PERSPECTIVE

Check for updates

First-principles calculations of hybrid inorganic–organic interfaces: from state-of-the-art to best practice

Oliver T. Hofmann, [ID] *[a] Egbert Zojer, [ID] [a] Lukas Hörmann, [ID] [a] Andreas Jeindl [ID] [a] and Reinhard J. Maurer [ID] [b]

The computational characterization of inorganic–organic hybrid interfaces is arguably one of the technically most challenging applications of density functional theory. Due to the fundamentally different electronic properties of the inorganic and the organic components of a hybrid interface, the proper choice of the electronic structure method, of the algorithms to solve these methods, and of the parameters that enter these algorithms is highly non-trivial. In fact, computational choices that work well for one of the components often perform poorly for the other. As a consequence, default settings for one materials class are typically inadequate for the hybrid system, which makes calculations employing such settings inefficient and sometimes even prone to erroneous results. To address this issue, we discuss how to choose appropriate atomistic representations for the system under investigation, we highlight the role of the exchange–correlation functional and the van der Waals correction employed in the calculation and we provide tips and tricks how to efficiently converge the self-consistent field cycle and to obtain accurate geometries. We particularly focus on potentially unexpected pitfalls and the errors they incur. As a summary, we provide a list of best practice rules for interface simulations that should especially serve as a useful starting point for less experienced users and newcomers to the field.

1 Introduction

Interfaces between organic and inorganic materials are special for a variety of reasons. From a technological point of view, inorganic materials and organic molecules exhibit complementary properties: For instance, inorganic materials tend to have higher charge-carrier mobilities, while organic materials exhibit stronger light-matter coupling and are easier to modify chemically.[1–3] Organic molecules may, for instance, be tuned to obtain a specific band gap for light absorption or emission. Moreover, hybrid inorganic–organic interfaces determine the functionality of essentially all organic (opto)electronic devices,[4–6] like OLED-TVs or AMOLED displays. For these reasons, they have been in the focus of research on organic semiconductors already for several decades.

Hybrid inorganic–organic interfaces are also highly interesting from the perspective of fundamental science, as they form a bridge between the traditional worlds of chemistry and physics, combining delocalized and localized electronic states.

The strength and nature of the interactions between an inorganic and an organic system can vary substantially, ranging from purely physisorptive systems, where van der Waals interactions dominate, to strongly chemisorptive interactions, where new bonds are formed (and, inevitably, other bonds are weakened or broken).[7,8] Accurately describing these different worlds on equal footing poses a formidable technical challenge, especially if such a description is based on first-principles computational simulations.

This challenge can be loosely separated into two main aspects: On the one hand, there is the "level of sophistication" of the computational method that needs to be sufficient to describe the relevant physics of the interface, *i.e.* the "methodological challenge". Although both density functional theory and the solution of the Schrödinger equation using a many-electron wavefunction method are formally exact, in practice every practically applicable variant of these approaches is based on certain fundamental approximations that limit the physical effects it can capture. Such approximations comprise, *e.g.*, the choice of a specific approximate density functional in DFT approaches or using trial wave functions of limited flexibility in (post) Hartree–Fock calculations. A second, often overlooked aspect is the "numerical challenge": This arises from the many

[a] Institute of Solid State Physics, Graz University of Technology, NAWI Graz, Petersgasse 16/II, 8010 Graz, Austria. E-mail: o.hofmann@tugraz.at
[b] Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK

8132 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

possible numerical strategies and the specific settings for solving the system of equations resulting from the chosen electronic structure method. These options include the choice of the type of basis set, the algorithms used to find the minima for the electronic and/or geometric structure, the precision with which the minima are determined, *etc.* Here, many choices that perform well for crystalline inorganic materials (such as metals) fare much worse for molecules and *vice versa*. Most electronic structure software packages employ robust default settings that ensure convergence for a variety of systems. These settings are sometimes more geared towards crystalline inorganic materials or towards molecular systems, making the code more efficient for one material class than for the other. However, the ideal choices of parameters and algorithms become more difficult when different classes of materials are combined, which is naturally the case for hybrid inorganic–organic interfaces. In the best case, this may mean that the defaults cause the calculations to run inefficiently, so that they take longer than necessary to complete. In the worst case, applying the wrong algorithm – or the right algorithm wrongly – can render calculations numerically unstable or can lead to results that are physically incorrect.

The goal of this article is to provide a detailed account of the "methodological" as well as the "numerical/technical" challenges associated with modelling hybrid inorganic–organic interfaces. To achieve this, we will first discuss (in Section 2) in which respects the inorganic and organic components of a hybrid interface differ from a physical point of view. This provides the basis for the remainder of the discussion: in Section 3, we will briefly address model building, *i.e.* the necessary considerations when creating a model of a system that will be simulated and eventually compared to a real-world experiment. Section 4 continues by briefly discussing the advantages, disadvantages, and limitations of different popular computational methods that are used to model hybrid inorganic–organic interfaces. Due to its practical relevance, we will mainly focus on density functional theory with the different exchange–correlation approximations and van der Waals correction schemes that can be employed to model hybrid inorganic–organic interfaces. In Section 5, we will then deal with the algorithms, numerical approximations and settings employed in common software packages. Using examples, we will show, *inter alia*, how the uncritical use of numerical default settings may yield incorrect results. In addition, we will highlight the problems associated with geometry optimizations and the different optimization strategies, which can lead to significant differences in the calculated interface geometries.

During the following discussion, we will especially address questions relevant to practitioners. These include aspects regarding the general strategy employed for modelling interfaces (*e.g.*, extended interfaces with periodic boundary conditions *versus* aperiodic interface models), as well as specific details such as the optimal choice of numerical parameters, *e.g.*, for a rapid convergence of the self-consistent field algorithm or for the geometry optimization. We provide a diverse, but not exhaustive, set of examples from our experience gathered over more than a

decade of interface simulations, but also show new simulations performed to illustrate certain arguments. This is done in an effort to establish a set of best practices. Wherever possible, we point out how chemical intuition or physical understanding can guide the choice of the optimal algorithms and their settings. We expect that this work will help newcomers in this field to obtain reliable results more efficiently, but we are confident that even seasoned veterans may still learn a thing or two. We definitely did when compiling this work!

# 2 Some fundamental differences between inorganic and organic materials

An inorganic–organic interface consists of two very dissimilar components: An inorganic component ("the substrate") and an organic component ("the adsorbate").† The substrate is either a metal, a semiconductor, or an insulator, and typically much thicker than the adsorbate, which consists of molecules adsorbed in a (sub)monolayer or as a thin film. In this section, we will briefly discuss the most salient differences in their electronic properties. These differences will accompany us through the rest of this work, as they are the reason for many of the numerical challenges the simulations face. Before we start with the discussion, it is important to note that the possible components of hybrid inorganic–organic interfaces encompass – quite literally – the whole of chemical space. To keep the discussion focused and avoid derailing the topic, we have to simplify and generalize several statements regarding the different materials classes, fully knowing that there are several exceptions and further physical differences that we cannot mention here. Overall, we highlight three important differences in electronic structure between organic and inorganic compounds:

2.1 Electronic states – orbitals and bands (Fig. 1a)

Electronic states in molecules are typically discussed in terms of molecular orbitals, *i.e.*, as discrete energy levels. Conversely, electronic states in crystalline inorganic materials are commonly described as continuous bands. On a formal basis, the two pictures are strictly equivalent: A band structure is just a sorting of orbitals (which, in a periodic structure, are infinitely extended) using the associated wave vectors (which can be obtained, *e.g.*, by a Fourier analysis). The band structure picture can be used to reveal a major difference between two materials classes: Metals and semiconductors (and, more rarely, also insulators) show appreciable band dispersion, *i.e.*, a pronounced dependence of the energy of the electronic states on the wavevector. The reason for that is that the individual building blocks (which are often just one or a few atoms) interact with each other over multiple unit cells. Conversely, in organic materials the wave-function overlap between neighboring

---

† Note that, of course, the role of substrate and adsorbate may be reversed, *i.e.*, the substrate could be organic and the adsorbate inorganic.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8133**
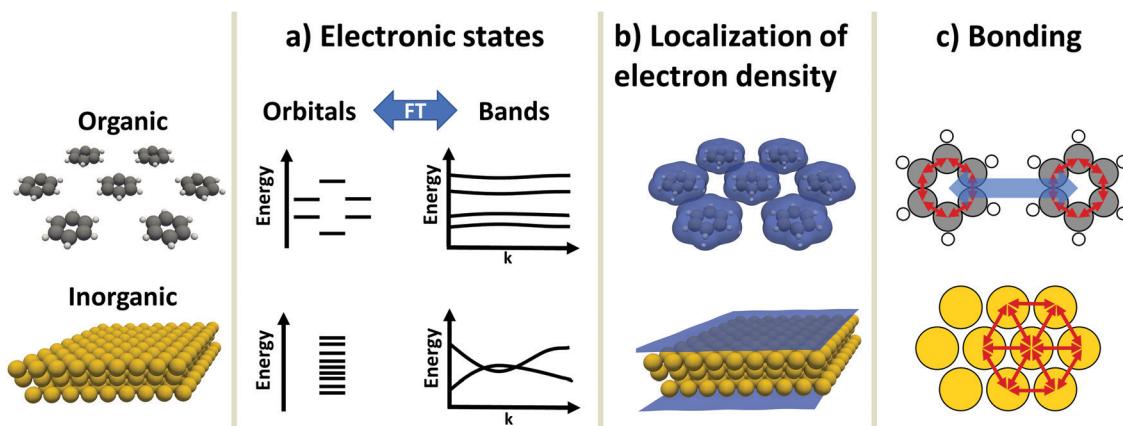
Fig. 1 Differences between organic and inorganic materials. Shown is a monolayer of benzene and a metal slab (here, Au) as prototypical organic and inorganic component, respectively. (a) Schematic depiction of the electronic states in the orbital picture and the band picture. Due to the interaction between the atoms, the metal exhibits many energetically similar states that form (dispersing) bands due to the interaction between the atoms. Conversely, each benzene molecule has several molecular orbitals, which hardly interact with each other, giving rise to (almost) flat bands. (b) Electron density, as calculated with the PBE functional, shown at an isovalue of 0.0025 e bohr^{-3}. While the metal shows an extended, rather smooth and homogeneous electron distribution, the electron density of the organic component is localized on the individual molecules and becomes small in between. (c) Schematic depiction of the bonding in the components. Each benzene molecule is held together by directed covalent bonds, while between the molecules, non-directed van der Waals interactions dominate the cohesion. The metal is held together by metallic bonds. Like covalent bonds, these originate from wave function overlap, but they contain multiple bonding partners at the same time (which also increases their isotropy).

molecules is typically much smaller, leading to bands with comparably smaller band dispersions. In fact, even in high-mobility materials like rubrene, band dispersions in π-stacking direction amount to only a few tenths of an eV.[9]

2.2 Variation of the valence electron density (Fig. 1b)

Crystals consisting of elemental metals and simple semiconductors typically have a relatively uniform, weakly varying valence electron density. Conceptually, their electron densities can, thus, be reasonably well described using models derived from the homogeneous electron gas, or weak perturbations to it (like the local density approximation or the generalized gradient approximations within density functional theory). This contrasts with organic molecules, where the charge densities in the regions between the individual molecules can drop by orders of magnitude, resulting in much larger gradients of the electron density. Hence, for organic materials hybrid functionals tend to yield more accurate properties than (semi-)local functionals. This situation is primarily a consequence of interatomic distances between atoms of different molecules being significantly larger than interatomic distances of neighboring atoms within the same molecule. Additionally, the periodic repeat units of the inorganic crystals relevant in the present context are typically small, encompassing only one or a few atoms, resulting in lattice vectors of only a few Å in length. Due to the size of molecules, the unit cells of molecular crystals or adsorbates are typically much larger than those of a metal or semiconductor substrate such that the above-described charge density variations occur on length-scales much larger than the size of the unit cell of the substrate.

A further reason for qualitatively different charge distributions in organic and inorganic materials is that metals and many (conventional) semiconductors exhibit only non-polar (or very

weakly polar) bonds. In contrast, organic molecules can contain highly polar bonds (such as O–H, N–H, C=O, etc.).

