for real-time electron dynamics to have a proper description of orbital energies, especially the higher-lying ones. For this, we present the computed HOMO energies with modified Leeuwen–Baerends (LB) exchange potential[188,189] along with the other four functionals. It is known to give a more satisfactory representation of the long-range nature of the XC potential, as reflected from a better matching with experimental results. This is formulated as:

$$v_{xc\sigma}^{LB\alpha}(\alpha, \beta: \mathbf{r}) = \alpha v_{x\sigma}^{LDA}(\mathbf{r}) + v_{c\sigma}^{LDA}(\mathbf{r}) + \frac{\beta x_\sigma^2(\mathbf{r})\rho_\sigma^{1/3}(\mathbf{r})}{1 + 3\beta x_\sigma(\mathbf{r})\ln\{x_\sigma(\mathbf{r}) + [x_\sigma^2(\mathbf{r}) + 1]^{1/2}\}},$$

(54)

where $\alpha = 1.19$, $\beta = 0.01$ are two empirical parameters, σ signifies up, down spins and last term containing gradient correction is reminiscent of the exchange functional of ref. 39, while $x_\sigma(\mathbf{r}) = |\nabla\rho_\sigma(\mathbf{r})|[\rho_\sigma(\mathbf{r})]^{-4/3}$ is a dimensionless quantity. This ensures the required long-range nature, *i.e.*, $v_{xc\sigma}^{LB\alpha}(\mathbf{r}) \to -1/r$, $r \to \infty$. A combination of LB exchange with the VWN correlation functional produces the HOMO energies of column 6, which abundantly establishes its superiority over all the other four functionals considered here. These are substantially lower and more accurate than four functionals. In case of atomization energy, the present results reveal

a considerable deviation from experimental values for both LDA and BLYP functionals. Note that latter contains zero-point vibrational corrections and relativistic effects. Interestingly, LDA atomization energies are apparently better than BLYP counterparts in some cases, but that should not be misconstrued to identify the former as a better choice. There could be some cancellation of errors. Moreover there are other important factors, such as using more complete accurate basis set or more appropriate pseudopotential (the current work uses one designed for HF calculations; those developed for DFT would be more suitable), which could circumvent some of these discrepancies. However, such deviations are not uncommon and found, even in all-electron calculations with more sophisticated, extended basis sets. More discussion on these issues could be found in ref. 46. Nevertheless, this is aside the primary objective of this chapter and does not directly impact the main conclusions drawn herein.

So far, we were mainly concerned with valence-shell electrons using some sort of effective core potentials to incorporate the effects of frozen core electrons. Undoubtedly, while the pseudopotential approximation is indispensable for larger systems (especially those containing heavy atoms), ''full'' calculations provide a more detailed and accurate result at the expense of a heavy computational cost. Nowadays, it is all about a balance between cost and accuracy. If the cost accuracy ratio permits, unquestionably, the latter remains the preferred choice for most chemical and physical problems. Thus, as a further extension of our method, we present some sample results on ''full'' calculations in Table 7. The electronic energy including the relevant components and N, are produced for two representative atoms, namely, N and O in their open-shell ground-state configuration. We employed an STO-3G basis set, LDA XC potential and grid Set I ($N_r = 128$, $h_r = 0.3$). In all cases computed quantities, once again, establish excellent agreement with reference values. This validates the effectiveness and level of performance of our method in the context of all-electron calculations. It also gives a comparison of orbital energies for both α and β spins for the same two atoms, which again proves the usefulness. It is a matter of fact that we use very crude basis functions and LDA functional for the purpose of illustration; involvement of better basis sets and non-local XC functionals should enhance the applicability and accuracy of full calculations.

Before passing, a few comments are in order. The computational scaling goes as follows: (a) construction of the localized basis set scales as N_g, where N_g is total number of grid points (b) formation of the basis set in grid scales as $N_b \times N_g$, with N_b giving the number of basis functions (c) generation of the electron density in grid scales as $N_b^2 \times N_g$ (d) one- and two-body matrix elements of the KS matrix scale as N_b^2 and $N_b^2 \times N_g$ respectively. One obvious way to achieve better scaling is by reducing N_b and N_g. Now, up to this point, all our calculations were performed in a uniform Cartesian grid; further there was a restriction of 2^{2N} grid points due to the FFT algorithm employed. To lift these restrictions, we pursue some calculations in a non-uniform, unequal box of arbitrary size other than a cubic shape. For treatment of larger systems, it is highly desirable,

Table 7 Energy components and orbital energies for two atoms (N, O) using Cartesian grid. All-electron calculations with LDA XC functionals using STO-3G basis set are given along with reference data. PR implies Present Result.

Energy	N		O		Orbital	N		O	
	PR	Ref. 180	PR	Ref. 180		PR	Ref. 180	PR	Ref. 180
$\langle T \rangle$	53.66407	53.66407	73.44497	73.44497	ϵ_{1s}^{α}	−13.6312	−13.6311	−18.3331	−18.1330
$\langle V^{ne} \rangle$	−127.32649	−127.32649	−176.32432	−176.32432	ϵ_{2s}^{α}	−0.6152	−0.6153	−0.7538	−0.7537
$\langle V^{ee} \rangle$	20.25538	20.25536	29.42801	29.42799	$\epsilon_{2p_x}^{\alpha}$	−0.1671	−0.1672	−0.1941	−0.1942
$\langle V \rangle$	−107.07111	−107.07113	−146.89631	−146.89634	$\epsilon_{2p_y}^{\alpha}$	−0.1671	−0.1672	−0.1941	−0.1942
$\langle E \rangle$	−53.40704	−53.40701	−73.45134	−73.45137	$\epsilon_{2p_z}^{\alpha}$	−0.1671	−0.1672	−0.1085	−0.1085
N	6.99999	6.99999	7.99999	7.99999	ϵ_{1s}^{β}	−13.5837	−13.5836	−18.2972	−18.2971
					ϵ_{2s}^{β}	−0.4449	−0.4450	−0.6301	−0.6302
					$\epsilon_{2p_x}^{\beta}$			−0.0378	−0.0379

Table 8 Electronic energy^a of Cl_2 in non-uniform grid ($h_r = 0.3$). LDA results in a.u.

Set I				Set II			
N_x	N_y	N_z	$\langle E \rangle$	N_x	N_y	N_z	$\langle E \rangle$
30	30	30	−25.32542	30	30	70	−29.66504
—	—	40	−29.48297	32	32	—	−29.68722
—	—	50	−29.65381	34	34	—	−29.69838
—	—	60	−29.66461	36	36	—	−29.70387
—	—	70	−29.66504	38	38	—	−29.70651
—	—	80	−29.66505	40	40	—	−29.70777
—	—	90	−29.66505	42	42	—	−29.70837
	36	40	−29.52104	44	44	—	−29.70865
—	—	60	−29.70343	46	46	—	−29.70877
—	—	70	−29.70387	48	48	—	−29.70884
—	—	80	−29.70388	50	50	—	−29.70886
—	—	90	−29.70388	60	60	—	−29.70888

^{a} Corresponding *grid*-DFT value is −29.70899 a.u.[180]

as it offers much flexibility. At this stage, we provide some preliminary results on a general non-uniform grid with arbitrary number of points. For this, Cl_2 is chosen as a representative in Table 8. These are performed with LDA XC functional, using a grid spacing of $h_r = 0.3$, keeping the same pseudopotential. A snapshot of computed energies, from two sets (labelled I, II) is offered for illustration. In Set I in the left-hand side, we first vary N_z, the number of grid points along internuclear axis (z) keeping same along xy plane static at certain reasonable value ($N_x = N_y = 30$). This can be appreciated from a glance at Table 1; the calculated total electronic energy of Cl_2 is −28.41626 a.u. in case of Set A using the LDA XC functional. As N_z is increased from 30 to 40, there is a huge jump in energy value (≈ 4.16 a.u); with any further increase, it slowly converges to −29.66505 a.u., at around $N_z = 70$. Applying the same procedure for fixed N_x, N_y at 36, offers an energy value of −29.70388, converging again nearly at $N_z = 70$. There is no need to show all these details. Rather it suffices to mention that for each such (N_x, N_y) pair, energies attain convergence for a certain N_z, which in this case happens to be around 70. Then Set II, in the right side, varies the number of points, N_x, N_y along xy plane keeping N_z constant at 70. It is apparent that formal convergence of energy takes place for $N_x = N_y = 48$. A re-look at Table 1 reveals that the comparable convergence in a uniform grid occurs for Set D corresponding to $N_x = N_y = N_z = 64$ and $h_r = 0.3$. This leads to a drastic reduction in the total number of points, which eventually helps to achieve an enhanced computational scaling compared to the uniform grid. This will have a tremendous bearing on routine calculations of larger systems. A more thorough analysis of these aspects will be undertaken later.

4 Future and outlook

We have demonstrated the validity and feasibility of a Gaussian-based LCAO-MO approach to DFT using CCG in the context of atomic and molecular properties. This produces practically identical results with

those obtained from other grid-based/grid-free quantum chemistry programs available in literature. The classical Coulomb potential is obtained by means of a Fourier convolution technique, accurately and efficiently in the real-space grid. The matrix elements of all two-body potentials are computed through numerical integration. Quantities such as HOMO energies, potential energy curves, orbital energies, atomization energies are reproduced very well. The method performs quite decently for both local and non-local XC density functionals. While we focused mainly on pseudopotential situations, it is equally applicable for all-electron calculations as well. So far, these works were restricted to a uniform grid; here we report some exploratory calculations for Cl_2 in a more general *non-uniform* grid, for the first time. A more thorough analysis is in progress. Energies are variationally well-founded.

The success of the CCG method, as delineated above, encourages us to employ (i) more appropriate effective core potentials suitable for DFT calculations (ii) elaborate, extended and sophisticated basis sets (iii) superior quality density functionals in our future works. It would be worthwhile to assess its merit and suitability for geometry optimization of molecules using CCG. In this regard, a non-uniform grid, could be quite useful, for which some early results have been presented here and needs further inspection. An important concern would be to lessen the computational cost by incorporating a linear scaling approach; apparently the most promising will be real-space multi-grid techniques. A useful application of this methodology could be in the real-time dynamical studies, particularly laser-atom/molecule interactions in intense/super-intense regimes within a TDDFT framework. Some of these works are currently being investigated by us.

Acknowledgements

References

1   *Modern Electronic Structure Theory*, ed. D. R. Yarkony, World Scientific, Singapore, 1995.
2   A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, Dover, New York, 1996.

3 J. Simons and J. Nichols, *Quantum Mechanics in Chemistry*, Oxford University Press, New York, 1997.

4 W. Kohn, *Rev. Mod. Phys.*, 1999, **71**, 1253.

5 T. Helgaker, P. Jørgensen and J. Olsen, *Modern Electronic Structure Theory*, John Wiley, New York, 2000.

6 M. Springborg, *Methods of Electronic Structure Calculations*, Wiley, Chichester, New York, 2000.

7 D. C. Young, *Computational Chemistry: A Practical Guide for Applying Techniques to Real-World Problems*, John Wiley, New York, 2001.

8 W. M. C. Foulkes, L. Mitas, R. J. Needs and G. Rajagopal, *Rev. Mod. Phys.*, 2001, **73**, 33.

9 E. Lewars, *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, Kluwer Academic, Netherlands, 2003.

10 C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, John Wiley, New York, 2004.

11 R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press, Cambridge, UK, 2004.

12 E. O. Hoffman, *Progress in Quantum Chemistry Research*, Nova Science Publishers, New York, 2007.

13 F. Jensen, *Introduction to Computational Chemistry*, John Wiley, New York, 2007.

14 S. Tolosa, J. A. Sansón and A. Hidalgo, in *Quantum Chemistry Research Trends*, ed. M. P. Kaisas, Nova Science Publishers, New York, 2007.

15 W. T. Borden, *J. Am. Chem. Soc.*, 2011, **133**, 14841.

16 *Specialist Periodical Reports: Chemical Modelling, Applications and Theory*, ed. M. Springborg, Royal Society of Chemistry, London, 2012, vol. 9.

17 *Specialist Periodical Reports: Chemical Modelling, Applications and Theory*, ed. M. Springborg and J.-O. Joswig, Royal Society of Chemistry, London, 2014, vol. 10.

18 S. Saebo and P. Pulay, *Annu. Rev. Phys. Chem.*, 1993, **44**, 213.

19 R. J. Bartlett, *J. Phys. Chem.*, 1989, **93**, 1697.

20 B. O. Roos, K. Andersson, M. P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot and M. Merchán, *Adv. Chem. Phys.*, 1996, **93**, 219.

21 R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.

22 *Recent Advances in Density Functional Methods*, ed. D. P. Chong, World Scientific, Singapore, 1995, vol. I.

23 *Recent Developments and Applications of Modern DFT*, ed. J. M. Seminario, Elsevier, Amsterdam, 1996.

24 *Density Functionals: Theory and Applications*, ed. D. Joubert, Springer, Berlin, 1998.

25 *Density Functional Theory: Recent Progress and New Directions*, ed. J. F. Dobson, G. Vignale and M. P. Das, Plenum, New York, 1998.

26 W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, John Wiley, New York, 2001.

27 C. Fiolhais, F. Nogueira and M. Marques, *A Primer in Density Functional Theory*, Springer, Berlin, 2003.

28 N. I. Gidopoulos and S. Wilson, *The Fundamentals of Electron Density, Density Matrix and Density Functional Theory in Atoms, Molecules and the Solid State*, Springer, Berlin, 2003.

29 H. Hu and W. Yang, *Annu. Rev. Phys. Chem.*, 2008, **59**, 573.

30 C. J. Cramer and D. J. Truhlar, *Phys. Chem. Chem. Phys.*, 2009, **11**, 10757.

31 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Chem. Rev.*, 2011, **112**, 289.

32 K. Burke, *J. Chem. Phys.*, 2012, **136**, 150901.
33 *Theoretical and Computational Developments in Modern Density Functional Theory*, ed. A. K. Roy, Nova Science Publishers, New York, 2012.
34 A. D. Becke, *J. Chem. Phys.*, 2014, **140**, 18A301.
35 R. O. Jones, *Rev. Mod. Phys.*, 2015, **87**, 897.
36 P. A. M. Dirac, *Math. Proc. Cambridge Philos. Soc.*, 1930, **26**, 376.
37 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200.
38 J. P. Perdew and Y. Wang, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1992, **45**, 13244.
39 A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098.
40 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785.
41 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
42 A. D. Becke, *J. Chem. Phys.*, 1988, **88**, 1053.
43 A. D. Becke and M. R. Roussel, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1989, **39**, 3761.
44 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.
45 A. K. Roy, *Int. J. Quantum Chem.*, 2008, **108**, 837.
46 A. K. Roy, *Chem. Phys. Lett.*, 2008, **461**, 142.
47 A. K. Roy, in *Handbook of Computational Chemistry Research*, ed. C. T. Collett and C. D. Robson, Nova Publishers, New York, 2009.
48 A. K. Roy, *Trends Phys. Chem.*, 2010, **14**, 27.
49 A. K. Roy, *J. Math. Chem.*, 2011, **49**, 1687.
50 G. J. Martyna and M. E. Tuckerman, *J. Chem. Phys.*, 1999, **110**, 2810.
51 P. Minary, M. E. Tuckerman, K. A. Pihakari and G. J. Martyna, *J. Chem. Phys.*, 2002, **116**, 5351.
52 W. R. Wadt and P. J. Hay, *J. Chem. Phys.*, 1985, **82**, 284.
53 P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 299.
54 P. O. Löwdin, *Phys. Rev.*, 1955, **97**, 1474.
55 L. H. Thomas, *Math. Proc. Cambridge Philos. Soc.*, 1927, **23**, 542.
56 E. Fermi, *Rend. Accad. Naz. Lincei*, 1927, **6**, 32.
57 E. Fermi, *Z. Phys.*, 1928, **48**, 73.
58 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864.
59 A. K. Roy and S.-I. Chu, *J. Phys., B*, 2002, **35**, 2075.
60 B. M. Deb and P. K. Chattaraj, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1989, **39**, 1696.
61 B. K. Dey and B. M. Deb, *Int. J. Quantum Chem.*, 1998, **67**, 251.
62 B. K. Dey and B. M. Deb, *Int. J. Quantum Chem.*, 1998, **70**, 441.
63 A. K. Roy and S.-I. Chu, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2002, **65**, 043402.
64 M. Sadhukhan and B. M. Deb, *Eur. Phys. Lett.*, 2011, **94**, 50008.
65 Vikas, *J. Comput. Chem.*, 2011, **32**, 2404.
66 Vikas, *Int. J. Quantum Chem.*, 2013, **113**, 139.
67 M. Sadhukhan, A. K. Roy, P. K. Panigrahi and B. M. Deb, *Int. J. Quantum Chem.*, 2016, **116**, 377.
68 Y. A. Wang and E. A. Carter, in *Theoretical Methods in Condensed Phase Chemistry*, ed. S. D. Schwartz, Kluwer, Dordrecht, 2000.
69 Y. A. Wang, N. Govind and E. A. Carter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 13465.
70 Y. A. Wang, N. Govind and E. A. Carter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **60**, 16350.
71 G. S. Ho, V. L. Ligneres and E. A. Carter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **78**, 045105.
72 V. Gavini, K. Bhattacharya and M. Oritz, *J. Mech. Phys. Solids*, 2007, **55**, 697.

73   C. Huang and E. A. Carter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **81**, 045206.

74   W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.

75   G. Te Velde and E. J. Baerends, *J. Comput. Phys.*, 1992, **99**, 84.

76   P. M. W. Gill, B. G. Johnson and J. A. Pople, *Chem. Phys. Lett.*, 1993, **209**, 506.

77   O. Treutler and R. Ahlrichs, *J. Chem. Phys.*, 1995, **102**, 346.

78   B. Delley, *J. Chem. Phys.*, 1990, **92**, 508.

79   R. Lindh, P. Malmqvist and L. Gagliardi, *Theor. Chem. Acc.*, 2001, **106**, 178.

80   M. Franchini, P. H. T. Philipsen and L. Visscher, *J. Comput. Chem.*, 2013, **34**, 1819.

81   V. I. Lebedev and A. L. Skorokhodov, *Russ. Acad. Sci., Dokl., Math.*, 1992, **45**, 587.

82   V. I. Lebedev and D. N. Laikov, *Dokl. Akad. Nauk*, 1999, **366**, 741.

83   A. D. McLaren, *Math. Comput.*, 1963, **17**, 361.

84   C. W. Murray, N. C. Handy and G. J. Laming, *Mol. Phys.*, 1993, **78**, 997.

85   A. El-Sherbiny and R. A. Poirier, *J. Comput. Chem.*, 2004, **25**, 1378.

86   S.-H. Chien and P. M. W. Gill, *J. Comput. Chem.*, 2006, **27**, 730.

87   M. E. Mura and P. J. Knowles, *J. Chem. Phys.*, 1996, **104**, 9848.

88   P. M. W. Gill and S.-H. Chien, *J. Comput. Chem.*, 2003, **24**, 732.

89   K. Kakhiani, K. Tsereteli and P. Tsereteli, *Comput. Phys. Commun.*, 2009, **180**, 256.

90   M. Mori, *J. Comput. Appl. Math.*, 1985, **12**, 119.

91   M. Mori and M. Sugihara, *J. Comput. Appl. Math.*, 2001, **127**, 287.

92   M. Muhammad and M. Mori, *J. Comput. Appl. Math.*, 2003, **161**, 431.

93   M. Krack and A. M. Köster, *J. Chem. Phys.*, 1998, **108**, 3226.

94   J. M. Pérez-Jordá, A. D. Becke and E. San-Fabián, *J. Chem. Phys.*, 1994, **100**, 6520.

95   V. Termath and J. Sauer, *Chem. Phys. Lett.*, 1996, **255**, 187.

96   M. Challacombe, *J. Chem. Phys.*, 2000, **113**, 10037.

97   J. I. Rodriguez, D. C. Thompson, P. W. Ayers and A. M. Köster, *J. Chem. Phys.*, 2008, **128**, 224103.

98   L. Füsti-Molnár and P. Pulay, *J. Chem. Phys.*, 2002, **117**, 7827.

99   L. Füsti-Molnár, *J. Chem. Phys.*, 2003, **119**, 11080.

100 S. T. Brown, L. Füsti-Molnár and J. Kong, *Chem. Phys. Lett.*, 2006, **418**, 490.

101 Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, 2003.

102 A. Brandt, *Math. Comput.*, 1977, **31**, 333.

103 C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, O. Diéguez and M. C. Payne, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2002, **66**, 035119.

104 N. D. M. Hine, J. Dziedzic, P. D. Haynes and C.-K. Skylaris, *J. Chem. Phys.*, 2011, **135**, 204103.

105 C.-M. Chang, Y. Shao and J. Kong, *J. Chem. Phys.*, 2012, **136**, 114112.

106 R. Beatson and L. Greengard, *Wavelets, Multilevel Methods and Elliptic PDEs*, 1997, **1**, 1.

107 R. D. Skeel, I. Tezcan and D. J. Hardy, *J. Comput. Chem.*, 2002, **23**, 673.

108 D. York and W. Yang, *J. Chem. Phys.*, 1994, **101**, 3298.

109 J. R. Chelikowsky, N. Troullier and Y. Saad, *Phys. Rev. Lett.*, 1994, **72**, 1240.

110 J. R. Chelikowsky, N. Troullier, K. Wu and Y. Saad, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **50**, 11355.