2.3 Chemical bonding (Fig. 1c)

Metals and most insulators have high coordination numbers, since they are held together by isotropic forces (i.e., metallic and ionic bonds). Semiconductors and organic molecules are primarily held together by covalent forces, which are strongly anisotropic. Conversely, van der Waals forces, which for most organic molecules are both the major source of cohesion in molecular crystals and a major factor in the interaction between substrate and adsorbate,[7] are (to a first approximation) isotropic.‡

These three major differences between organic and inorganic systems each require special consideration when performing first-principles calculations: For instance, the localization of states and the variation of the electron density influences the applicability of electronic structure methods and basis sets used to describe the components (see Section 3). This poses a sizable challenge when combining differently bonded systems at a hybrid interface. Moreover, the different nature and directionality of the above-mentioned forces complicates the structure optimization, as will be illustrated in Section 4. A further challenge in modelling periodic hybrid inorganic–organic interfaces is the requirement of describing organic and inorganic components in the same unit cell. This requirement of commensurability often leads to very large super-cells, where a large number of substrate atoms need to be accounted for.

Finally, it should be emphasized that the nature of the interaction between inorganic substrates and organic

‡ Note that, technically, van der Waals forces are anisotropic as a consequence of the anisotropy of the polarizability of the system. However, this anisotropy is only weak and mostly ignored in modelling approaches.

8134 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

adsorbates strongly depends on the considered interface.[4–6,8,10] The interaction can perturb electronic states only weakly, being almost exclusively driven by van der Waals interactions, or it can be very strong and include the formation of covalent bonds.[7,11] Additionally, charge is often transferred across hybrid inorganic–organic interfaces.[4–6,8,10] At interfaces, the presence of dipoles can also play a different role than for bulk materials or isolated molecules,[10] impacting charge- and energy-transfer across the interface,[4,12] modifying the lateral charge-distribution at the interface,[13] and reshaping molecular states.[14] Regular assemblies of dipoles also massively impact the electronic structure of interfaces through so-called "collective electrostatic effects".[4,8,10,15]

3 The structural model

The first step of any first-principles calculation is to set up a structural representation of the system to be modelled. Realistic systems extend over mesoscopic length-scales, typically containing too many atoms to be explicitly considered in a quantum-mechanical simulation. Furthermore, only in very few cases the atomistic structure is known *a priori* in full detail. Therefore, an (idealized) abstraction of the real-world situation must be used. The choices made in this abstraction are crucial, as they will determine the outcome of the computation, potentially resulting in fundamental deviations from the actual situation. Even if a perfect simulation method was available, it could not yield sensible results if the structural representation oversimplifies reality. It is, therefore, imperative that the structural model contains all the physical and chemical aspects that are relevant for the considered real-world system. Thus, in this section, we will review several common approaches used to set up models for hybrid inorganic–organic interfaces and discuss their impact on the outcome of the calculations.

Interface calculations can be separated into two major categories, depending on the spatial extent of the chosen model system. In an open boundary or "cluster" calculation, illustrated in Fig. 2a, a finite-size cluster is cut from the actual, extended interface. It is then modelled as a finite system, either on its own or "embedded" into a surrounding medium that is then typically treated at a lower level of theory. Alternatively, a slab type

calculation can be performed. It employs periodic boundary conditions to infinitely repeat a central unit cell, as illustrated in Fig. 2b. Both open and periodic boundary calculations have their merits. Which approach is preferable depends on the system and the scientific question that should be addressed. At the same time, both approaches also face their own (notably different) challenges, as will be discussed in the following.

3.1 Cluster models of hybrid inorganic–organic interfaces

The vast majority of studies on extended hybrid inorganic–organic interfaces use periodic boundary conditions. However, finite cluster models of interfaces are well-suited for interface problems where the process or property to model is spatially localized, exhibits quantum confinement, or only affects a single molecule.[16] These could be, for example, interfaces formed by nanoconfined systems such as nanoclusters or small nanoparticles. For extended interfaces, cluster models can be preferable when, *e.g.*, the modelled coverage is extremely low or when there is only an individual molecule on a surface, which is often the case in catalysis-related problems. A similar situation is encountered in single molecule junctions, *e.g.*, in break-junction experiments, where one typically aims at a situation in which only a single molecule forms the bridge between two electrodes.[17] Modelling a system as "non-periodic" is also the method of choice when dealing with situations where the objects of interest are at such large distances that they do not significantly influence each other. This, for example, applies to adsorbates at moderate coverage in the absence of island formation, as well as to surface defects (again, for relatively low coverage). For such systems the unit cells in periodic boundary conditions would have to be made enormously big to prevent effects associated with the limited unit cell size, which introduce spurious interactions (especially if the defects or adsorbate molecules are charged).[18–21] In cluster calculations such spurious interactions cannot exist by design.

Nowadays it is possible to deal with clusters containing hundreds[22,23] or even thousands[24,25] of atoms. However, sometimes there are benefits to creating relatively small cluster models containing "only" a few tens of atoms. This allows to study interface properties with high-level methods, including highly correlated wave-function-based methods or many-body perturbation theory.[26,27] For certain spectroscopic observables, this arguably yields more accurate properties than when relying solely on DFT.[28] A further decisive advantage is that the calculated systems do not need to be charge-neutral (in contrast to periodic boundary calculations, as will be discussed below). This is particularly useful when considering, *e.g.*, charged defects and core-level excitations, or when calculating ionization energies/electron affinities from the energy difference between a charged and an uncharged cluster (ΔSCF-approach).[29,30]

There are two main challenges associated with creating a cluster representation: One has to decide which part of the substrate is cut out, *i.e.* the size and the shape of the cluster. Additionally, one has to decide how the fringe of the cluster is treated, *i.e.* whether (or how) it needs to be saturated or
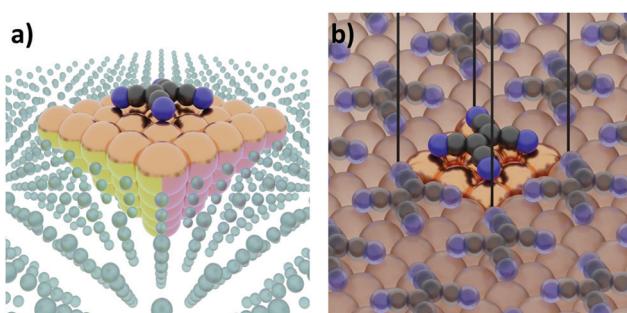
Fig. 2 Schematic representation of tetracyanoethylene (TCNE) on a fcc(100) surface in (a) a cluster calculation and (b) with periodic boundary conditions.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8135**

embedded into a realistic environment in order to avoid artefacts. A further obvious problem when using small substrate clusters are so-called finite-size effects, where the spatial confinement of the wavefunction in the cluster increases the band gap.[29,31] For semiconductors, this leads to quantitatively too large band gaps. The situation is, however, more problematic for metal clusters, which form a finite band gap (and are, thus, not metallic at all).[29,32] Particularly problematic in this context is that electronic and optical properties of metal clusters converge very slowly with cluster size.[33] For some properties, this problem can, in principle, be mitigated by using very large clusters or using appropriate extrapolation techniques.[34] However, such large clusters are often computationally unfeasible. A second, related problem is that the interaction between substrate and adsorbate can change non-trivially with the size and shape of the cluster.[35,36] Interestingly, this problem appears to be less serious when dealing with large, π-conjugated, flat lying molecules.[29] Another issue is that clusters cut out from the bulk structure (say, an fcc crystal structure of a coinage metal) are often not stable in this form.[37] This can outright prevent a geometry optimization of the substrate, since the cluster would disintegrate. In practice, this is sometimes solved by constraining the positions of the substrate atoms,[37] which, however, impedes a comprehensive determination of adsorbate geometries, since surface relaxations can be accounted for only to a limited degree.

Furthermore, cutting a cluster from a substrate creates additional surfaces at its rim. For semiconducting substrates, this requires "breaking" several covalent bonds. The remaining "dangling bonds" often give rise to states in the band gap, which artificially increase the reactivity of the cluster. To obtain reliable results, it is, thus, necessary to remove these dangling bonds. This can be done either by artificially saturating them, e.g. with (pseudo)hydrogen atoms,[38–44] or by employing density embedding methods.[16,31,45–50] In the latter case, the cluster, which is computed using an expensive method, is embedded into a surrounding material that is significantly more extended, but described by a computationally less demanding approach (often, specially designed pseudopotentials). Several density embedding schemes have been developed for metal clusters,[50,51] which are mostly focused on predicting adsorption energies and excited-state properties.

As mentioned above, it is important to be aware that most cluster models focus on only a single (or at most a few) individual molecules adsorbed on the surface. They are, therefore, generally not well-suited for mapping a situation in which inter-molecular interactions are relevant or even dominant. This is, for example, the case when the adsorbate structure is a direct consequence of adsorption-induced surface reconstructions[52–58] or intermolecular interactions (as for upright-standing molecules in ordered self-assembled monolayers, SAMs).[59–62] Such SAMs typically consist of long aliphatic or aromatic backbones, which are chemically bonded to the substrate via docking groups. In these systems, the chemical docking of a single molecule to the substrate might be properly described by a cluster model, but to obtain reliable electronic properties of the interface, also suitable embedding schemes

need to be applied.[63–65] At the same time, these systems also have sizable inter-molecular interactions between molecules, which are mostly determined by electrostatic and van der Waals forces.[66] Removing the surrounding molecules in a cluster model also removes these interactions. As a result, rather than remaining upright, as they would be in a real-world sample, in a simulation the isolated molecules would fall over, which fundamentally changes the physical properties of the interface (see also Section 5). Even for flat-lying molecules, neglecting the interactions with neighboring molecules in a cluster calculation can be particularly problematic, if their electronic nature changes due to intermolecular interactions. This is, for example, the case for dyes such as indigo, where the formation of hydrogen bonds triggers an extension of the aromatic π-system.[67] Notably, even if the size of the cluster is extended to accommodate several interacting molecules, edge effects would likely be unavoidable.

Besides the two rather specific examples of self-assembled monolayers and H-bonded dyes, cluster calculations, by design, miss so-called collective (also termed cooperative) electrostatic effects.[68–74] This term refers to the fact that extended sheets of dipoles, which are commonly encountered at interfaces, affect the electrostatic potential landscape in a way that is fundamentally different from the impact of an isolated dipole.[10,75] This directly affects the energies of the electronic levels of the materials and has a strong impact on several observables, such as X-ray photoemission spectra (i.e., core-level energies)[76] or ballistic electron transport through molecules,[77] as well as on the overall magnitude of the charge transfer.[78] In principle, these collective effects could be modelled in a cluster approach by using a sufficiently large cluster containing multiple molecules, but whenever an extended sheet of dipolar entities is replaced by a finite-size cluster, massive electrostatic edge-effects occur.[70,79] Thus, to date the only strategy for correctly capturing collective electrostatics in conjunction with a cluster approach appears to be the use of suitable embedding schemes.[65]

3.2 An overview of the repeated slab approach

Most calculations for hybrid inorganic–organic interfaces are, arguably, performed using periodic boundary conditions as indicated in Fig. 2b. Therefore, the peculiarities of this approach shall be discussed in more detail here. There are two subclasses of this ansatz, "buried interfaces" and "exposed interfaces". As shown in Fig. 3a, in a "buried" interface, the whole unit cell is filled with inorganic and organic material. Due to the periodic boundary conditions in the direction perpendicular to the buried interface, there are two separate interfaces between the organic and inorganic material, which may not be identical.