111 M. Mundt and S. Kümmel, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **76**, 035413.

112 M. A. L. Marques, N. T. Maitra, F. M. S. Nogueira, E. K. U. Gross and A. Rubio, *Fundamentals of Time-Dependent Density Functional Theory*, Springer Science & Business Media, 2012.

113 X. Andrade and A. Aspuru-Guzik, *J. Chem. Theory Comput.*, 2013, **9**, 4360.

114 J. R. Chelikowsky, A. T. Zayak, T. Chan, M. L. Tiago, Y. Zhou and Y. Saad, *J. Phys.: Condens. Matter*, 2009, **21**, 064207.

115 T. L. Beck, *Rev. Mod. Phys.*, 2000, **72**, 1041.

116 A. Natan, A. Benjamini, D. Naveh, L. Kronik, M. L. Tiago, S. P. Beckman and J. R. Chelikowsky, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **78**, 075109.

117 D. R. Bowler and T. Miyazaki, *Rep. Prog. Phys.*, 2012, **75**, 036503.

118 J. Bernholc, M. Hodak and W. Lu, *J. Phys.: Condens. Matter*, 2008, **20**, 294205.

119 S. Mohr, L. E. Ratcliff, P. Boulanger, L. Genovese, D. Caliste, T. Deutsch and S. Goedecker, *J. Chem. Phys.*, 2014, **140**, 204110.

120 L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman and R. Schneider, *J. Chem. Phys.*, 2008, **129**, 014109.

121 E. L. Briggs, D. J. Sullivan and J. Bernholc, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1995, **52**, R5471.

122 J.-L. Fattebert and J. Bernholc, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **62**, 1713.

123 M. M. G. Alemany, M. Jain, M. L. Tiago, Y. Zhou, Y. Saad and J. R. Chelikowsky, *Comput. Phys. Commun.*, 2007, **177**, 339.

124 J.-I. Iwata, D. Takahashi, A. Oshiyama, T. Boku, K. Shiraishi, S. Okada and K. Yabana, *J. Comput. Phys.*, 2010, **229**, 2339.

125 Y. Fujimoto and A. Oshiyama, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **81**, 205309.

126 M. Heiskanen, T. Torsti, M. J. Puska and R. M. Nieminen, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2001, **63**, 245106.

127 J. J. Mortensen, L. B. Hansen and K. W. Jacobsen, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2005, **71**, 035109.

128 N. A. Modine, G. Zumbach and E. Kaxiras, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1997, **55**, 10289.

129 I.-H. Lee, Y.-H. Kim and R. M. Martin, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **61**, 4397.

130 J. E. Pask and P. A. Sterne, *Modell. Simul. Mater. Sci. Eng.*, 2005, **13**, R71.

131 E. Tsuchida and M. Tsukada, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1995, **52**, 5573.

132 S. R. White, J. W. Wilkins and M. P. Teter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1989, **39**, 11355.

133 J. Ackermann, B. Erdmann and R. Roitzsch, *J. Chem. Phys.*, 1994, **101**, 7643.

134 J. E. Pask, B. M. Klein, C. Y. Fong and P. A. Sterne, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**, 12352.

135 H. Yu, A. D. Bandrauk and V. Sonnad, *Chem. Phys. Lett.*, 1994, **222**, 387.

136 E. Hernández, M. J. Gillan and C. M. Goringe, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1997, **55**, 13485.

137 E. L. Briggs, D. J. Sullivan and J. Bernholc, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 14362.

138 F. Gygi and G. Galli, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1995, **52**, R2229.

139 D. Braess and R. Verfürth, *J. Numer. Anal.*, 1990, **27**, 979.

140 S. Brenner and R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer Science & Business Media, 2007.

141 T. A. Arias, *Rev. Mod. Phys.*, 1999, **71**, 267.

142 J. C. Slater, *Phys. Rev.*, 1930, **36**, 57.

143 S. F. Boys, *Proc. R. Soc. London, Ser. A*, 1950, **200**, 542.

144 K. Cho, T. A. Arias, J. D. Joannopoulos and P. K. Lam, *Phys. Rev. Lett.*, 1993, **71**, 1808.

145 J. Ihm, A. Zunger and M. L. Cohen, *J. Phys., C*, 1979, **12**, 4409.

146 S. Obara and A. Saika, *J. Chem. Phys.*, 1986, **84**, 3963.

147 H. Sambe and R. H. Felton, *J. Chem. Phys.*, 1975, **62**, 1122.

148 B. I. Dunlap, J. W. D. Connolly and J. R. Sabin, *J. Chem. Phys.*, 1979, **71**, 3396.

149 B. I. Dunlap, J. W. D. Connolly and J. R. Sabin, *J. Chem. Phys.*, 1979, **71**, 4993.

150 J. A. Pople, P. M. W. Gill and B. G. Johnson, *Chem. Phys. Lett.*, 1992, **199**, 557.

151 K. Huang and C. N. Yang, *Phys. Rev.*, 1957, **105**, 767.

152 M. L. Cohen, M. Schlüter, J. R. Chelikowsky and S. G. Louie, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1975, **12**, 5575.

153 A. J. Leggett, *Rev. Mod. Phys.*, 2001, **73**, 307.

154 M. L. Cohen, *Phys. Rep.*, 1984, **110**, 293.

155 P. Schwerdtfeger, B. Assadollahzadeh, U. Rohrmann, R. Schäfer and J. R. Cheeseman, *J. Chem. Phys.*, 2011, **134**, 204102.

156 L. Maron and C. Teichteil, *Chem. Phys.*, 1998, **237**, 105.

157 P. Schwerdtfeger, T. Fischer, M. Dolg, G. Igel-Mann, A. Nicklass, H. Stoll and A. Haaland, *J. Chem. Phys.*, 1995, **102**, 2050.

158 P. Schwerdtfeger, J. R. Brown, J. K. Laerdahl and H. Stoll, *J. Chem. Phys.*, 2000, **113**, 7110.

159 I. G. Batyrev, J.-H. Cho and L. Kleinman, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2001, **63**, 172420.

160 S. Huzinaga and A. A. Cantu, *J. Chem. Phys.*, 1971, **55**, 5543.

161 T. Zeng and M. Klobukowski, *J. Chem. Phys.*, 2009, **130**, 204107.

162 D. G. Fedorov and M. Klobukowski, *Chem. Phys. Lett.*, 2002, **360**, 223.

163 J. Andzelm, E. Radzio, Z. Barandiarán and L. Seijo, *J. Chem. Phys.*, 1985, **83**, 4565.

164 W. H. E. Schwarz, *Theor. Chim. Acta*, 1971, **23**, 147.

165 L. R. Kahn and W. A. Goddard III, *J. Chem. Phys.*, 1972, **56**, 2685.

166 A. Redondo, W. A. Goddard III and T. C. McGill, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1977, **15**, 5038.

167 H. Preuss, H. Stoll, U. Wedig and T. Krüger, *Int. J. Quantum Chem.*, 1981, **19**, 113.

168 P. Schwerdtfeger, H. Stoll and H. Preuss, *J. Phys., B*, 1982, **15**, 1061.

169 L. R. Kahn, *Int. J. Quantum Chem.*, 1984, **25**, 149.

170 L. E. McMurchie and E. R. Davidson, *J. Comput. Phys.*, 1981, **44**, 289.

171 P. Fuentealba, H. Stoll, L. Von Szentpaly, P. Schwerdtfeger and H. Preuss, *J. Phys., B*, 1983, **16**, L323.

172 D.-K. Lee, G.-H. Jeung and Y. S. Lee, *Int. J. Quantum Chem*, 2009, **109**, 1975.

173 C. Thierfelder and P. Schwerdtfeger, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2010, **82**, 062503.

174 P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 270.

175 D. Figgen, K. A. Peterson, M. Dolg and H. Stoll, *J. Chem. Phys.*, 2009, **130**, 164108.

176 P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 299.

177 D. R. Hamann, M. Schlüter and C. Chiang, *Phys. Rev. Lett.*, 1979, **43**, 1494.

178 E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney and D. Sorensen, *LAPACK Users' Guide*, SIAM, 1999, vol. 9.

179 M. Frigo and S. G. Johnson, *Proc. IEEE*, 2005, **93**, 216.

180 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Hensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347.

181 K. R. Glaesemann and M. S. Gordon, *J. Chem. Phys.*, 1998, **108**, 9959.

182 B. Miehlich, A. Savin, H. Stoll and H. Preuss, *Chem. Phys. Lett.*, 1989, **157**, 200.

183 M. Filatov and W. Thiel, *Int. J. Quantum Chem.*, 1997, **62**, 603.

184 M. Filatov and W. Thiel, *Mol. Phys.*, 1997, **91**, 847.

185 Density Functional Repository, *Quantum Chemistry Group*, CCLRC Daresbury Laboratory, Daresbury, Cheshire, UK, 2001.

186 *NIST Computational Chemistry Comparisons and Benchmark Database, NIST Standard Reference Database, Number, Release 14*, ed. R. D. Johnson III, NIST, Gaithersburg, MD, 2006.

187 H. Y. Afeefy, J. E. Liebman and S. E. Stein, in *NIST Chemistry Webbook, NIST Standard Reference Database, Number 69*, ed. P. J. Linstrom and W. G. Mallard, NIST, Gaithersburg, MD, 2005.

188 R. Van Leeuwen and E. J. Baerends, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1994, **49**, 2421.

189 P. R. T. Schipper, O. V. Gritsenko, S. J. A. Van Gisbergen and E. J. Baerends, *J. Chem. Phys.*, 2000, **112**, 1344.

Molecular-level visualisation for educational purposes[†]

Roy Tasker

1 Introduction

The distinctive feature of chemistry is its molecular-level perspective to both explain observable phenomena, and to create new substances to produce new phenomena. Award-winning chemist Richard Zare[1] characterizes chemists as "...highly visual people who want to 'see' chemistry and to picture molecules and how chemical transformations happen." Another eminent chemist, Peter Atkins,[2] expressed the need to convey this to novices before they can begin to understand the great ideas of chemistry – "The principle target of our education should be to find a way to bridge the imagined to the perceived. By that, I mean, we should show people how to look at a lump of matter, and in their mind's eye, see it as a collection of atoms and molecules." This priority is often lost in chemistry education where novices are drowned in the symbolism and mathematical descriptions of chemical behaviour *before* they have the mental models[‡] of the molecular world to make meaning from them. If chemical species are shown, they are represented in isolation, ignoring the importance of chemical speciation,[3] so important for understanding reactivity, bioavailability, chemical toxicity, environmental fate and transport.

This chapter will start by describing the need to help novices to build mental models of the molecular level, and then focus on the scientific and pedagogical challenges of portraying the invisible molecular world with visualisations[2] for educational purposes.