Buried interfaces are the model of choice for the simulation of hybrid inorganic–organic interfaces with thick adsorbate films. This includes, for example, multilayer adsorption of small molecules or truly three-dimensional heterostructures.[80] Conversely, "exposed" interfaces are the model of choice when the simulations are connected to surface-science experiments of

8136 | Phys. Chem. Chem. Phys., 2021, 23, 8132–8180

This journal is © the Owner Societies 2021
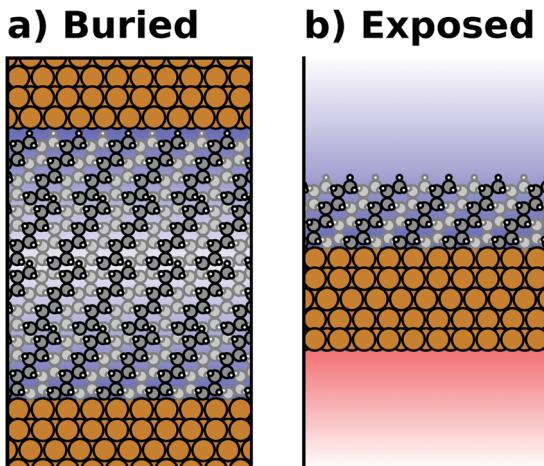
a) Buried   b) Exposed

Fig. 3 Schematic representations of (a) a buried interface and (b) an exposed interface, consisting of alkyl chains on a metal slab. The background gradients indicate the electrostatic potential.

solid/vacuum and solid/gas interfaces. While most issues that we will discuss below pertain to buried as well as exposed interfaces, the focus in the following will be on exposed interfaces due to their prevalence in literature.

Exposed interface calculations require the inclusion of a region of vacuum, as shown in Fig. 3b. The purpose of this vacuum region is to decouple the periodic replicas perpendicular to the interface, which are generated as a consequence of the employed boundary conditions (see below). The semi-infinite substrate is modelled by a finite number of atomic layers and the exposed interface(s) frequently contain only a (sub)monolayer of the adsorbate, usually only on one side of the metal slab. In the past, adsorbate layers have sometimes been placed on both sides of the substrate to prevent spurious polarizations. This approach is now outdated, as alternative approaches to account for an asymmetry of the slab have been implemented in the majority of the commonly applied codes, as discussed below.

Most band structure codes employ periodic boundary conditions in all three dimensions. As the buried and exposed interfaces are intrinsically only periodic in two dimensions (unless the buried interface is meant to model a periodic organic–inorganic heterostructure), in the simulations there is also the above-mentioned periodic repetition of the substrate(s) and the adsorbate(s) in the direction perpendicular to the interface. In conjunction with exposed interfaces, this strategy is commonly referred to as the "repeated slab approach".

Despite their name, not only cluster, but also periodic calculations may suffer from various finite size effects:

• When the cell contains too little vacuum, the periodic replicas in z-direction may interact with each other. Provided that the vacuum is large enough to prevent wave-function overlap (approx. 15–20 Å), this interaction is not of quantum-mechanical nature. Rather, electrostatic interactions and polarization effects play a role, as will be discussed in detail in the next section.

• Unit cells can be laterally too small to represent the actual experimental situation. This can be the case, for example, when a low-coverage situation should be modelled, but the size of the unit cell in the computation needs to be restricted due to computational limitations. Too tight packing of adsorbates then results in spurious interactions between lateral replicas, which can lead, e.g., to geometrical distortions of the adsorbate molecules or to inaccurate adsorption energies, especially when encountering spurious electrostatic interactions between charged adsorbates.

• Finite-size effects can also occur when the number of atomic layers chosen to represent a semi-infinite substrate is too small. As in the above-discussed case of cluster calculations, this can result in an artificial band gap and/or artificial surface states of the substrate. In contrast to clusters, where no clear construction recipe exists to eliminate artifacts due to such finite-size effects, for slab calculations the thickness of the substrate can be systematically increased by adding further layers (using the bulk crystal structure of the substrate) until these finite size effects disappear.

An advantage of periodic boundary conditions is that substrate slabs are usually stable structures. It is, thus, both possible and sensible to relax the geometry of the substrate when it comes into contact with the adsorbate. Often, these geometry optimizations lead to partial extractions of surface atoms or small-scale surface reconstructions.[81,82] This can notably improve the stability of the interface, leading to a substantially larger (more exothermic) adsorption energy. It is common practice not to relax the complete slab. Rather, only the layers closest to the adsorbate are allowed to rearrange, while the other atoms are constrained to their bulk position. The reasoning behind that strategy is that the layers away from the interface are meant to represent the bulk of the substrate. In fact, if all metal layers were allowed to relax, the atoms at the surface of the slab not covered by an adsorbate would adopt the geometry of a clean metal surface, rather than that of the bulk atoms. In this way one would model adsorption on a free-standing metal film rather than on an exposed surface. As a rule of thumb, half or more of the layers present in the slab should be kept constrained. In Section 5 we will showcase, how such seemingly unimportant "details" (like fixing the positions of all slab atoms completely) can critically influence the resulting adsorption geometry.

3.3 A detailed discussion of the repeated slab approach

There are additional consequences of employing periodic boundary conditions, as they imply also full translational periodicity of all observables (including the electrostatic potential, the electron density, and, of course, the ionic positions). This imposes some constraints and limitations on the simulations, which are discussed in the following paragraphs.

3.3.1 Determining interface dipoles. Periodic boundary conditions require that the electrostatic potential must have the periodicity defined by the unit cell. Therefore, the potential must not change for translations by (multiples of) a unit cell vector. In particular, the electrostatic potential must be identical at opposite edges of a unit cell. In the example shown
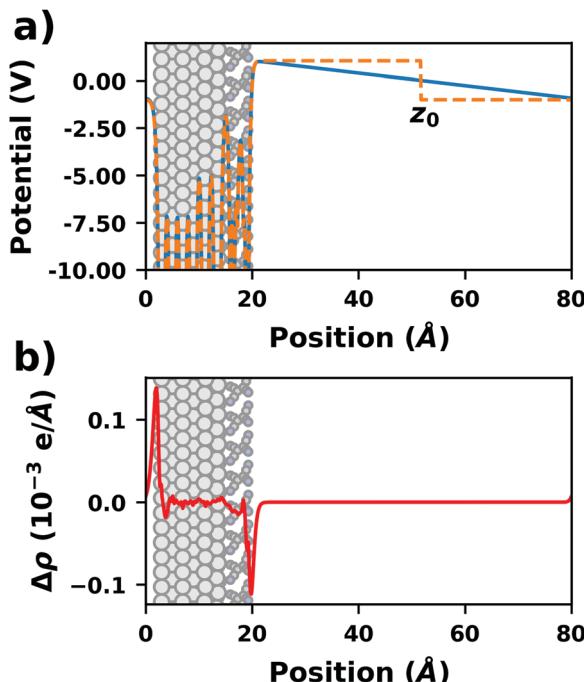
Fig. 4 Tetracyanoethylene on Ag(111) using a unit cell with 80 Å in z-direction. (a) Evolution of the (plane-averaged) electrostatic potential without (blue) and with (orange, dashed) dipole correction. (b) Polarization induced by the spurious field, calculated as the difference of the electron density with and without dipole correction. The Ag/TCNE system extends from 0 to ~20 Å, with 60 Å of vacuum above it. The data were obtained with FHI-aims using the PBE+vdW^{surf} method and a $3 \times 3 \times 1$ Γ-centered k-grid. The original calculations can be found at http://dx.doi.org/10.17172/NOMAD/2019.10.16-1.

in Fig. 4a, which displays the electrostatic potential of a tetracyanoethylene (TCNE) monolayer adsorbed on Ag(111), this requires identical values of the electrostatic potential at 0 and 80 Å (the two limits of the unit cell in the direction perpendicular to the surface). Thus, there can be no net dipole moment within a unit cell. In practice, when calculating a periodic structure that would have a dipole moment, *e.g.*, because of charge rearrangements at the interface or due to polar adsorbates, an electric field appears that generates a potential gradient in the unit cell such that the potential shift induced by the system's dipole moment is exactly cancelled. This is shown by the blue curve in Fig. 4a. The magnitude of the artificial electric field depends linearly on the magnitude of the dipole density (*i.e.*, dipoles per surface area of unit cell) and inversely on the thickness of the vacuum region. A small vacuum region therefore leads to a large field, a large vacuum to a small field. Note that this field is not a feature specifically implemented into certain program packages. Rather, it occurs automatically because of the periodic boundary conditions. The spurious field is most clearly visible in the vacuum region separating two periodic replicas of the slab. Intimately linked to that field is an artificial polarization of the slab. This artificial polarization is visualized in Fig. 4b, which shows the plane-integrated difference between the electron densities of the TCNE/Ag(111) interface with and without the spurious field.

Interestingly, although the metal is much more polarizable than the TCNE molecule, the polarization leads to notable (if relatively small) charge rearrangements between the ends of the entire slab, *i.e.* from the top of the TCNE monolayer to the bottom of the metal slab.

The spurious field also affects the total energy of the system. A first estimate of that contribution (neglecting the polarization of the metal and the adsorbate) can be derived from the interaction between the dipole of the interface and a homogenous field.[83,84] As pointed out by Bengtsson[84] (because the field is internal, not external), it is given by $E = 1/2 \times \mu \times F$, where μ is the dipole and F the field within the unit cell. In most cases, this energy contribution is small but non-negligible and, thus, might modify the calculated relative stabilities of specific configurations. The magnitude of the effect shall be illustrated by the following, example: Let us consider a dipole of 1 e Å (approx. 5 Debye), which is packed at such a density (here, *ca.* 180 Å^2) that the resulting step in the electrostatic potential amounts to 1 V. Using a unit cell height of 40 Å, the total energy contribution from the dipole in the field is $1/2 \times 1$ e Å $\times$ 1 V/40 Å = 0.01 eV. Although that value is small, it is in the same range as the typical energy differences between different polymorphs of organic molecules.[85] Therefore, it is important to never compare energies between different structures when the spurious field is not compensated. This is, for example, sometimes done for time-consuming geometry optimizations, since the correction schemes discussed below can slow down the SCF convergence.

The spurious field and its ensuing artefacts can be avoided in several ways, which are graphically depicted in Fig. 5:

• Size of the vacuum region: in principle, one possible approach would be to increase the vacuum region until the field essentially vanishes. Unfortunately, increasing the vacuum size also increases the computational cost, especially when working with a plane-wave basis set (for a discussion on basis sets, see Section 4). Moreover, since the field decays with one over thickness, the convergence is very slow, making this not a good approach. Moreover, in this approach a surface vacuum level never forms, making an accurate determination of properties like the work function impossible. Thus, this strategy is not advisable at all.

• Symmetric slab: a second approach is to construct symmetric model systems. There, the interface is mirrored perpendicular to the xy plane (*i.e.*, the adsorbate is added on the bottom side of the substrate, too), causing the net potential step across the overall slab to vanish for symmetry reasons. This approach has two major disadvantages. To avoid the two sides interacting with each other, both the substrate and the adsorbate must be duplicated. Therefore, the system must be at least twice as large as for an interface with an adsorbate layer bonded to only one side of the slab. Furthermore, care must be taken to construct truly symmetric systems. This is not always straightforward. For example, the (111) surfaces of fcc-metals are stacked in an ABC manner, making it impossible to generate strictly symmetric slabs. In order to at least have identical surfaces (terminating, *e.g.*, with "A"-type layers), it is necessary

8138 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021
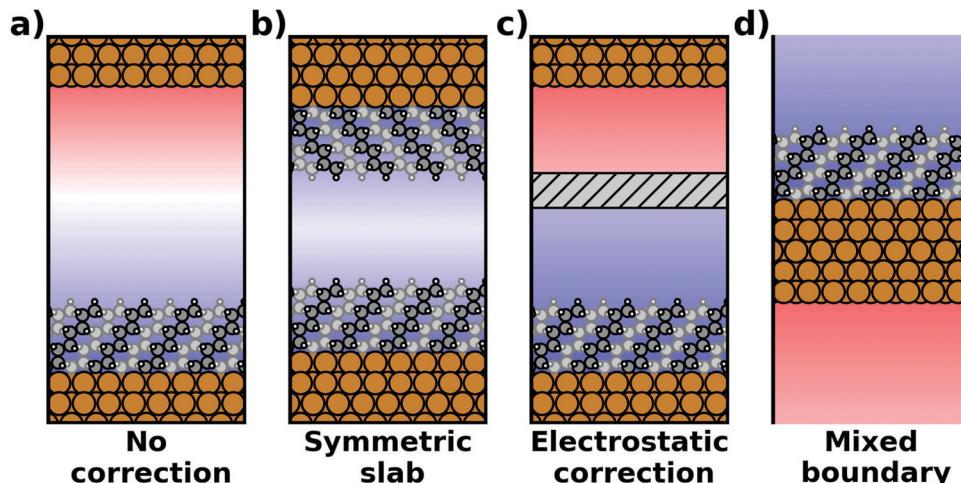
Fig. 5 Schematic visualization of the electrostatic situation of exposed interfaces: (a) no electrostatic correction (b) a symmetric construction to prevent the formation of a net dipole in the unit cell, (c) an electrostatic correction scheme (*i.e.*, either the dipole correction or a Coulomb cutoff) and (d) a truly 2D-periodic interface with mixed boundary conditions (periodic parallel to the interface, open boundary conditions perpendicular to it). The color in the background schematically indicates the evolution of the electrostatic potential.

to use $3n + 1$ layers (n being an integer number), *i.e.*, four layers (ABCA), seven layers (ABCABCA), ten, *etc.* For other types of materials, like, for example the polar surfaces of wurtzite-type materials including the popular ZnO, (close to) symmetric slabs cannot be constructed at all.

● Dipole correction: the most common way to deal with the spurious polarization due to periodic boundary conditions is to introduce a discontinuity in the electrostatic potential within the vacuum region.[83,84] This is also frequently referred to as the "dipole correction". Since it is by far the most common approach used in practice, we discuss it in more detail in the following: the first step is to define the position of the plane (parallel to the interface) where the discontinuity of the electrostatic potential is created. In Fig. 4, we designate this position as z_0. For all practical intents and purposes, this position separates the unit cell into a region that is "below" and a region that is "above" the slab (the red and blue regions in Fig. 5c). The choice of the position is, to some extent, arbitrary, but it should be far away from the electron density of the slab to ensure that lateral variations in the potential due to the finite extent of the adsorbate molecules have decayed. While z_0 can be chosen manually, today many codes determine this position automatically, choosing either the center of the vacuum or the position in the unit cell where the electron density is minimal (for numerical reasons, plane-wave basis sets exhibit very small but finite electron density in the vacuum region). The next step is to determine the magnitude of the potential discontinuity that is necessary to compensate for the asymmetry of the slab. The specific implementation differs between different code packages, but there are two main strategies how this can be done: Either, (1) the dipole moment of the unit cell in z direction is explicitly calculated by integrating over the plane-averaged charge density (electrons and nuclei) between z_0 and $z_0 + a_3$ (where a_3 is the length of the lattice vector perpendicular to the surface). Alternatively, (2) rather than the dipole

moment, the gradient of the electrostatic potential at z_0 (*i.e.*, the electrostatic field at this position) is numerically calculated, and a field with opposite polarity is applied by introducing the appropriate step in the potential. Either method needs to be repeated self-consistently throughout the SCF cycle until the system's dipole is fully compensated by the correction. Although this procedure generally increases the number of cycles required to reach SCF convergence, the additional computational cost tends to be relatively small, especially compared to the other strategies described before.

When applying the dipole correction in practice, some technicalities should be considered. First, the approach requires that z_0 is in a region where there is no electron density. Otherwise, especially for unit cells with a large vacuum region, moving z_0 by a small distance can result in significantly different dipole moments when using method (1). Similarly, also method (2) will fail, since the potential gradient differs for different values of z_0 if there is non-vanishing electron density in the vacuum region. The second technicality to keep in mind is that both procedures only work if the potential at z_0 is spatially homogenous, *i.e.* if there are no significant variations in the potential within the xy-plane. To achieve this, z_0 needs to be placed sufficiently far away from the interface. Due to collective effects, the field emanating from the dipoles decays very quickly.[68] Therefore, choosing z_0 at a distance from any atom in the slab that is larger than the distance between the dipoles on the surface can already be sufficient. Consequently, to be able to find a suitable position for z_0 in cases where there is only one dipole (*i.e.*, typically one adsorbate molecule) per unit cell, the height of the vacuum region should be in the same range as (and ideally larger than) the lateral extent of the unit cell. For large supercells, this provides the limiting factor for how small the vacuum region can be chosen in a slab calculation.

● Other approaches: there are many more approaches to achieve a similar electrostatic decoupling of periodic replicas:

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8139**

for example, the long-range electrostatic interactions across images perpendicular to the surface can simply be truncated in Fourier space. This so-called Coulomb cutoff method is fast, simple to implement[86] and effectively equivalent to the dipole correction approach.[87] Other approaches include the minimum image convention, where the Coulomb operator is still periodic, but contributions from neighboring unit cells are removed.[88,89] Yet another approach to eliminate spurious electrostatic interactions across images is the use of open boundary conditions perpendicular to the surface. Then, the system is periodically repeated in x and y directions, while it is non-periodic perpendicular to the surface (mixed boundary conditions). This can be achieved by solving the electrostatic potential on a grid and imposing reflecting (Dirichlet) conditions when solving the Poisson equation, which is used to evaluate the Hartree potential.[90,91] All these methods display a slightly different behavior[92] and are occasionally used for more non-standard situations that involve complex boundary conditions or electrostatic environments.[93]

Periodic boundary conditions directly imply that a sample is infinitely extended. While this is obviously not the case perpendicular to the surface, it is also an approximation parallel to the surface, since every real-world sample is finite in size. One consequence of this is that, if an adsorbate has a dipole moment parallel to the surface, it vanishes when employing periodic boundary conditions. For metallic substrates, this is less of an issue, since any dipole moment is fully screened by the metal. On non-metallic substrates, the screening is incomplete and in the actual sample lateral potential gradients prevail, which in the simulations would be eliminated by the periodic boundary conditions. However, because in practice the potential gradients are typically very small (and correlate inversely with the domain size),[94] this artefact is typically of little relevance (unless observables like piezoelectricity are of interest).

This brings up another problem associated with employing the repeated slab approach: naturally, this approach is not suited to describe interfaces that do not display periodicity in a two-dimensional plane. One of the reasons for that can be the lack of commensurability between substrate and adsorbate, an aspect that will be discussed below. Another reason is that one might not be dealing with a flat substrate or with a homogeneous adsorbate layer, as would be the case for nanoparticles with adsorbed molecules or for single molecule junctions. In such cases, problems associated with collective electrostatics (*i.e.* steps in the electrostatic energy due to periodic arrangements of dipoles) come into play.[4,8,10] As discussed above, cluster calculations without electrostatic embedding will fail to account for these effects of extended polar surfaces. Conversely, the periodic boundary conditions in repeated slab simulations automatically include them. This, however, becomes a problem, when the investigated system does not contain a periodic arrangement of dipoles. This, for example, applies to single-molecule junctions. Their correct description employing periodic boundary conditions requires large unit cells to prevent spurious interactions between neighboring dipoles

and to avoid collective electrostatic effects in the simulations that do not occur in the actual junctions.[12,77] Similar considerations apply when modelling local excitation processes at a surface, as in the case of final-state simulations of core-level binding energies. With periodic boundary conditions, they artificially model a core hole in each unit cell, in sharp contrast to the actual situation encountered in the experiments.[21] In such cases, where a periodicity of the interface is assumed that does not exist in the actual system, repeated slab calculations can produce even qualitatively wrong results.

3.3.2 Charged unit cells. Another issue that arises due to periodic boundary conditions is the problem of describing charged moieties. Non-neutral entities arise in several physical problems, *e.g.*, when dealing with defects in semiconductors[18,31] or when modelling ionization processes.[95] Within the repeated slab approach, it is generally not possible to charge only a single unit cell. Rather, periodic boundary conditions imply that all unit cells in the system carry a charge. A periodic arrangement of (non-compensated) charges, however, results in a diverging energy. This prevents convergence of the SCF algorithm for charged unit cells.

In practice, by default all band structure codes compensate unit cells containing an excess (or a deficiency) of electrons by assuming a homogenous background charge with opposite polarity, which exactly compensates the excess charge. This is a feature of the Ewald summation, which neglects the "$G = 0$ term". Including that term would cause numerical problems and neglecting it has no consequences for neutral unit cells, as for those it is exactly zero. For charged unit cells, neglecting the "$G = 0$ term" corresponds to restoring charge neutrality by the above-mentioned homogeneous background charge. Therefore, analogous to the electric field in dipolar cells, this homogenous background charge is not a particular feature of a specific code (which could be switched on or off *via* a keyword), but rather appears automatically.

A common scenario, where homogeneous charge backgrounds are used are, *e.g.*, first principles studies of charged defects in bulk systems. Although these calculations suffer strongly from a dependence on the size of the supercell (*i.e.*, the charge density of the compensating background), several efficient correction and extrapolation schemes have been developed to compensate for these problems.[96–100] Unfortunately, they are not applicable to interfaces:[20] A particular problem for interfaces is that there is a Coulomb interaction between the excess charge, that is localized in the slab, and the delocalized homogenous background charge. This gives rise to a significant contribution to the total energy of the system, as shown in Fig. 6b for the example of a 4-layer Cu(111) slab containing one excess charge per 1×1 surface unit cell. The spurious energy contribution originates from the spurious net dipole of the unit cell and, hence, scales linearly with the thickness of the vacuum region,[101] as shown schematically in Fig. 6a.

Two classes of approaches have been developed to deal with this problem. The first corrects for the electrostatic potential of the excess surface charge by interfering with the Poisson equation that describes the electrostatic potential[104–106] or *via*

8140 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021
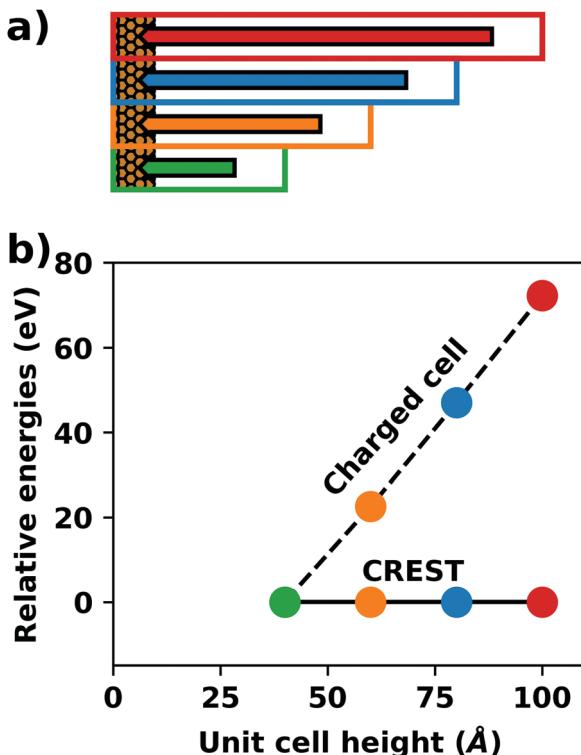
Fig. 6 (a) Schematic representation of the slab and the evolution of the dipole when varying the unit cell height for a charged unit cell. Provided that the vacuum region is much larger than the slab, the dipole of the cell amounts to $Q \cdot z_0 / 2$, with Q being the net charge of the slab and z_0 referring to the unit-cell height. (b) Dependence of the energy (relative to a 40 Å thick cell) on the unit cell height of a charged Cu(111) (1 × 1) surface unit cell where one extra electron was added and compensated with a homogenous background (dashed line) or with the Charged Reservoir Electrostatic Sheet Technique CREST[102,103] (solid line). All calculations were done using the PBE functional and 33 × 33 × 1 k-points and can be found under https://dx.doi.org/10.17172/NOMAD/2020.12.11-1.

a posteriori correction schemes based on the dielectric profile of the interface.[20] A variety of modified Poisson solvers have recently been proposed to tackle charged surface systems in the context of electrified interfaces[107] such as the Solvated Jellium approach[108] and the metallic boundary conditions proposed by Otani and Sugano.[104] Both approaches neutralize the interaction between charged cells perpendicular to the substrate. Note, however, that even there the lateral interaction between charged cells persists, requiring a careful converge of the supercell size. The downside of these approaches is that they are not yet widely available.