1.1 The need to build accurate mental models of the molecular level

Johnstone[4] offered an explanation for why the sciences, and chemistry in particular, is so difficult to learn. He proposed that a chemistry expert thinks and moves seamlessly between three levels; the *macro* (referred to as the *observable* level in this chapter), the *sub-micro* (referred to as the *molecular* level here), and the *representational* (referred to as the *symbolic* level here). The observable level involves phenomena that are visible, tangible, and often perceptible with other senses. Mental models of the imperceptible molecular level consist of dynamic, imaginary images used to explain observations in terms of structures and processes involving

Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, IN 47907, USA. E-mail: rtasker@purdue.edu

[‡]Dynamic internal images in the 'mind's eye' are visual *mental models*, while the external images seen with the eyes will be referred to as *visualisations*.
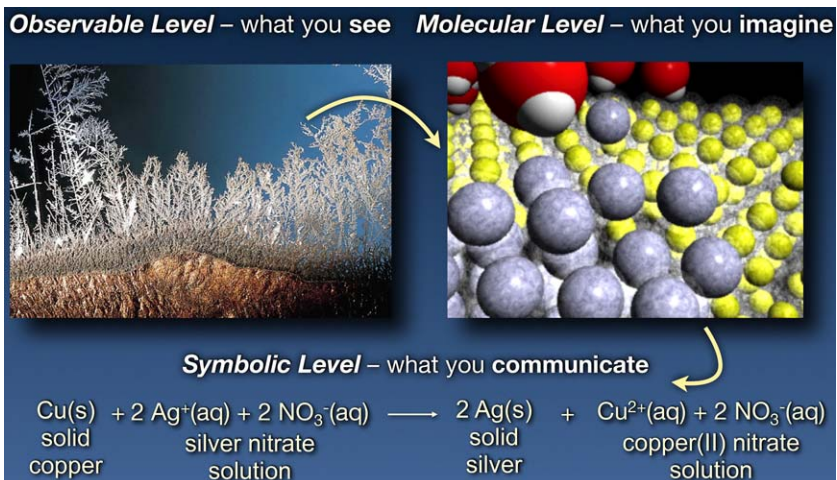
Fig. 1 The reaction of copper with aqueous silver nitrate can be represented at three thinking levels.

atoms, ions and molecules. Observed phenomena, and molecular-level structures and processes, are both represented using chemical notation and language, and rationalised mathematically, at the symbolic level.

Often a term like "oxygen" is used without a clear indication of the thinking level involved; for example, does "water contain the element oxygen"? Does this refer to the elemental substance dioxygen, O_2, dissolved in water, or the oxygen atoms in water molecules? With the frequent confusion between thinking levels, and lack of attention to specificity in chemical language and symbolism, no wonder novices have difficulty with chemistry!

Figure 1 illustrates the three thinking levels for the redox reaction that occurs when silver nitrate solution is added to solid copper. Dendritic silver crystals growing on the surface of the copper can be perceived at the observable level. At the molecular level an animation can portray the dynamic, imperceptible formation of silver atoms adhering to a growing cluster of silver atoms. The equation summarises the reaction at the symbolic level.

In a study of our students' understanding of what is happening at the molecular level to give rise to the three observations in this reaction; namely, the solution turning blue, shiny crystals forming on the copper surface, and the amount of copper metal decreasing, we asked them to describe in words and drawings what was happening at the molecular level. A typical response from a student with an above-average high school chemistry background is shown in Fig. 2. Very often such students have a mental model of the molecular level that is determined by the chemical equation for the reaction (Fig. 2a). With an appropriate intervention[5] that requires students to declare their prior ideas, engage with carefully produced and presented visualisations, and apply their new model to a new but similar reaction, it is possible to improve a student's mental model (Fig. 2b).

Nakhleh[6] defined the term "misconception" as "any concept that differs from the commonly-accepted scientific understanding of the term". There is convincing evidence in the literature that many student difficulties and misconceptions in chemistry result from inadequate or inaccurate models at the molecular level.[7] Moreover, many of the misconceptions are common to students all over the world, and at different educational levels. Lack of meaningful learning is demonstrated by the fact that many students can solve traditional-style chemistry problems without understanding the underlying molecular processes.[8,9] The most
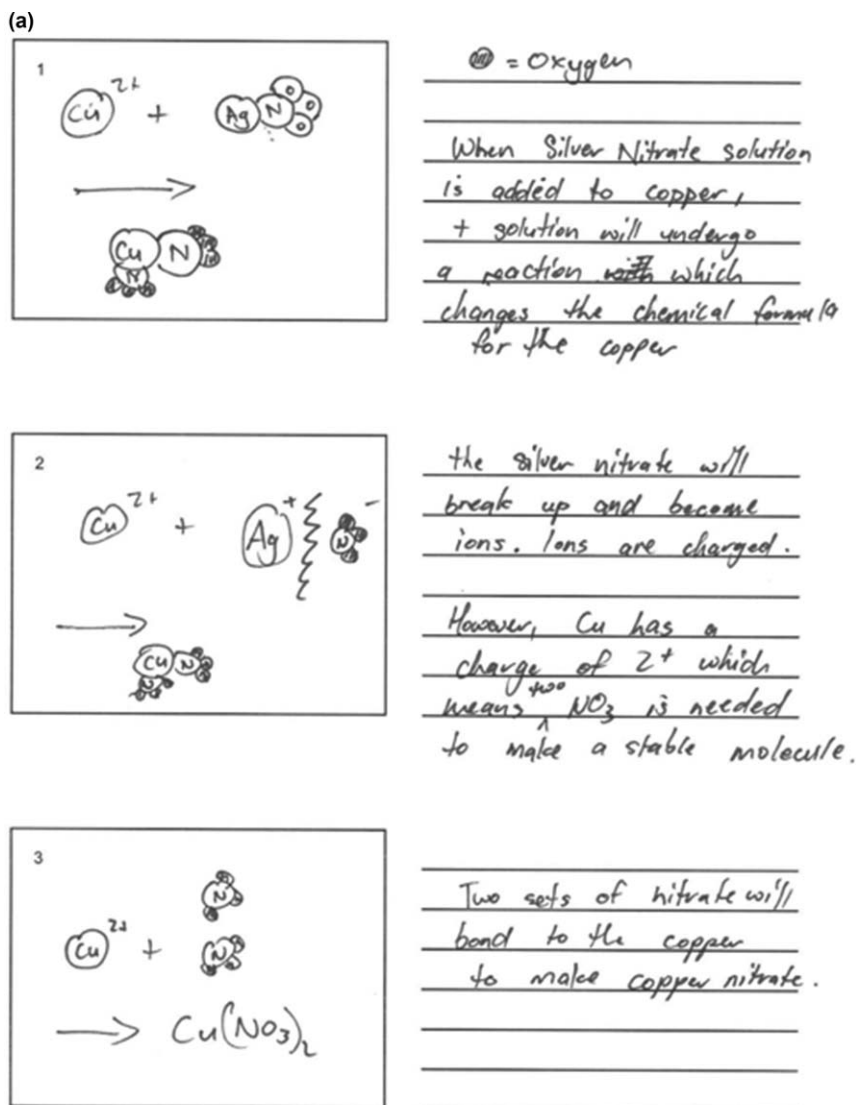
Fig. 2 A student drawing and description of what is happening in the reaction between copper and aqueous silver nitrate at the molecular level, before (a) and after (b) engaging with molecular-level animations in the *VisChem* learning design.[3]

(b)
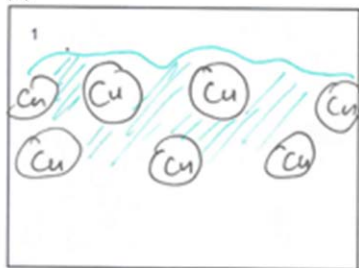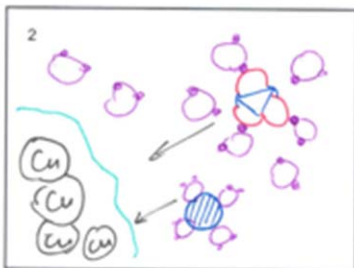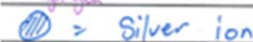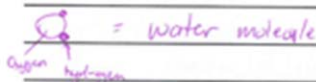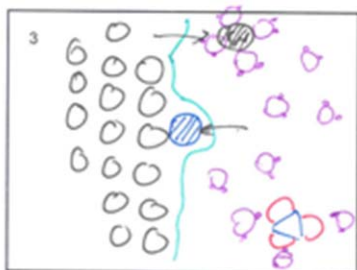
Fig. 2   (Continued.)

important finding is that many misconceptions are extraordinarily resistant to change.

Many misconceptions can be addressed by developing students' mental models of substances and reactions at the molecular level *before* representing them with chemical formulas and equations. After this step qualitative concepts can be expressed quantitatively, harnessing the enormous predictive power of mathematics. The challenges of portraying and presenting the molecular level are two-fold – scientific and pedagogical, as described below.

2 The scientific challenge of modelling the molecular level

The molecular world is multi-particulate, dynamic and, in the liquid and solid states, crowded; and the interactions are often subtle (*e.g.*, electron transfer, bonding rearrangement), complicated, and probabilistic. With changes occurring at or faster than the picosecond timescale, and occurring randomly in space between molecules and ions with fluxional shapes, any attempt to portray these changes requires artistic license. Consequently, any form of modelling at this level risks generating misconceptions unless care is taken to declare assumptions, approximations, and conventions used in the models. With these constraints, the dynamic molecular world can be visualised with models using scripted animations or calculated simulations.

2.1 Modelling with scripted animations

Animations are choreographed depictions, informed by theoretical considerations, but designed to show key events on 'centre-stage', with visual clarity, albeit with judicious 'artistic license' described below. In our *VisChem* project[10,11] the challenge was to balance the often-competing demands of:

• scientific accuracy – the crowding of molecules in the liquid state which would otherwise prevent viewing interactions between reacting species; the complicated internal molecular bond vibrations and rotations which would distract the viewer and add to the cognitive load; and the representation of molecules as diffuse electron clouds with embedded nuclei which would make distinguishing molecules too difficult

• 'artistic license' required for clear visual communication – such as depicting slightly less than realistic crowding in the liquid state to enable visibility beyond the nearest molecules; the absence of internal molecular bond vibrations to reduce the degree of movement; use of reflective boundary surfaces on atoms at their van der Waals radii; and greatly reduced speed of molecules, particularly in the gaseous state

• technical computing constraints on rendering times and file size – such as using a close-up view to limit the number of moving objects to be rendered; and not depicting the randomness in productive collisions to reduce the number of animation frames.

All these artistic license features have the potential to produce misconceptions, or at least mislead students, unless they are declared clearly and repeated often. With these reservations, the key animations produced in the *VisChem* project are summarised in Fig. 3, and listed in Table 1 with the misconceptions they target. These misconceptions have been identified in chemistry education research in studies with various age groups, from high school to university level, and in different educational systems. The *VisChem* animations are available for viewing and downloading from the *VisChem* web site.[12]

The *VisChem* animations have been incorporated into online learning resources associated with university-level chemistry textbooks, most
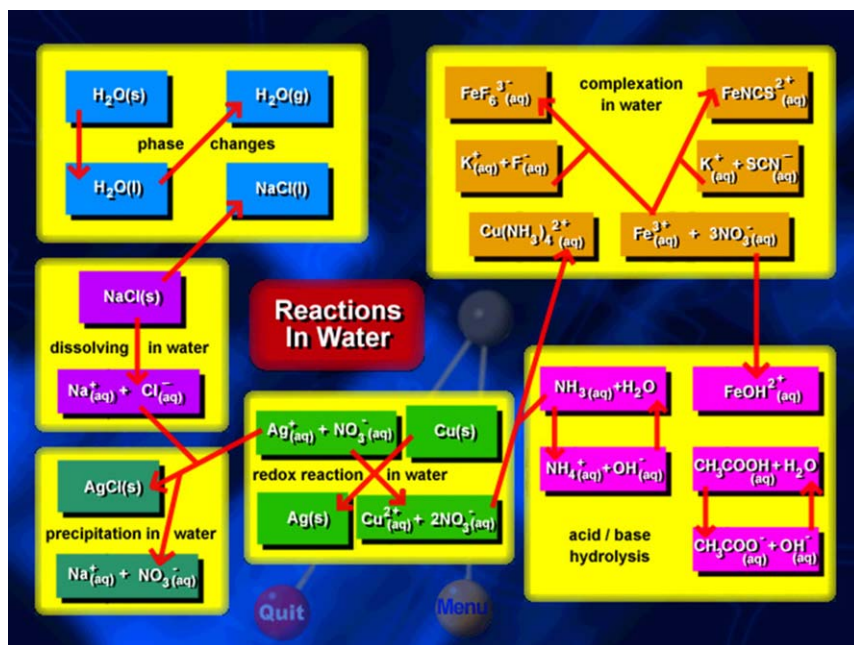
Fig. 3 Each substance and solution shown above is depicted in a *VisChem* animation. The physical and chemical changes indicated with arrows are also animated.

comprehensively and strategically in one in particular.[13] The animations discussed here portray inorganic substances, some in different states of matter, some undergoing physical changes, and some involved in common chemical reactions, as summarised in Fig. 3. All the individual atoms, molecules, ions, and hydrated ions in these animations are available as separate graphics or animations for use as 'symbol legends'. This is important because, as we will see in the next section, the animations are visually complex and familiarity with the iconography should never be assumed.

VisChem animations are novel in that they portray the vibrational movement in solid substances (*e.g.*, copper, sodium chloride in Table 1). This is important because the degree of movement is correlated with temperature, and students need to understand this correlation to interpret the significance of melting and boiling points in molecular-level terms. The confusion between the terms 'heat' and 'temperature' results from inadequate mental models of what they mean at the molecular level.

Many diagrams in textbooks depicting particles in the solid, liquid and gaseous states are misleading because the relative spacing between particles is inaccurate. Little wonder that students develop poor mental models of states of matter. The *VisChem* animations balance crowding with visual clarity.

Few students have a 'feel' for the average distance between ions in a solution of a given concentration. *VisChem* animations portray ionic solutions at a concentration of about 1 mol L^{-1}, with ions separated from

Table 1 List of *VisChem* animations (a–w) with chemical descriptions and targeted misconceptions. These animations are available for viewing and downloading from the *VisChem* web site and in the ESI.[†12]
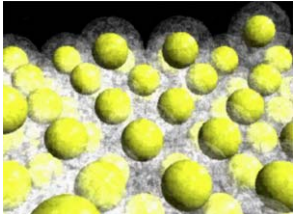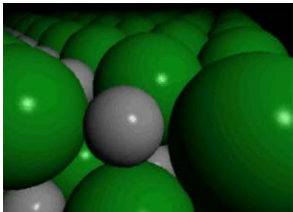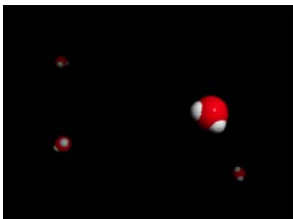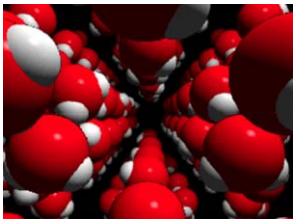
Selected frame	Description	Misconceptions targeted
Metal solid		
 a. Solid copper, Cu(s)	Close-packed copper atoms vibrate in an ordered lattice. Each atom is represented as a yellow copper(II) ion with its two electrons delocalised in a cloud.	Solids are conceived as composed of static atoms, ions, or molecules.
Ionic solids		
 b. Solid sodium chloride, NaCl(s)	Close packed sodium and chloride ions vibrate as net electrostatic attractions hold them together.	Discrete ion groups matching the ionic compound's chemical formula, *e.g.* NaCl, exist in ionic solids. Students believe it is not possible to point to where the ionic bonds are unless you know which chloride ions have accepted electrons from which sodium ions.
Molecular substances		
 c. Gaseous water, $H_2O(g)$	Water molecules move quickly in straight lines unless they collide, almost elastically. The average distance between molecules in a gas is much larger than in the liquid and solid states.	Students have difficulty imagining empty space. Matter is conceived as continuous. There is no vacuum.
 d. Solid water (ice), $H_2O(s)$	We move into one of the hexagonal channels in the ice structure, look around at the vibrating molecules attracted together by hydrogen bonds, and then move back out of the channel.	Students confuse *intra*molecular bonds with *inter*molecular bonds, and van der Waals forces with hydrogen bonds.

Table 1 (*Continued*)
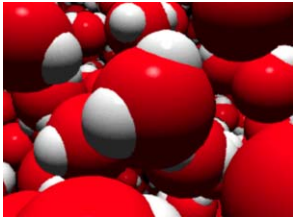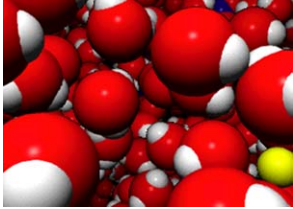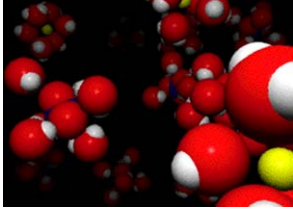
Selected frame	Description	Misconceptions targeted
 e. Liquid water, $H_2O(l)$	Water molecules move around, closely packed and attracted together by hydrogen bonds, with some molecules in transient ice-like clusters.	There is a tendency to believe that the molecules in ice are more densely packed than in liquid water. Students conceive of molecules in a liquid as spaced far enough apart to make the liquid compressible.
 f. Liquid oxygen, $O_2(l)$	Oxygen molecules in the liquid state move *almost* randomly with respect to one another.	Students conceive of molecules in a liquid as being reasonably spaced such that it could be compressible.
 g. Gaseous oxygen, $O_2(g)$	Oxygen molecules moving quickly in space, occasionally colliding.	Students believe that there is little reduction in density when a liquid changes to a gas.
Aqueous solutions		
 h. Aqueous copper(II) nitrate (∼1 M), $Cu^{2+}(aq) + 2NO_3^-(aq)$ 	Hydrated copper(II) and nitrate ions, and water molecules, in a 1 : 2 : 55 ratio, roam amongst the water molecules, with the occasional formation and dissociation of a transient ion pair. Solvent water molecules omitted in the version below show the proximity of hydrated ions at ∼1 M. Other *VisChem* animations show ∼1 M solutions of iron(III) nitrate, sodium nitrate, potassium thiocyanate, sodium chloride, and potassium fluoride.	Many students do not show separate ionic species in their representations of aqueous ionic solutions. They believe they exist in ion groups matching the misleading chemical formula $Cu(NO_3)_2(aq)$. Molecules and ions in aqueous solutions are not generally drawn in contact with each other. Ion hydration is not important. Some students think that dissolved particles go into empty spaces *inside* water molecules.

Table 1 (*Continued*)
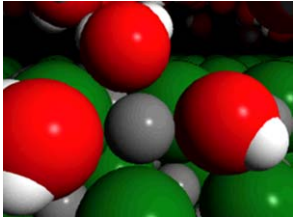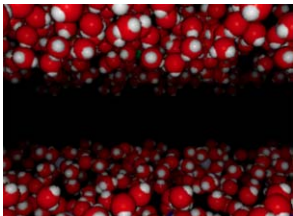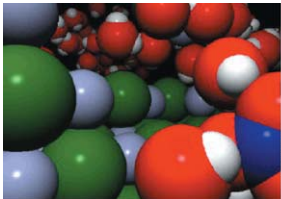
Selected frame	Description	Misconceptions targeted
Dissolution		
i. Solid sodium chloride dissolves $NaCl(s) \rightarrow Na^+(aq) + Cl^-(aq)$	Water molecules approach the vibrating ions in the lattice, hydrating the ions in a 'tug o' war' competition between ion-dipole forces and the electrostatic forces in the lattice.	The two terms *dissolving* and *melting* are often used interchangeably. Students rarely acknowledge the role of the polar nature of the water molecule in the process of dissolution. Students generally do not see dissolving as an interactive process but rather as a spontaneous separation of ions, followed by dispersal of of the ions amongst the water molecules.
Precipitation		
j. Mixed aqueous solutions $Na^+(aq) + Cl^-(aq) + Ag^+(aq) + NO_3^-(aq)$	The animation begins at the molecular surface of the silver nitrate solution, just prior to mixing with sodium chloride solution added from above. The mixing of solutions at the molecular level enables new combinations of ionic collision to occur.	Students cannot explain why a precipitate forms immediately when the solutions are mixed.
k. Silver chloride precipitation $Ag^+(aq) + Cl^-(aq) \rightarrow AgCl(s)$	In a solution containing silver, sodium, nitrate and chloride ions a silver ion and a chloride ion collide, and form a stable, neutral, poorly-hydrated ion pair. Another ion pair joins, and the resulting cluster joins a growing ionic lattice of silver chloride, with nitrate spectator ions in the background.	Students imagine that a precipitate is composed of 'ionic molecules', each containing a neutral ion pair or group of ions.