The second class of methods ensures that the slab as a whole is charge neutral, such that no compensating background charge is generated. This is typically achieved by intentionally adding spatially localized countercharges into the system. Two representatives of this class are the virtual crystal approximation (VCA)[18,109] and the Charge Reservoir Electrostatic Sheet Technique (CREST):[102,103] The VCA modifies the charge of the atomic nuclei, such that the excess charge in the cell is completely accounted for. This effectively localizes the countercharge to the slab, mitigating the size divergence

problem (but not completely eliminating it, since for a fixed amount of charge there is now a dependence on the substrate thickness). Hence, the VCA is commonly used to simulate doping of bulk semiconductors, where it provides a fixed number of free charge carriers per volume, rather than per area. CREST is an extension of the VCA and models the countercharges as a charged sheet, which is placed below the substrate. The additional energy contribution due to introducing that charged sheet at a specific position in the vacuum region is then corrected analytically,[103] eliminating the size dependence. This approach can also be used to mimic the impact of band-bending and charge transfer for adsorbates on (doped) semiconductor substrates.[43,110,111] A conceptually similar approach is the generalized dipole correction approach,[112] which introduces a monopole sheet as a "computational electrode" and a dipole layer in the vacuum region to decouple charged periodic images to produce boundary conditions equivalent to isolated slabs.

3.3.3   The commensurability conundrum. Slab-type calculations can only employ a single set of periodic boundary conditions (*i.e.*, one size of the lateral unit cell). Therefore, both the substrate and the adsorbate are subject to the same lattice periodicity and must share a common supercell. This is only possible if the adsorbate unit cell is an integer multiple of the substrate unit cell (or *vice versa*) such that a common supercell can be defined,[113] as shown in Fig. 7a and c. More precisely, the epitaxy matrix, that defines the relation between the adsorbate and substrate unit cells, must contain only integer entries. More details on how epitaxy matrices are defined and a
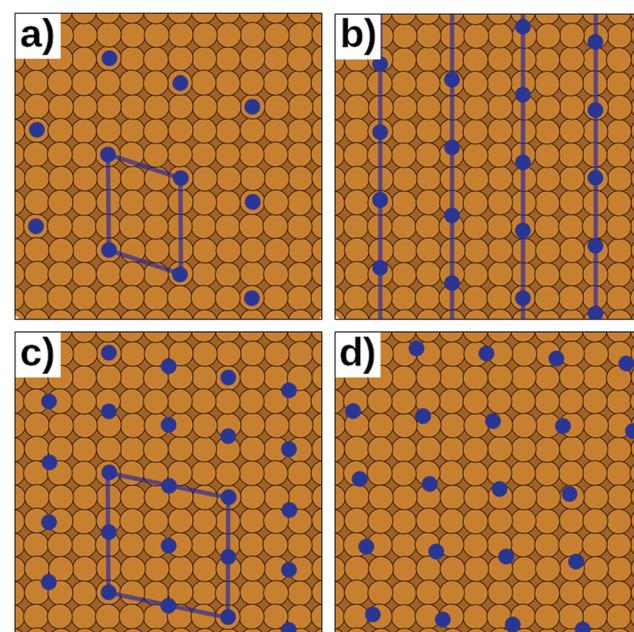

Fig. 7 Different types of epitaxy. Orange circles represent the substrate, blue circles the adsorption site of the adsorbate. (a) Commensurate interface. (b) Point-on-line epitaxy. (c) Higher-order commensurate interface. (d) Incommensurate interface. Only a and c can be modelled with periodic boundary conditions. b and d require streching or compressing of either the substrate of the adsorbate.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8141**

discussion of the different types of commensurability (including epitaxy, point on line, line on line, etc.) can be found in a recent review by Forker et al.[114] Following the Frenkel–Kontorova model,[115] commensurate structures form only under certain conditions (coincidence notwithstanding): The interaction of the substrate with the adsorbate must be sufficiently strong such that there is a clear, energetically preferred adsorption site for each molecule. At the same time, the interaction within the molecular adsorbate layer must be sufficiently weak such that the molecules are not expelled from the potential well of that adsorption site.

In reality, of course, not all hybrid inorganic–organic interfaces are commensurate in all directions (see Fig. 7b and d). Still, when modelling an interface using periodic boundary conditions, commensurability must be enforced. If the deviation from commensurability is not too large, this can be done by compressing and/or expanding the lattice constants of either of the interface components. In practice, this implies that an appropriate supercell of the substrate and the adsorbate (probably containing 2 or more molecules) should be sought that minimize the relative lattice mismatch. When modifying the lattice constant(s), care has to be taken not to perturb the properties of the material too much, e.g., by verifying that the degree to which the band structure, the work function, and similar properties are affected by the change of the unit cell parameters is still acceptable.

Even then, one should keep in mind that any distortion of the lattice parameters will make covalently bonded systems more reactive (as they are no longer in their energetic minimum). A second effect is that in a truly incommensurate system every adsorbate molecule resides at a different relative position on the substrate. Consequently, in the full layer all adsorption sites are equally sampled independent of whether they are energetically favorable or not. Conversely, when artificially enforcing commensurability, only favorable adsorption sites will be occupied. This can be most easily understood, when considering only a single molecule in a fixed supercell. Moving the modelled molecule over the substrate does not change the molecule–molecule interaction. Hence, the whole layer can be translated relative to the surface, until the most favorable molecule–adsorbate interaction is found. This and the higher reactivity of the modified sub-systems tend to lead to an overestimation of the substrate–adsorbate interaction, and, thus, to overestimated adsorption energies. However, also the opposite effect may occur, if, e.g., the stretching of the substrate introduces a lattice mismatch between the docking groups of the adsorbate and the docking sites of the substrate. The magnitude of either effect can be appreciable[116] and can easily be larger than the energetic differences between different adsorption sites or polymorphs.[117] Consequently, for such interfaces authoritative first-principles structure search or the reliable simulation of surface phase diagrams is currently an unsurmountable challenge.

Presently, there are few approaches of which we are aware that overcome the commensurability conundrum. If the molecule–substrate interaction is very weak, certain properties can be described by modelling a freestanding monolayer defined by the periodicity of the adlayer alone, for example when simulating STM topographies. An alternative is to map the DFT interactions onto more approximate methods like interatomic potentials[116] (i.e., force fields) or semi-empirical methods. Today, this mapping can be done efficiently with the help of machine-learning methods.[118–123] The computationally cheaper methods can then be efficiently evaluated for extensive domains containing thousands of molecules, potentially providing a more accurate representation of incommensurate interfaces.

3.4 Structure of the interface

The properties of an interface are strongly affected by its atomistic structure. While setting up the initial structural model of the interface (i.e., the input geometry of the calculation), several assumptions need to be made that can have a fundamental impact on the obtained results. These include, e.g., the size and shape of the unit cell, the number of molecules to be placed in that cell, the adsorption sites and orientations of the molecules, etc. Many of these decisions are immutable and will not change during the calculations. This is obvious for factors like the coverage, but also applies (at least to some extent) to aspects like the orientation of the molecules and their adsorption site, as will be discussed in the following.

The choice of the initial setup can be massively simplified when structural input from experiments is available. Unfortunately, such input is typically limited, since most experimental methods yield only incomplete atomistic insight. The immediate interface region typically comprises only a monolayer or at most a few layers of organic material, which often provides too few scattering events for X-ray diffraction to yield the full structural information. Low-energy electron diffraction (LEED) provides information about the unit cell of the organic adsorbate, but yields only limited information on the geometry of the adsorbed molecules regarding their adsorption sites or on their orientation relative to the substrate. Scanning tunneling microscopy (STM) can provide this information when performed with atomic resolution, but imaging both the substrate and the adsorbate simultaneously works only in fortunate circumstances. Also, STM can only be applied to conducting substrates. Moreover, it should be kept in mind that STM measures the local density of states, not the positions of the atoms, and that the appearance of an adsorbate in STM can be deceiving.[124–131]

Also, the composition of the substrate surface and its impact on the interface structure is often not a priori clear. Many metal–organic interfaces incorporate adatoms from the substrate,[8,82,132–139] which have a decisive impact on the geometry of the organic layers as well as on the interface properties.[8,82,132,133,140,141] A similar problem occurs for semiconductor surfaces, which can exhibit different reconstructions and surface terminations. One example for this is the ZnO(0001)-surface, which, depending on the sample history, may contain a different number of hydrogen or hydroxyl groups on the surface or form triangular pits with missing surface atoms.[39,101,103,142–144] Like for metal surfaces, these surface modifications can substantially change the interface geometry and chemistry.[145,146] A particular challenge in this context is

8142 | Phys. Chem. Chem. Phys., 2021, 23, 8132–8180

This journal is © the Owner Societies 2021

that adatoms or hydroxyl group are frequently not detectable in (especially LEED or STM) experiments. Rather, their presence can often be inferred only indirectly, *e.g.*, through their impact on the geometry, which requires a combination of theoretical and experimental techniques.

In lieu of experimental information, scientific studies often consider only a single molecule in a large supercell in order to "avoid spurious interactions" or to model a theoretical low-coverage regime. This strategy does not account for the fact that the absence of interaction may be spurious in itself, or that at nominally low coverages there could be island growth, with layers that locally display dense packing. Spuriously large unit cells particularly affect the adsorbate geometry: For (mostly) upright standing molecules, the coverage and packing motif crucially determines their tilt angle,[147] but also for simple, weakly interacting, flat-lying molecules, a dependence of the adsorption height on the molecular coverage has been reported.[148,149] Moreover, a surprisingly large number of molecules shows coverage-dependent re-orientations, phase transitions, or conformational and chemical changes.[62,150–160]

Even if the "single molecule in a supercell" approach is the best suited model (*e.g.*, because the molecules repel each other on the surface), finding the correct geometry (*i.e.*, adsorption site, orientation, conformation, *etc.*) is not necessarily straightforward. Even for an isolated molecule, multiple possible adsorption sites and conformations exist, often with notably different properties. This point is exemplified in Fig. 8 for the case of 6,13-pentacenequinone on Ag(111). Systematically scanning the potential energy surface for a single molecule in an approx. 252 Å^2 large supercell, we detect 10 different stable local minimum structures. The displayed data have been calculated with PBE+vdW^{surf}[161] and were taken from Jeindl *et al.*[162] The different structures and their adsorption energies are depicted in Fig. 8a. Notably, six of the ten found structures are almost isoenergetic (*i.e.*, within 50 meV); only four structures are substantially less stable. Despite the very similar adsorption energies, the adsorption heights vary substantially between the structures (especially for the oxygens), as shown in Fig. 8b.

We emphasize that the existence of a variety of different adsorption structures is not a peculiarity of this example. Even for the relatively small molecule tetracyanoethylene (TCNE), a large variety of different structures has been found on various metal surfaces both theoretically[153,158,163,164] and experimentally[134,153] and there are plenty of additional examples.[162,165–167] The existence of multiple minima of the potential energy surface for even a single adsorbed molecule illustrates a fundamental problem for first-principles studies of interfaces: Unless a
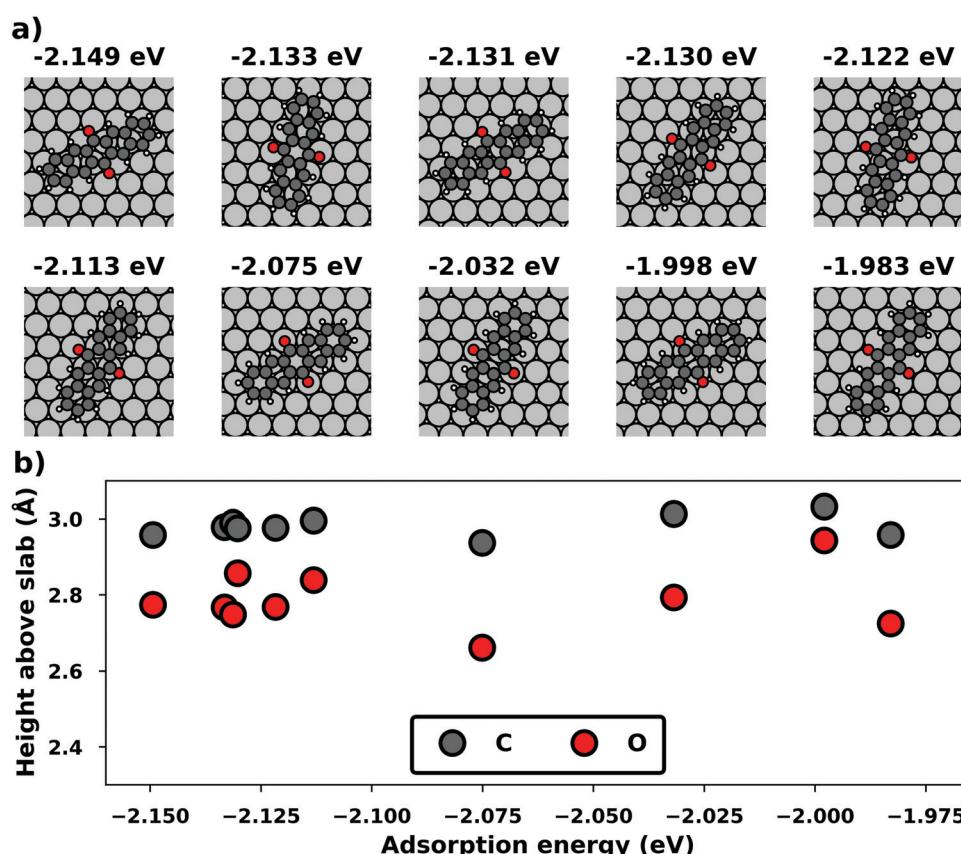

Fig. 8 Overview of the different local minimum structures for the adsorption of pentacenequinone on Ag(111), calculated with 3 × 3 × 1 *Γ*-centered *k*-point grid using PBE+vdW^{surf}. (a) Top view of the different structures, together with their adsorption energies, calculated as the energy difference between the system and the sum of slab and molecule on their own. (b) Average position of carbon (grey), and oxygen (red) atoms above the topmost silver layer, ordered according to their adsorption energy. Further calculation details can be found in Jeindl *et al.*[162]

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8143**

systematic structure search is performed, it is easy to miss the global energetic minimum, which is arguably the most relevant structure. A survey of literature indicates that systematic searches are the exception rather than the rule. (Although, admittedly, it is often not clear to which extent some search based on trial-and-error was performed, but not reported. Thus, we recommend that the details of such a search are documented more often in the published literature). Even when systematic evaluations are done, they mostly rely on creating different starting points based on physical and chemical intuition followed by local geometry relaxations, rather than employing more unbiased and systematic approaches like Basin Hopping[168] or machine learning-based algorithms.[166] As a further complication, local relaxations have their own challenges, as will be shown in Section 5. Yet, without any systematic search, a (more or less) random local minimum will be found. Consequently, different scientists studying identical interfaces will not necessarily find the same (local) minimum structures, which negatively impacts the reproducibility of first-principles interface studies, at least unless the full geometry is provided (e.g., via data repositories). Moreover, the variation of adsorption geometries can introduce a notable (perceived) error bar on the calculations and significantly complicates the evaluation of the performance of the employed methodology compared to real-word experiments (see Section 4). Interestingly, recent advances in global structure optimization have finally put global structure search at hybrid interfaces including predictions of their atomistic structures[168–170] and even an assessment of different polymorphs, within reach.[165,171–173] In passing, the authors note that their algorithm for global structure search at interfaces (SAMPLE)[165] is available as a python package that can be downloaded from the author's homepage (www.if.tugraz.at/hofmann). Notably, global structure search also provides an overview of the potential energy surface, which then allows to either perform Monte-Carlo simulations in order to predict the morphologies of (sub)monolayers under kinetically determined growth conditions,[171] to determine thermodynamically stable structures and phase diagrams,[165] or to account for the impact that defects with low formation energies have on the structure.[163]

4 The electronic structure method

Despite recent advances in the efficiency of correlated wave-function-based methods, their application to hybrid inorganic–organic interfaces remains expensive and is mostly restricted to small (embedded) clusters.[47] Thus, band-structure calculations based on density functional theory remain the workhorse for first-principles studies of interfaces. There exists, however, a plethora of different exchange–correlation (xc) functionals (empirical and non-empirical) that could be applied. This raises the question which xc functional should be used in practice.

Kieron Burke, one of the pioneers of modern xc functionals, stated that "Users should stick to standard functionals [i.e., PBE for materials], or explain carefully why not".[174] This statement certainly also holds true for interfaces, although it

should be amended by stating that a van der Waals correction is almost always required. Nonetheless, there are sometimes good reasons to deviate from the standard approach. Since the performance of different electronic structure methods is frequently reviewed, also in the light of interfaces,[47,175–181] here we will focus on a quick, rather general overview of some of the most relevant xc functionals available as well as the options to account for van der Waals interactions. Note that for the sake of brevity, some of the following statements may be (over)generalizing. For a more comprehensive discussion on how to construct and choose xc-functionals (including a benchmark of their performance) the reader is referred to the recent reviews of Burke[182] and of Truhlar.[183]

4.1 Density functionals for interfaces

The core idea of density functional theory, as laid out by Hohenberg and Kohn,[184] is that the ground state properties of a system are encoded in its electron density, i.e. there is a functional that directly connects the systems' electron density to its total energy. Although this greatly simplifies first-principles calculations in theory, the correct functional is still unknown. Later, Kohn and Sham demonstrated that the correct electron density can also be obtained from an ensemble of non-interacting particles.[185] For such a system, it is straightforward to write down the equations for the kinetic and the Coulomb energy. Only the functional for exchange and correlation, for which no general form is known, must be approximated. Today, it has become customary to group the different approximations to the exchange–correlation functional into several rungs of the so-called Jacob's ladder, as coined by Perdew.[186] Within each rung, many xc functionals exist. These are either constructed to obey certain theoretical limits and sum rules, or parameterized to reproduce the properties of specific materials that were calculated with higher-level methods. Notably, these rungs roughly group the xc functionals according to their computational cost, but ascending the rungs does not necessarily mean that also the accuracy of the results increases, as we will discuss below. It is also important to note that many developments of the last 15 to 20 years do not neatly fit into these rungs, e.g. self-interaction corrected functionals or van-der-Waals-inclusive functionals. A detailed review of these recent developments in materials research was recently given by Maurer et al.[175]

The local density approximation (LDA) constitutes the 1st rung of Perdew's ladder and is the simplest of all constructions, where the xc-functional depends only on the value of the electron density at a given point in space. Although different versions of LDA functionals exist, they are all numerical parametrizations to the exchange and correlation energies of the uniform electron gas. In practice, they differ only in the functional form and parameterization, and all yield very similar results. It is well known that LDA gives a good account of the lattice constants and band structures of simple metals, but does not perform as well for molecules and semiconductors. Nonetheless, twenty years ago, LDA calculations were used routinely for interface calculations,[29] because they exhibit an

8144 | Phys. Chem. Chem. Phys., 2021, 23, 8132–8180

This journal is © the Owner Societies 2021

artificial energy minimum between subsystems, even if they are not covalently or ionically bonded. This was often taken to "mimic" van der Waals interactions, although this overbinding occurs for the wrong physical reason.[176,187–189] A nice example is the case of PTCDA on Ag(111), where LDA predicts an adsorption height in good agreement with X-ray standing wave measurements.[190] However, for most cases of hybrid inorganic–organic interfaces the adsorption distances computed employing LDA are too low compared to experiments.[176,190–194] Furthermore, LDA also shows a strong overbinding for intra-molecular bonds in molecules (resulting in too short inter-atomic distances[195]) and, generally, performs poorly for molecular properties.[196] Since the computational advantage over the next higher rung (GGA, see below) is very small, and due to the unsatisfactory geometries it predicts, LDA is hardly used today for modelling interfaces.

The first improvement over LDA is to account also for the gradient of the local electron density, giving rise to the Generalized-Gradient-Approximation (GGA). Due to the inclusion of the gradient, these xc functionals are called semi-local. In contrast to LDA, many different functional forms and parameterizations are in frequent use. For inorganic bulk materials, GGA functionals provide significantly better cohesive energies and lattice constants compared to LDA.[197,198] For interface simulations, the PW91 functional[199] has seen frequent use, since it recovers some of the artificial binding of LDA. The PBE functional,[200] which is a non-empirical simplification of PW91, is considered today's default for solid state physics problems and is the most widely used xc functional. Multiple modified versions of PBE exist, which involve a re-parameterization to other theoretical constraints, (rPBE[201]) to constraints specifically relevant for solids (PBEsol),[202] or to match atomization energies (revPBE[203]). In PBE, most molecules that are purely van der Waals bonded to the surface do not stick to the surface at all[175,191] (potential artefacts from basis set superposition errors and/or the geometry optimization algorithms notwithstanding).[204] Indeed, PBE generally tends to underbind,[205,206] even within molecules.[207] Its reparameterizations rPBE and revPBE are said to improve the description of chemisorption,[203] while PBEsol slightly overestimates the adsorption energies of chemisorbed moieties.[208] PBE has been shown to yield excellent predictions for the electronic properties of metals, especially the work function (within the experimental uncertainties).[209] It does, however, generally overestimate the polarizability and underestimate the dipole moment of isolated molecules.[210] Although one would expect that this poses a problem for interfaces, when the work function is determined by polar adsorbates, it is our general experience that (provided that the experimental geometry is used) the obtained results agree rather well with the experimental values obtained in ultra-high vacuum.[82,124,150,151,191,211–216]

Meta-GGAs are the next rung of functionals. They also consider the kinetic energy density, which is equivalent to the second derivative of the electron density. This makes them the logical next step up from GGAs. Multiple variants exist, such as

TPSS[217] (and its revised version revTPSS[218]) or SCAN[219] and its revised versions rSCAN[220] and SCAN-L.[221] Although the description of cohesive energies of metals as well as atomization energies of molecules is improved with these xc functionals, they are, to the best of the authors' knowledge, hardly ever applied to investigate the adsorption of large organic molecules on metal surfaces. This may be in part because accounting for dispersion forces by *a posteriori* correction schemes (see next section) is more difficult for this class of functionals (apparently because meta-GGAs already are relatively non-local, thus, undermining the concept of these *a posteriori* corrections). Still, dispersion interactions have been incorporated *via* non-local corrections of meta-GGAs[222] (see next section). These have been applied successfully to small, prototypical interfaces.[222,223] Unfortunately, at present we are unaware of more extensive, systematic tests for hybrid inorganic–organic interfaces, making it too early to assess the performance of meta-GGAs in the context of this work.

Hybrid functionals constitute the highest rung of functionals that can currently realistically be applied to hybrid inorganic–organic interfaces. In this rung, a fraction of Hartree–Fock (HF)-like exchange (sometimes referred to as exact exchange) is admixed to the complementary part of semi-local exchange. This exchange is non-local, which increases the computational effort over semi-local functionals typically by at least one order of magnitude. The computational cost is furthermore increased by the fact that hybrid functionals yield larger band dispersions, which has a two-fold impact: For metals, the larger dispersion leads to a smaller density of states (and, thus, available charge carriers) near the Fermi energy. Thus, slab-type interface calculations generally require more metal layers, *i.e.* larger systems, to provide a stable Fermi energy of the substrate, as exemplarily shown in Fig. 9 for a Cu(111) slab. Secondly, the larger dispersion generally[224] requires a denser sampling of the reciprocal space (*i.e.*, more k-points) to yield converged results.
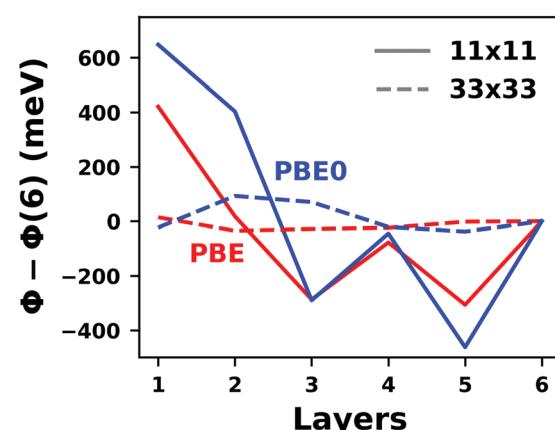

Fig. 9 Work function Φ of a Cu(111) slab (relative to the work function of a 6 layer-slab) as a function of the number of layers in the slab calculated with GGA-type xc functional PBE (red) and the hybrid functional PBE0 (blue). The calculations were done for a primitive surface unit cell containing 1 atom per layer with 11×11 (solid line) and 33×33 (dashed line) k-points. Calculations were done with FHI-aims using the "tight" defaults. Further details and results can be found at https://dx.doi.org/10.17172/NOMAD/2020.12.07-1.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8145**

Furthermore, hybrid functional-based calculations are more difficult to parallelize efficiently, which affects their scalability towards larger systems. Although these challenges are continuously addressed by new, improved algorithms, the large computational cost is an effective barrier that deters many users from routinely using hybrid functionals for interface simulations.