Table 1 (*Continued*)
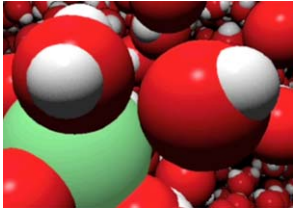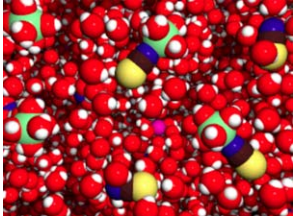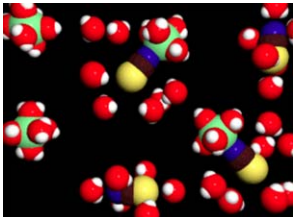
Selected frame	Description	Misconceptions targeted
Complexation		
 l. Water exchange on iron(III) $[(H_2O)_5Fe-OH_2]^{3+}(aq) + H_2O \rightleftharpoons$ $[(H_2O)_5Fe-OH_2]^{3+}(aq) + H_2O$	Exchange of an Fe(III)–bonded water molecule with a nearby solvent water molecule. The Fe(III) ion is represented here with its van der Waals radius.	Students do not realize that water molecules in the first coordination sphere exchange with surrounding water molecules. This is a necessary first step to understanding ligand substitution complexation reactions.
 m. Formation of a Cu(II) complex $Cu^{2+}(aq) + 4NH_3(aq) \rightleftharpoons [Cu(NH_3)_4]^{2+}(aq)$	Successive substitution of four Cu(II)–bonded water molecules with ammonia molecules, with Jahn Teller lengthening of axial bonds to coordinated water molecules. Cu(II) is represented by its ionic radius.	Students have difficulty imagining how a square planar complex can form in solution, or if they do, they imagine it occurs in one multi-substitution event.
Equilibrium		
 n. Iron(III) thiocyanate complexation equilibrium $[Fe(H_2O)_6]^{3+}(aq) + SCN^-(aq) \rightleftharpoons$ $[Fe(H_2O)_5 NCS]^{2+}(aq) + H_2O$ 	Formation and dissociation of the isothiocyanatoiron(III) complex (each available as a separate animation) occurs at the same rate at equilibrium. 　Potassium and nitrate spectator ions are also present. The version below leaves out solvent water molecules and spectator ions to focus attention on the two reactions, occurring simultaneously.	The use of everyday terms, "shift", "equal", "stress", "balance" when referring to equilibria can conjure up different visual ideas to students from those intended by the teacher. 　"Equilibrium" is seen as a static two-sided picture. This can be unintentionally reinforced by misleading metaphors and analogies. 　Equilibrium is seen as oscillating like a pendulum, and Le Chatelier's stress-then-shift logic reinforces this misconception. 　Lack of awareness of the dynamic nature of the equilibrium state.

Table 1 (*Continued*)

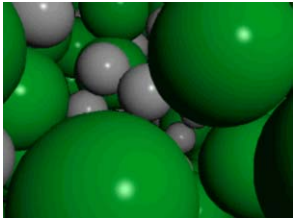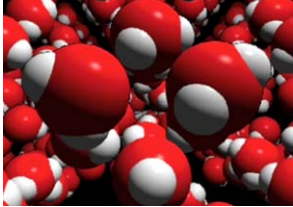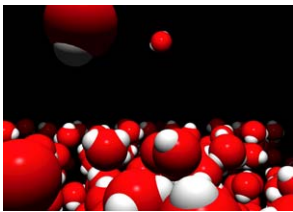| Selected frame	Description	Misconceptions targeted
Acid/Base Hydrolysis		
o. Acetate hydrolysis $CH_3COO^-(aq) + H_2O \rightarrow CH_3COOH(aq) + OH^-(aq)$	An acetate ion removes a proton from a water molecule, with some apparent difficulty, to form an acetic acid molecule and a hydroxide ion.	A base is just something that reacts with an acid.
p. Dissociation of acetic acid $CH_3COOH(aq) + H_2O \rightarrow CH_3COO^-(aq) + H_3O^+(aq)$	An acetic acid molecule transfers a proton to a water molecule, with some apparent difficulty, to form an acetate ion and a hydronium ion.	An acid is something which eats material away or which can burn you.
q. Autoionisation of water $H-O-H + H_2O \rightleftharpoons HO^- + H_3O^+$	Amongst the bustle of water molecules two come together with sufficient energy and appropriate orientation *very rarely* to exchange a proton, forming a hydronium ion and a hydroxide ion.	Students have difficulty imagining how pure water can contain any hydronium ions and hydroxide ions.
r. Iron(III) hydrolysis $[(H_2O)_5Fe-OH_2]^{3+}(aq) + H_2O \rightleftharpoons [(H_2O)_5Fe-OH]^{2+}(aq) + H_3O^+$	One of the coordinated water molecules on iron(III) loses a proton to a solvent water molecule. The charge density of the metal ion increases the acidity of the coordinated water molecule through polarisation.	Students have difficulty understanding how an aqueous solution containing metal ions could produce hydronium ions.

Table 1 (*Continued*)
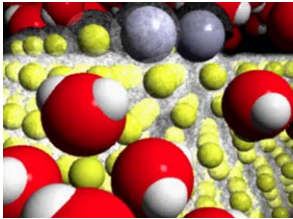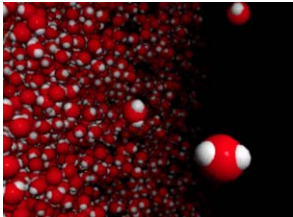
Selected frame	Description	Misconceptions targeted
Redox reaction		
s. Reduction of silver(I) by copper $Cu(s) + 2Ag^+(aq) \rightarrow Cu^{2+}(aq) + 2Ag(s)$	Hydrated silver ions migrate towards the copper surface. In a competition between electron transfer and ion-dipole attractions, electron 'cloud' (equivalent to one electron) moves onto each silver ion to form a silver atom. For every two Ag atoms formed there is concomitant release and hydration of one Cu^{2+} ion from a remote location on the metal lattice. Both anodic and cathodic sites are represented.	This reaction is often imagined as copper atoms each spontaneously donating two electrons to two silver ions to form atoms. Students have a "reluctance" to perceive or represent chemical reactions as multi-particulate. Students "cannot grasp the interactive nature of a chemical reaction" or that it often involves competition between opposing influences.
Physical changes		
t. Sodium chloride melting $NaCl(s) \rightarrow NaCl(l)$	To show the difference between dissolving and melting we see the thermal energy of the ions in solid NaCl increase until the ordered structure collapses to a more disordered liquid state.	Melting is just another word for dissolving. Students believe that there is a significant reduction in density when a solid melts.
u. Ice melting $H_2O(s) \rightarrow H_2O(l)$	Starting within the ice structure the camera moves down to the molecules on the lower surface closer to the heat source, just prior to melting. The thermal energy increases until the ordered structure collapses to a more disordered liquid state, even though the number of hydrogen bonds only decreases slightly.	There is a tendency to suggest that ice is more densely packed than liquid water.

Table 1 (*Continued*)

Selected frame	Description	Misconceptions targeted
 v. Evaporation of water $H_2O(l) \rightleftharpoons H_2O(g)$	Starting within liquid water the view moves up to the surface. Molecules occasionally break away, with some apparent difficulty, and some return. More leave than return.	Students believe that there is little reduction in density when a liquid changes to a gas. Students believe that molecules increase in size when moving from solid to liquid to gas.
 w. Inside a boiling water bubble $H_2O(l) \rightleftharpoons H_2O(g)$	Moving through the water molecules in the liquid bubble wall, the view suddenly breaks into the interior of the bubble, revealing only water molecules in the gaseous state. Some of the bubble wall can be seen in the background.	Bubbles in boiling water are made up of "heat" or "air" or "oxygen and hydrogen". Melting and boiling of molecular compounds are processes in which covalent bonds within molecules are broken.

each other by, on average, about three water molecules (Table 1h). Students are encouraged to imagine dilution of a solution in terms of separation of ions by more water molecules.

Animations of the molecular world can stimulate the imagination, bringing a new dimension to learning chemistry. One can imagine being inside a bubble of boiling water, or at the surface of silver chloride as it precipitates, as depicted in Table 1w and k respectively.

Most molecular-level processes involve competition between opposing processes. Atkins[1] recommended that educators should "teach [students] to judge between conflicting influences. That is the essence of our subject, for it is rare that a single property governs the outcome of a reaction. We need to train our students to judge the likely outcome of conflict". *VisChem* animations portray the 'tug-of-war' competition for a proton between the lone pair on the O atom in a water molecule and the electron pair on the O atom in the acetic acid molecule (Table 1p); and between the electrostatic lattice forces and ion–dipole interactions when water molecules interact with the surface ions in solid sodium chloride (Table 1i).

To the knowledge of the author there are no other collections of visually self-consistent animations portraying these common substances and reactions.

2.2 Modelling with molecular dynamics simulations

The focus of this section is to compare the features of calculated simulations and choreographed animations for showing *multimolecular* systems. There are many molecular dynamics software programs that

simulate substances and reactions but in the opinion of the author the most innovative use of a molecular dynamics force-field engine for educational purposes was developed by *Wavefunction Inc., Irvine, CA*, and used in *Odyssey* software.

The most obvious advantage that calculated simulations have over scripted animations is their scientific accuracy. Molecules in the gaseous state move with a Maxwell–Boltzmann distribution of speeds, and molecules in condensed states are appropriately crowded according to their freedom of movement.

The ability for students to build their own simulations to explore and discover new insights is a major advantage of simulation over animation. For example, students can develop their own 'feel' for the crowding of ions in a solution at a given concentration by building a simulation of say, 1 mol L^{-1} NaCl. Once they can see the ions are only separated from each other by about three water molecules on average, it is easy to explain that in such solutions ion pairing can be a factor to consider in calculations of colligative properties (see Fig. 4). This brings meaning to the molarity concentration scale, in much the same way that people have a 'feel' for a length of one metre.