All hybrid functionals contain the fraction of employed Hartree–Fock-like (HF) exchange as one parameter. Similar to semi-local functionals, many different hybrid functionals with different design strategies exist. Here, only a few should be mentioned: PBE0 constitutes an extension of PBE, with a fraction of HF-exchange of 25%, as motivated on theoretical grounds.[225] Another popular approach is the B3LYP functional, which is heavily parameterized (to molecular data) and employs 20% HF-exchange. However, B3LYP is rarely used for hybrid inorganic–organic interfaces as its molecule-focused para-metrization does not provide consistent accuracy for metals or interfaces.[226] Generally, hybrid functionals significantly improve the quality of the prediction of the properties of isolated molecules.[227] They are, however, fundamentally problematic for metallic systems,[208,226] because of a divergence near the Fermi-energy (a problem they share with many wave-function based methods). This problem can be mitigated by using range-separated functionals, specifically short-range hybrid functionals, such as HSE06,[228] in which exact exchange is attenuated in the long range. This attenuation is well justified for metallic bulk systems, where exchange is effectively screened and the electron density "in the long-range" resembles a homogenous electron gas.[229] Conversely, for systems surrounded by vacuum (such as isolated molecules), theory would require that hybrid functionals, which employ 100% HF-exchange in the long range, are applied.[230] The latter approach is frequently used in optimally-tuned range-separated functionals,[231–234] such as the LC-w-PBE functional.[230] Such functionals try to solve the general band-gap problem of DFT by choosing the amount of exact exchange (and possibly the range-separation parameter) such that the IP-theorem[235] is fulfilled, i.e. such that the energy of the highest occupied Kohn–Sham state is equal to the ionization energy. For interfaces, this poses a clear challenge: Individual, small molecules typically require large fractions of Hartree–Fock-like exchange to fulfill the IP theorem,[236] while for metals, smaller values suffice.[212] Although this problem is partly mitigated by the fact that for interacting systems, i.e. for molecules near a surface, the ideal amount of exchange is significantly reduced,[237] the conundrum remains that different values would be required for the substrate and the adsorbate. A more detailed discussion of the role of exact exchange for modelling interfaces, how to potentially determine a suitable value for its relative weight, and its impact on the transfer and distribution of charges at interfaces is given by Wruss et al.[237] Generally, hybrid functionals (of all kinds) tend to improve the description of molecular properties, including geometries, dipole moments and polarizabilities.[210] Kresse et al. showed for pristine substrates that work functions calculated with hybrid functionals are generally smaller than those obtained from semilocal approximations.[208] This can be ascribed to the fact that hybrid functionals yield more localized electron densities, which decreases the surface dipole of metals. Notably, the trend of smaller work functions for hybrid functionals does not necessarily transfer to the simulation of metals covered by an adsorbate layer, where the surface dipole is changed due to pushback and other effects, such as charge transfer.[212]

4.2 Long-ranged dispersion interactions

Van der Waals (vdW) interactions between individual atoms are quite weak, typically amounting to about 0.1 eV per atom. However, since they act between all atoms of a system, the grand total of the vdW energy for extended molecules adsorbing on surfaces or forming molecular crystals becomes quite substantial. Generally, van der Waals interactions are the main contribution to the cohesion of organic crystals[238] and organic layers.[66] Without them, many interfaces would be inherently unstable.[191] Even for covalently bonded systems, a correct description of vdW interactions is crucial, e.g., for the adsorption height of flat-lying molecules.[176] In short, vdW interactions play a decisive role at hybrid inorganic–organic interfaces; maybe even so much that it is fair to say that the advent of van der Waals inclusive functionals and a posteriori corrections, that became commonplace for interface calculations between 2006–2010, has created a schism in the modelling of hybrid inorganic–organic interfaces. Whereas before, large deviations between experimentally and computationally determined adsorption geometries were readily accepted as a reasonable agreement, today adsorption heights within 0.1–0.2 Å of the experiment can be readily obtained[176] and even these deviations are potentially due to numerical rather than due to physical approximations, as we show in Section 5.

The first rung of density functional theory that intrinsically includes (long range) van der Waals interactions would be the random phase approximation (RPA). Due to its large computational cost and difficulty of generating converged and numerically robust results, this method is rarely used to describe relevant hybrid inorganic–organic interfaces.[239,240] In addition to the RPA, there are presently two conceptually different approaches that allow treating van der Waals interactions within density functional theory: a posteriori dispersion corrections that add energies derived from analytical expressions describing van der Waals interactions to the DFT energies and functionals containing an explicit non-local correlation contribution (so-called van der Waals functionals). Here, we will only briefly summarize the basic ideas behind these approaches, and discuss the effect they have on interface calculations. A comprehensive discussion of these approaches goes beyond the scope of this work, and the interested reader is referred to pertinent reviews.[175,176,178,189,241–248]

4.2.1 A posteriori correction schemes. Today, the most common way to account for van der Waals interactions is via pairwise-additive dispersion correction schemes.[249–253] Here, the vdW energies are computed via an analytical expression after the electronic self-consistency cycle has been converged. This approach is justified by the observation that, with a few

8146 | Phys. Chem. Chem. Phys., 2021, 23, 8132–8180

This journal is © the Owner Societies 2021

exceptions,[254] vdW interactions have a very small impact on the electron density. In their simplest form, the vdW interaction energies (E_{vdW}) are given as

$$E_{vdW} = -\frac{1}{2}\sum_{A \neq B} f_{damp}(R_{AB})\frac{C_{6,AB}}{R_{AB}^6}, \tag{1}$$

where the sum runs over all possible combinations of atoms A and B. R_{AB} is the distance between those atoms, C_6 the effective interaction parameter and f_{damp} an (empirical) damping function. Several different dispersion correction schemes exist, which vary (mainly) in their approach for creating dispersion correction coefficients (C_6) that reflect the chemical environment of an atom in a molecule or solid. The most commonly used schemes for hybrid inorganic–organic interfaces are that of Grimme (the DFT-Dx[250] class of vdW correction schemes, with their most recent versions DFT-D3[255] and DFT-D4[256]), the exchange-hole dipole moment approach of Becke and Johnson,[257] and the methods developed by Tkatchenko and coworkers (vdW(TS)[251] and its successors vdW^{surf} [161] and MBD-NL[258]).

The Grimme-type correction schemes were the earliest widely adopted correction schemes. In their earlier, now outdated variants (DFT-D[249] and DFT-D2[250]), they rely on fixed coefficients for a given atomic species, neglecting the impact of the chemical environment. More contemporary variants renormalize the tabulated C_6-coefficients based on the molecular geometry, which substantially boosts the accuracy of this scheme.[255] The Tkatchenko–Scheffler-type corrections obtain the C_6 coefficient from the local polarizability of the atoms, and rescale the interaction parameters based on a Hirshfeld charge partitioning scheme in order to account for the chemical environment of the atoms.[251] For both pairwise-additive schemes, several improvements over the basic variants exist. This includes variations of the D3 scheme to account for effects beyond the dipole-approximation,[248] three-body interaction schemes,[259] or the use of alternative damping functions.[252] For both the DFT-D and the vdW(TS) scheme variants exist which implicitly account for the screening of dispersion interactions in extended metals *via* modified parameters.[161,260] Importantly, both schemes have been specifically reparametrized for interfaces, which has been done separately for metal substrates, like in the vdW^{surf} scheme,[161,261] and for ionic crystals and surfaces.[262,263] Using these reparametrized schemes seems highly advisable: For metals, because the corresponding parameterization partly restores the correct physics;[161] for ionic crystals, because using the conventional parametrizations yields incorrect lattice constants or causes unphysical surface reconstructions. The vdW^{surf} scheme has shown reliable performance in describing the adsorption structure of large conjugated molecules at metal surfaces, but yields overestimated adsorption energies.[176,264] An extensive list of parameters for the vdW^{surf} scheme has been published.[265]

When used out of the box, both the Grimme and the Tkatchenko–Scheffler approaches have known difficulties treating systems which undergo strong charge transfer, which is a common scenario at hybrid organic inorganic interfaces.[248]

In principle, D3 allows dealing with atoms that are far from neutral (*i.e.*, that experience strong charge transfer), but this requires manual tinkering with the reference systems used for the parameterization.[262] This situation has been recently improved with the D4 variant, which rescales the interaction parameters based on an atom-in-molecule charge partitioning scheme.[266] One could naively expect that the vdW(TS) scheme would automatically capture charge-transfer effects, since it relies on a charge-partitioning scheme. However, the Hirshfeld partitioning sometimes fails to yield correct charges (especially for negatively charged moieties and cases of strong charge transfer). The reason for this is that in the Hishfeld partitioning scheme the density is partitioned according to the densities of free, neutral atoms. Thus, the parameters in the vdW(TS) scheme are not sufficiently rescaled in cases, where a free ion would be a much better reference.[267] Here, the situation can be substantially improved by using an iterative Hirshfeld scheme,[268,269] which includes such an improved reference combined with an interpolation of tabulated C_6 parameters between neutral and charged free-atom species. In fact, we observed for complex interfaces with multi-component adsorbate layers comprising organic acceptors and alkali metal atoms that this approach yields clearly more accurate adsorption heights[270] (compared to X-ray standing wave experiments) than simulations based on the standard vdW^{surf} parameters.[125]

The Many Body Dispersion (MBD) method by Tkatchenko and coworkers can be seen as a method that neither fits into the pairwise additive methods nor the non-local correlation functionals.[271] Based on the vdW(TS) or vdW^{surf} C_6 rescaling approach,[161] the MBD method and its variants[272,273] create a set of quantum harmonic oscillators at the positions of the atoms, parametrized by atomic polarizabilities derived from Hirshfeld partitioning. These oscillators are coupled within the dipole approximation to calculate a long-range correlation energy. Although this approach has shown promise for hybrid organic–inorganic interfaces,[274] the description of the polarizability response *via* atom-localized harmonic oscillators is not able to properly capture the collective polarizability response of metals. A recent, non-local extension of the MBD correction (the MBD-NL scheme)[258] exists that replaces the atom-in-molecules parametrizations of atomic polarizabilities with a non-local functional to compute the polarization.[275] The approach remains to be systematically tested, but it is expected that this correction is universally applicable to molecules and solids[258] and, thus, should perform well for hybrid inorganic–organic interfaces.

4.2.2 Non-local van der Waals functionals. An alternative approach to include long-range van der Waals interactions is to use an xc functional that directly captures non-local correlation. At the same time, it should avoid a summation over unoccupied states, in order to keep the functional computationally tractable. Today, most non-local van der Waals functionals contain correlation of the form

$$E_{corr}^{nl} = \frac{1}{2}\iint \rho(r)\Phi(r,r')\rho(r')\mathrm{d}r\mathrm{d}r'. \tag{2}$$

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8147**

Here, ρ is the electron density and Φ is a pre-computed kernel. Different types of functionals differ in the exact form of Φ, which determines how the physics of non-local correlation is captured. Modern functionals base it on the local polarizability, which, for algebraic reasons, is approximated in a way that neglects screening effects.[242,248,276,277] Like the post-processing schemes, the non-local correlation can, in principle, be added to any xc functional, although this is much less straightforward here. However, the non-local form (requiring a double integration over space, similar to exchange in hybrid functionals) made their application within periodic boundary condition calculations expensive. It was shown that a reparameterization of the local exchange and correlation energy can substantially improve the accuracies of vdW-DF.[245] The recent reparameterizations by Vydrov and van Voorhis (VV09 and VV10) made these functionals computationally more efficient.[278,279] In contrast to the post-processing schemes, no dedicated re-parameterizations of the kernels for interfaces exist, nor should they conceptually be necessary. A new generation of vdW-DF methods has recently been proposed, that is specifically build to deal with various competing interaction mechanisms.[280,281] However, to the best of our knowledge, their performance has not yet been systematically assessed in the context of hybrid interfaces.