Filtering what you see in a simulation provides new insights. In a simulation of ions in aqueous solution it is possible to filter out the solvent water molecules to focus on an ion and its first coordination sphere. For example, Fig. 5a shows the first coordination sphere of water molecules around a magnesium ion, and outer coordination molecules
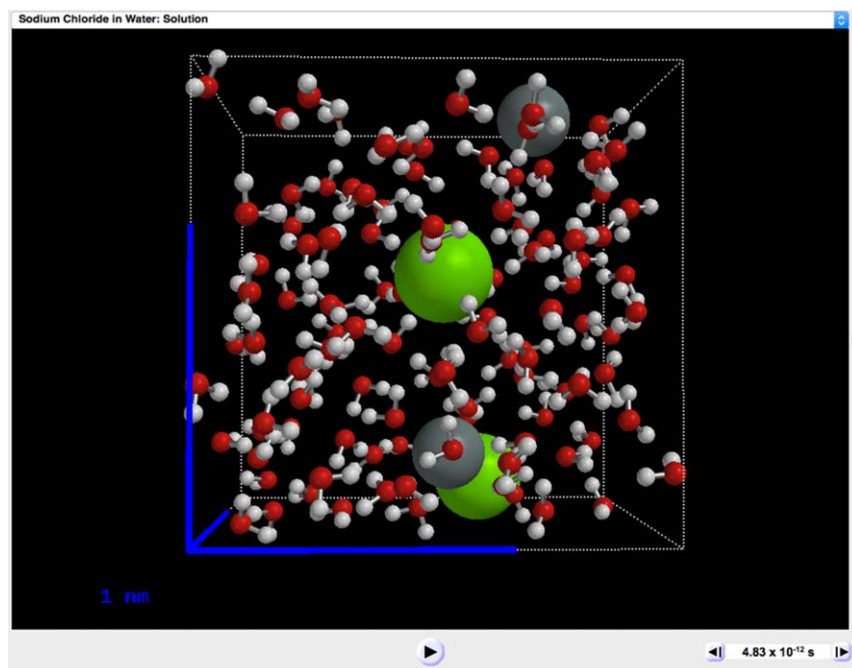

Fig. 4 A simulation of a solution with a known concentration, here 1 mol L^{-1} NaCl, in a cube containing Na$^+$ ions, Cl$^-$ ions, and H$_2$O molecules in a ratio of 1 : 1 : 55. The ions are separated from each other by only about three water molecules.
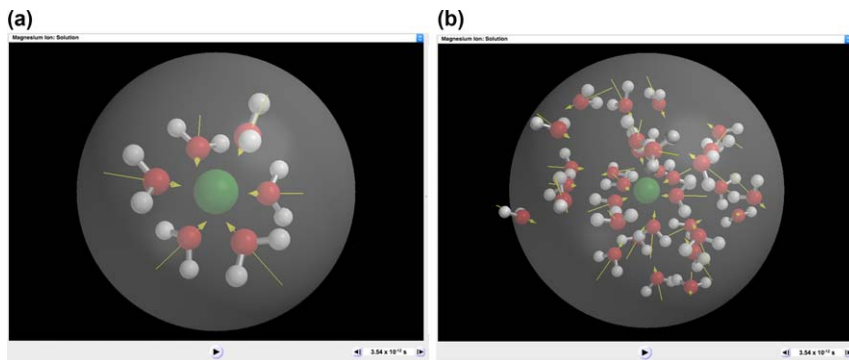
Fig. 5 (a) By selecting a hydrated ion like Mg^{2+}(aq) in a solution simulation, and making it a 'clipping centre', the first coordination sphere of hydration can be revealed clearly, with molecular dipoles visible. (b) By adjusting the size of the clipping sphere, water molecules outside the first coordination sphere can also be revealed.
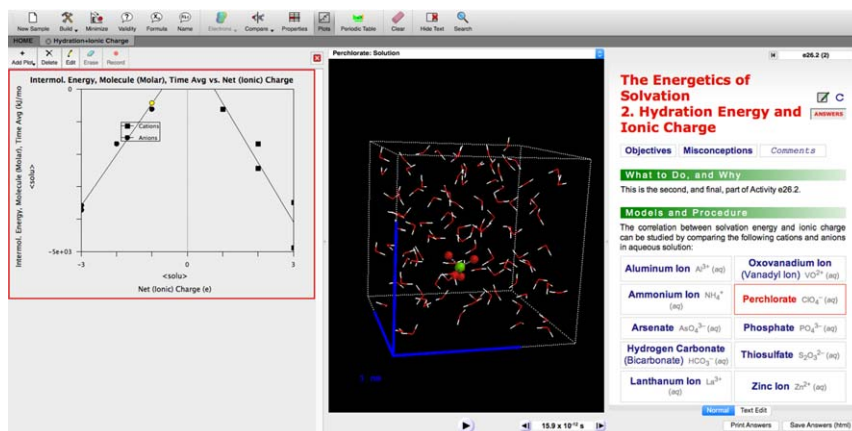

Fig. 6 An *Odyssey* experiment to study the dependence of time-averaged energy of solvation on net ionic charge after running simulations of ions with different charges surrounded by water molecules.

can be viewed by expanding the 'clipping centre' sphere around the ion (Fig. 5b).

Since all movements are calculated, experiments can be conducted by varying parameters. For example, Fig. 6 shows the dependence of time-averaged energy of solvation on net ionic charge after running simulations of ions with different charges surrounded by water molecules. In this way students can build their mental model of hydrated ions in solution, and apply the powerful explanatory concept of charge density to rationalise a lot of ionic reactions in aqueous solution.

Reaction mechanisms are usually described by listing the elementary steps involved as chemical equations. These are often misinterpreted by novices as interactions of just one molecule with another leading directly to the products. Reaction kinetics only makes sense if we visualise a dynamic, many-particle reaction mixture in which events are controlled at least partly by probabilities. One of the experiments in *Odyssey* involves
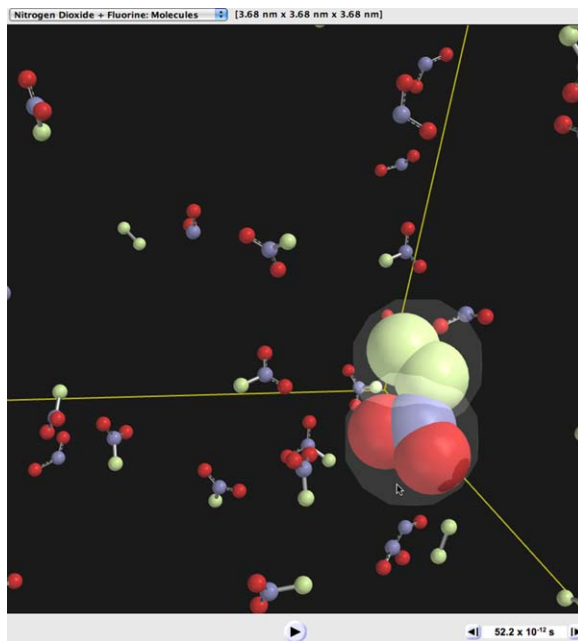
Fig. 7 In this simulation of the reaction between NO_2(g) and F_2(g) it is possible to 'freeze-frame' collisions to show whether intermediates are formed or not.

simulating the gaseous reaction of nitrogen(IV) oxide and fluorine where productive and unproductive collisions can be 'captured' by freeze-framing, revealing any formation of transient intermediates (see Fig. 7).

Simulations of many-particle systems do have limitations. Some phase changes such as freezing, and chemical changes like dissolution and precipitation are difficult to simulate with small sample sizes and pico-second time scales. Simulations of reactions involving proton transfer, electron transfer, and bond breaking/forming are possible, but only with special algorithms and a very detailed description of the reaction. Consequently, *optimum molecular modelling of substances and reactions for educational purposes requires a judicious combination of scripted animations—to show chemical changes that cannot be simulated—and calculated simulations—to show realistic, appropriately filtered representations with control over experimental variables.*

3 The pedagogical challenge of modelling the molecular level

The challenge of modelling the molecular world for educational purposes is to produce and present visualisations to enable novices to build useful mental models of structures and processes, and apply them in the way experts do. However, novices have difficulty making sense of visualizations of molecular phenomena without guidance, and there is an extensive research literature on the challenges of learning from dynamic visualisations.[14–17] Novices need to understand the visual symbols and significance of movements in each representation, the relationship

among different representations, and how the visualization relates to the target concepts.[18] Without such knowledge, students may only see bouncing balls when we show them a molecular-level visualisation that is intended to vividly demonstrate an observable phenomenon like melting. Interested readers are encouraged to consult these references for the general findings and recommendations, but this section will discuss their specific application to a specific complex animation (Table 1s) in the *VisChem* collection. This animation has been modified and then embedded in a learning sequence (learning design) to address most of the recommendations, *making the necessary compromises to ensure that the chemistry is presented accurately*. Both the learning design and the modified animation can be viewed in the *YouTube* video demonstration on the Learning Design page of the *VisChem* web site.[12]

The animation portrays the redox reaction between copper metal and aqueous silver nitrate summarized in Fig. 1. The animation is embedded in a learning design that starts with a demonstration of the reaction, seen either in the laboratory or on video (*YouTube* video $t = 1:40$), and the task of explaining the three observations ($t = 2:18$) – solution colour change to blue, growth of crystals on the copper, and the apparent 'dissolving' of some copper metal. Students are asked to describe what is going on at the molecular level in a storyboard ($t = 2:35$, and see a sample initial storyboard in Fig. 2a). In terms of the cognitive learning model[19] that is the framework informing this learning sequence, these activities are designed to "prime" the perception filter by activating the attention networks in the brain. Before showing the animation of the reaction it is essential to review the visual icons used to represent the chemical species in the reaction ($t = 3:55$), and refresh the memory of previous visualisations of substances (in this case, copper metal, $t = 4:31$) and solutions (in this case, aqueous silver nitrate, $t = 4:48$). The unmodified animation in the *VisChem* collection portraying the redox reaction was found to be far too cognitively demanding, even when shown in short segments. The modified version ($t = 5:22$) uses pre-emptive narration, highlighting, and repetition to focus attention on the key features of the animation. The next step in the learning sequence ($t = 10:18$) is to link the three thinking levels (see Fig. 1), explaining the observations in terms of the molecular-level events, and expressing these in symbolic language. The final step is for the novices to reflect on the whole process and transfer ideas in their new mental model of a redox reaction to another similar reaction with some differences. For example, novices should be able to analyse the corrosion of iron in air where the oxidising agent is a gas rather than dissolved ions.

The following sections will explain why the animation was modified and presented with this learning design, taking into account the recommendations and findings in the references cited above, and the need for students to engage with the essential complexity of the chemistry.

3.1 Inclusion of the solvent – extraneous cognitive load, or necessary for the chemistry?

Some chemistry education researchers have been critical of the inclusion of solvent in molecular-level visualisations, indicating that it contributes

unnecessarily to cognitive load. They recommend that the visualisations should be "simple", showing only selected processes (e.g., electron transfer) in a reaction, rather than showing a "realistic" depiction including the solvent.[20–22] A simplified animation portraying the reactions occurring at the electrodes in an electrochemical cell against a misty blue background, supposedly representing the aqueous solvent, is provided in the multimedia supplement for a leading chemistry text-book.[23] A matching figure exists in the corresponding chapter. Indeed, this relegation of the solvent to the status of 'molecular wallpaper' is very common in visualisations in educational publications and multimedia resources. However, this practice ignores the central role of solvation in many reactions in solution, and in this and other redox reactions in particular, and ignores the importance in chemical speciation in chemical reactions.

A thermodynamic analysis of the $Cu/Ag^+(aq)$ redox reaction shows that the reaction is not spontaneous without the solvent, and that the enthalpy of hydration of the copper(II) ions (-2099 kJ mol^{-1}) is a driving influence in the reaction. In other words, the water molecules play an *active* role in this reaction, and in many other reactions in solution, through enthalpy and entropy contributions. This suggests that educators should emphasise that ions exist as hydrated species in solution (Fig. 8), rather than as isolated species. One way to do this is to use the ionic formula representation (e.g., $Mg^{2+}(aq) + Cl^-(aq)$) instead of the potentially misleading 'molecular' formula representation (e.g., $MgCl_2(aq)$) for a dissolved ionic compound. *A major function of chemical education, and molecular modelling in particular, is to communicate chemistry "as simply as possible, but not simpler" (Albert Einstein).*

In a comparison study of a simplified animation with the *VisChem* animation, both portraying the $Cu/Ag^+(aq)$ redox reaction, and shown *without* narration and a legend to explain the meaning of the visual symbols involved, Rosenthal and Sanger[24] unsurprisingly identified student errors in explanations, misconceptions and misinterpretations from the animations. The importance of building familiarity and visual literacy using consistent conventions for how the molecular level is
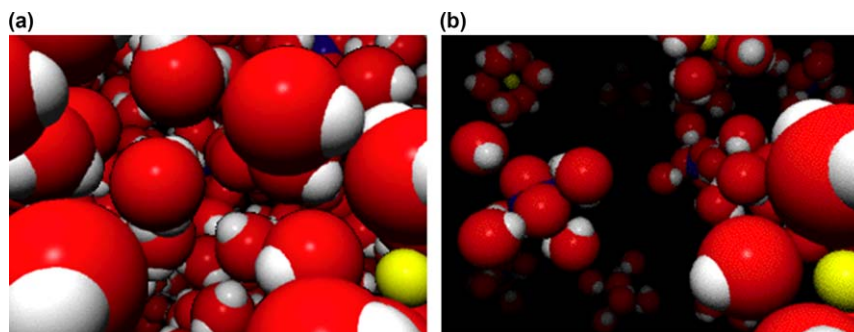
(a) (b)

Fig. 8 Some *VisChem* animations portraying ions in solution exist in one version with all solvent molecules (a), and in a simplified version (b) without most of the solvent water molecules (see Table 1h). However, the ions are still shown as hydrated species to emphasise the importance of their chemical speciation. (Animations available in the ESI.[†])

portrayed, and using narration and other attention focussing strategies to draw attention to the key features in a visualisation, are essential strategies for ameliorating the cognitive load issue, and making molecular modelling meaningful for novices.

3.2 Importance of drawing for interpretation and integration of complex visualisations

A number of studies have found that drawing is an effective way to help students interpret complex visualizations and integrate information from molecular modelling.[25–27] Generating drawings can help students develop criteria to distinguish various ideas demonstrated in a visualization and promote integrated understanding of chemical reactions. Storyboarding before and after showing visualisations is an important part of the *VisChem* learning design above, as it helps novices to explicitly express their ideas and new insights, and reveals otherwise hidden novice misconceptions and misinterpretations to educators (see Fig. 2a and b).

3.3 Minimising split attention effects

An important consideration in the production and presentation of molecular-level visualisations is the issue of minimising split-attention effects, in which attention to one type of presentation component may result in information being missed in a different, accompanying presentation component in the field of view.[28] This was an issue in the *VisChem* $Cu/Ag^+(aq)$ reaction animation (*YouTube* video $t = 7{:}33$) where reduction of the silver ions occurs at sites remote from, but synchronously with, sites of oxidation of copper atoms, in a ratio of two silver ions reduced for every copper atom oxidised (Fig. 9). In addition, the high perceptual salience[15] of the bright red spheres representing oxygen atoms in the water molecules, oscillating on and off the silver ions in the
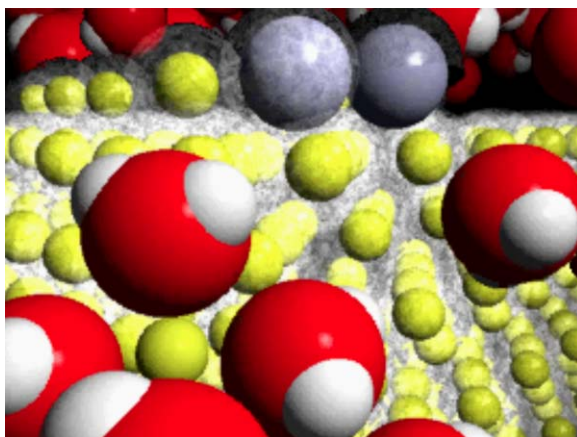

Fig. 9 A frame from the *VisChem* $Cu/Ag^+(aq)$ reaction animation (*YouTube* video $t = 7{:}33$) where there is a split in attention between the reduction and oxidation processes occurring, as well as a distracting high perceptual salience of the bright red oxygen atoms in the water molecules. (Animations available in the ESI.[†])

competition between ion hydration and electron transfer, was another source of distraction.

3.4 Avoiding apparent directed intent in animations

A common error in molecular-level animations, not found in calculated simulations, is the apparent 'directed intent' of chemical species moving toward each other to react, then leaving the reaction site afterwards. We discovered this flaw in the *VisChem* animation portraying precipitation of AgCl (Table 1k) during interviews with students. One perceptive student thoughtfully drew this to our attention: *"This animation … shows water molecules … sort of carrying this structure* [AgCl ion pair] *along … like a bunch of little robots … The animation depicts something that … I think really happens by chance, as a very deliberate sort of process and I think that's slightly misleading."* The reason that animation frames are not usually 'wasted' on depicting *unsuccessful* encounters (which are the majority) are related to the technical imperative to reduce rendering times, the animation time period, and to minimise file size to enable rapid delivery online. However, we need to explicitly point out to students that this is a form of 'artistic license', and can be likened to the conventional use of a chemical equation to summarize a reaction, rather than to list all the elementary steps as separate equations in the reaction mechanism.

4 Conclusion

Novices need useful mental models of the molecular world to explain observable phenomena, and to make sense of the chemistry symbolism and language used to describe them. Indeed, there is convincing evidence in the chemistry education literature that many novice difficulties and misconceptions result from inadequate or inaccurate models of the molecular level. Novices should be encouraged to communicate their models using storyboards (drawings with written elaboration) because conventional assessment often fails to expose these misconceptions.

Modelling structures and processes at the molecular level, using animations and simulations, requires a judicious compromise of scientific accuracy and artistic license for clear communication. Choreographed *VisChem* animations portray the competition between forces involved in proton and electron transfer reactions, precipitation, and dissolution—reactions that are difficult to simulate in reasonable time and with visual clarity. Calculated simulations in *Odyssey* depict multimolecular systems with realistic randomness in the system, instead of the often directed, apparently 'intentional' movement of species in animations. Simulations also give novices the ability to experiment by manipulating variables; filter selected molecules from view while still observing 'trails', hydrogen bonds and molecular dipoles; and construct their own simulations to discover new insights. A combination of animations and simulations is ideal for developing mental models of molecular level structures and processes.

The high cognitive load and split-attention effects in molecular visualizations depicting dynamic movement of brightly coloured icons constitute a pedagogical challenge for their use in education. Strategies such as pre-emptive narration, highlighting, and repetition, while providing student control over the frame rate and sequence order can ameliorate the cognitive load, as illustrated in the modified *VisChem* Cu/Ag$^+$(aq) redox animation. However, great care is necessary when attempting to simplify the visual model. For example, many animations and diagrams available in chemistry textbooks and online videos treat solvent molecules as 'visual noise', when in fact these molecules are active participants in many chemical reactions, and are critical to their thermodynamic spontaneity. The focus always needs to be on depicting the appropriate chemical speciation of the reactants and products in a reaction, simplifying where possible, but never oversimplifying.

References

1 R. N. Zare, *J. Chem. Educ.*, 2002, **79**, 1920.
2 P. Atkins, *Pure Appl. Chem.*, 1999, **71**, 927.
3 *IUPAC. Compendium of Chemical Terminology*, ed. A. D. McNaught and A. Wilkinson, Blackwell Scientific Publications, Oxford, 2nd edn, 2006.
4 A. H. Johnstone, *J. Comput. Assist. Learn.*, 1991, **7**, 75.
5 R. Tasker and R. Dalton, *Chem. Educ. Res. Pract.*, 2006, **7**, 141.
6 M. B. Nakhleh, *J. Chem. Educ.*, 1992, **69**, 191.
7 K. Taber, *Chemical Misconceptions: Prevention, Diagnosis and Cure*, Royal Society of Chemistry, Cambridge, 2002.
8 S. C. Nurrenbern and M. Pickering, *J. Chem. Educ.*, 1987, **64**, 509.
9 M. B. Nakhleh, *J. Chem. Educ.*, 1993, **70**, 52.
10 R. Tasker, *Educ. Chem.*, 2015, **May**, 20.
11 R. Tasker, *J. Chem. Educ.*, 2016, **93**, 1152.
12 R. Tasker, *VisChem Project*. Available from: http://www.vischem.com.au [Accessed: 1st July 2016].
13 P. Mahaffy, R. Bucat, R. Tasker and J. Kotz, *Chemistry: Human Activity, Chemical Reactivity*, Nelson Education, Canada, 2nd edn, 2015.
14 M. Hegarty, *Learn. Instr.*, 2004, **3**, 343.
15 R. Lowe, *Learn. Instr.*, 2003, **13**, 157.
16 R. E. Mayer and R. Moreno, *Educ. Psychol.*, 2003, **38**, 43.
17 B. Tversky, J. B. Morrison and M. Betrancourt, *Int. J. Human-Comput. Stud.*, 2002, **57**, 247.
18 S. Ainsworth, *Comput. Educ.*, 1999, **33**, 131.
19 R. Tasker and R. Dalton, in *Visualization Theory and Practice in Science Education*, ed. J. K. Gilbert and M. Reiner, 2008, Springer, p. 103.
20 J. P. Suits, in *Chemistry Education: Best Practices, Opportunities and Trends*, ed. J. García-Martínez and E. Serrano-Torregrosa, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2015, p. 595.
21 J. P. Suits and M. J. Sanger, in *Pedagogic Roles of Animations and Simulations in Chemistry Courses*, ed. J. P. Suits and M. J. Sanger, American Chemical Society, Washington, DC, 2013, p. 447.
22 J. P. Suits and M. J. Sanger, in *Pedagogic Roles of Animations and Simulations in Chemistry Courses*, ed. J. P. Suits and M. J. Sanger, American Chemical Society, Washington, DC, 2013, p. 1.

23   M. Silberberg and P. Amateis, *Chemistry: The Molecular Nature of Matter and Change*, McGraw-Hill Education, New York, 7th edn, 2015, p. 932.

24   D. P. Rosenthal and M. J. Sanger, *Chem. Educ. Res. Pract.*, 2012, **13**, 471.

25   Z. H. Zhang and M. C. Linn, *J. Res. Sci. Teach.*, 2011, **48**, 1177.

26   Z. H. Zhang and M. C. Linn, *Int. J. Sci. Educ.*, 2013, **35**, 2174.

27   S. Ainsworth, V. Prain and R. Tytler, *Science*, 2011, **333**, 1096.

28   R. E. Mayer and R. Moreno, *J. Educ. Psychiatry*, 1998, **90**, 312.