4.3 Consistency of the computational method and benchmarks

Many of the aforementioned functionals, combined with various dispersion-correction methods, have been applied to interfaces, including interfaces between rather large organic molecules and metallic or semiconducting substrates. Their success varies from case to case, but generally, all modern approaches for treating long-range dispersion interactions tend to perform reasonably well. However, that is not to say that all approaches yield the same results. Anecdotally, here we report the impact of the computational method on the structure of for PTCDA on Ag(111) (based on ref. 176 and 282), for which also a wealth of experimental information exists.[283]

Fig. 10a shows the adsorption height obtained with various functional/vdW-correction combinations. For the chosen system pure GGA-functionals (PW91/PBE/refPBE) yield the unphysical result that PTCDA barely binds. Conversely, combining the GGAs with the D3 or the vdW^{surf} correction yields results in rather good agreement with experiments,[284,285] as does the non-local VV10-functional. The values for the adsorption height (ca. 3.05 Å) are similar to those obtained with the Random Phase Approximation.[176] The more outdated vdW-DF and the vdW(TS) methods also yield adsorption geometries that are reasonable, but deviate more significantly from the experiment. Similar conclusions regarding the differences between the functionals are reached for the adsorption energies shown in Fig. 10b, although for the adsorption energy the "experimental value" should not be taken at face value (as it had to be extrapolated from a different molecule).[161] Nevertheless, it becomes clear that the values obtained with the different methods scatter notably. Already within the PBE functional, the different vdW-correction schemes yield adsorption energies
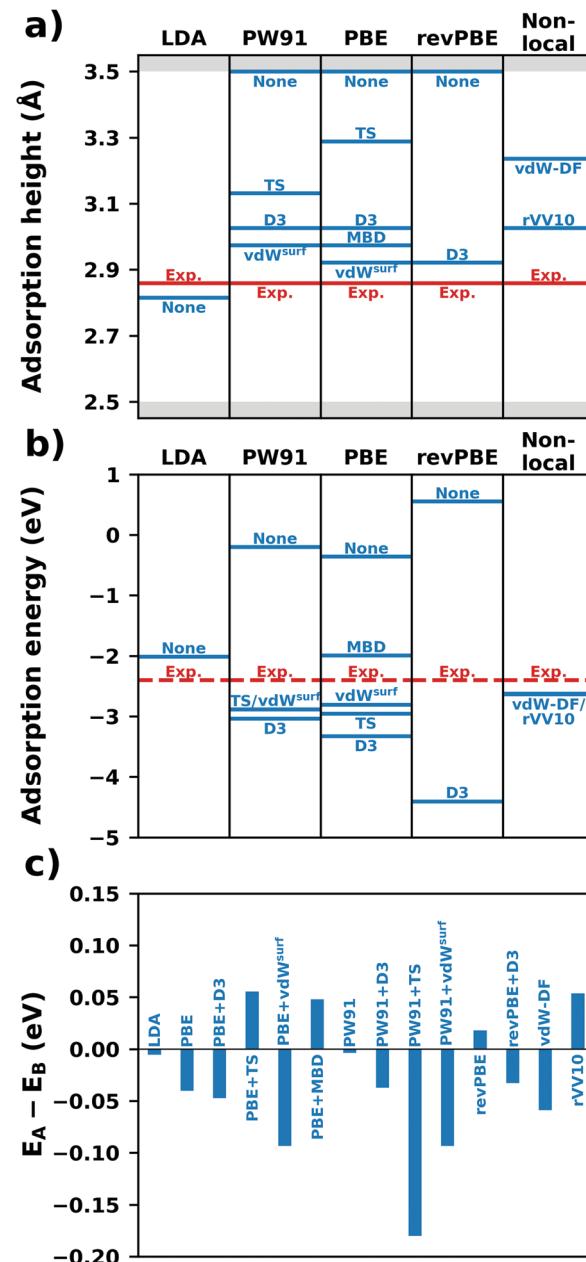

Fig. 10 Optimized adsorption heights (a) and corresponding adsorption energies (b) for PTCDA/Ag(111) with different methods. Results for different functional/vdW-correction scheme combinations. The different functionals (LDA, PW91, PBE, revPBE), were combined with the Grimme-D3 scheme and the Tkatchenko–Scheffler correction method in their original (TS), surface (vdW^{surf}) and MBD variant. The vdW-DF and the SCAN+rVV10 functional were used as representatives for non-local functionals (right column in a and b). (c) shows the energy difference between the two stable adsorption sites of the molecule. Calculations details are given in Hörmann et al.[282] The experimental values are reproduced from literature:[161,284,285] For the adsorption height, they were obtained by X-ray standing wave experiments,[284,285] the adsorption energy was extrapolated based on the chemically related NTCDA molecule.[161]

between −2.0 eV and −3.5 eV. The two tested non-local functionals SCAN+VV10 and vdW-DF, as well as PW91 with various correction schemes, are also within this energy window.

8148 | Phys. Chem. Chem. Phys., 2021, 23, 8132–8180

This journal is © the Owner Societies 2021

Unfortunately, these energy differences are not just rigid offsets between different methods, but affect different adsorption positions differently. We illustrate this again for PTCDA, which has two inequivalent stable adsorption positions when adsorbed on Ag(111). They differ mostly with respect to their orientation relative to the substrate.[283] As shown in Fig. 10c, the relative energies between the two minima assume values between −150 meV and +50 meV depending on the chosen approach. That means that even the relative ordering becomes methodology-dependent. Interestingly, Fig. 10c suggests that these variations depend more strongly on the vdW-correction than on the underlying functional.

On a more positive note, at least qualitatively the potential energy surface for interfaces seems to be relatively insensitive to the method. As shown in the Supporting Information of ref. 165, almost all methods consistently yield the two minima with very similar alignments of the molecules with respect to the substrate. Some of the methods yield additional, very shallow and energetically higher lying minima. Those geometries are not observed in the experiments, but that does not necessarily mean that they are not real.

Although PTCDA/Ag(111) is just one example, the considerations above show that the choice of the method matters. This has two important implications for practical purposes. First, it shows that the literature results obtained with different methods cannot be directly compared, at least not on a quantitative basis. Secondly, if different methods yield different results, the natural question arises whether there is one method that is clearly better than the others and could, therefore, be recommended for interface calculations?

We reckon that for a definite answer to this question not enough systematic tests have been performed to date (which is also not straightforward considering that the number of reliable experimental datasets is limited). In practice, at the time of this writing, the largest body of data exists for the PBE+vdW^{surf} method. For metallic substrates and both physisorbed and weakly chemisorbed organic adsorbates, it usually gives good results for the electronic structure[212] (work function within 0.2 eV of the experiment) and the adsorption height (typically within 0.2 Å of the experiment).[176] Similarly good results for interfaces have been obtained using PBE+D3.[260,286]

4.4 Basis sets (plane waves *versus* atom-centered)

In addition to the choice of the structural model and the computational approach (which for interfaces typically boils down to the choice of the functional and the van der Waals correction), it is also relevant to decide onto which basis functions the (Kohn–Sham) orbitals are to be mapped (which, in practice, is typically another factor that determines which computational code is applied). Most modern band-structure codes can be separated into two different classes depending on the type of basis functions they use to represent the electronic structure: One type are plane-wave-based approaches, which rely on delocalized basis functions and where the basis can be systematically improved to achieve monotonic convergence.

However, they require a special treatment for the core electrons. For a detailed discussion of the various forms of that treatment, such as pseudopotentials,[287,288] projector-augmented waves (PAWs),[289] or linearized augmented plane waves (LAPWs),[290] the interested reader is referred to pertinent reviews and books. The other type are atom-centered basis functions, which are centered on the nuclei and straightforwardly describe core electrons, but for which no clear systematic improvement schemes exist, as the convergence of results with basis-set complexity is typically not strictly monotonic. For sparse systems (such as surfaces exposed to vacuum) and low-dimensional systems, atom-centred basis functions have the benefit that basis functions are only placed around atoms and do not cover vacuum regions. We note that also a variety of other types of basis sets exist (and especially real-space grid representations appear to gain traction),[291–294] but these will not be considered here. As shown by the seminal paper by Lejaeghere *et al.*[295] for bulk materials, simulations employing plane-wave and atom-centered basis functions give the same results, provided well converged settings are used. Note that convergence is always defined with respect to a certain observable. This means that the mere fact that convergence has been reached for a given property does not mean that this applies to all quantities of interest, as discussed in more detail in Section 5. It is, thus, generally important to perform careful convergence tests, that ensure that the chosen basis set, as well as related settings, such as cutoff potentials and Brillouin zone sampling, affect the quantities of interest for a specific study by less than an "acceptable" error margin.

Thus, it is interesting to discuss briefly how the expected errors due to underconverged settings differ for the two types of basis sets, when it comes to the computational description of interfaces. Two issues deserve particular attention: (i) The description of the electron density at the surface and its decay into vacuum, which determines the interface dipole and, thus, the sample work function, and (ii) the adsorption energy of a molecule on the surface.

4.4.1 Influence on the interface dipole. As a first step, the implications of the choice of the basis set for the interface dipole shall be discussed: By definition, plane waves are periodic basis functions that encompass the whole space. At an interface, the same basis functions are, therefore, responsible for describing the electron density in the slab and the (lack of) electronic density in the vacuum region. Because the number of used basis functions is finite, their values in the vacuum region do not perfectly cancel. This leads to patches of spurious electron density. Although that electron density is very small, under certain circumstances it can have a noticeable impact on the electrostatic potential and the work function. The spurious electron density can be reduced by increasing the number of plane waves and by a careful convergence of the SCF procedure, as will be discussed in more detail in Section 5. Conversely, atom-centered basis functions (which exist in many flavors, *e.g.*, tabulated numerically or as Gaussian-type basis sets) are only present in the vicinity of matter. Thus, with this approach, no electron density far away from the slab can exist. However, for

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8149**
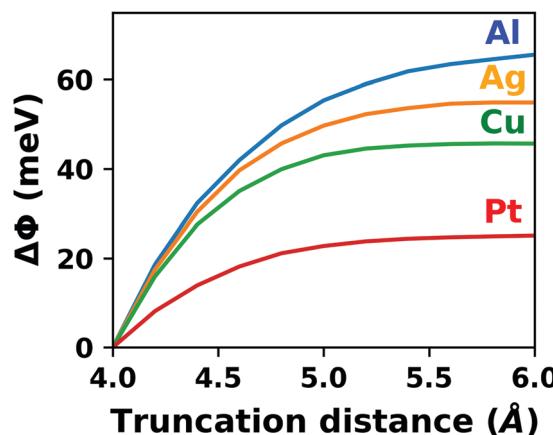
Fig. 11 Change of the work function ($\Delta\Phi$), relative to the default truncation distance of 4.00 Å, for different (111) metal surfaces as a function of the employed truncation distance (increased from the default value of 4 Å) for atom-centered basis functions, calculated with FHI-aims using the PBE functional and a $33 \times 33 \times 1$ k-point grid. More details and the full calculations can be found at https://dx.doi.org/10.17172/NOMAD/2020.12.07-2.

computational efficiency, atom-centered basis functions are truncated at a certain distance from the nucleus. This reduces the number of overlapping basis functions in regions where the overlap would be very small anyways, and in this way the natural sparsity in real-space can be used to significantly speed up calculations. For surfaces, this truncation limits how far electrons can spill out from the surface. However, accurately describing this spill-out is absolutely crucial for correctly describing the surface dipole (see Fig. 11), especially for metallic substrates. It can also subtly affect the tail of the electron density above the surface, which is often used to interpret STM topographs within Tersoff–Hamann simulations.[296,297] Particular care has to be taken here, as the default values for the truncation distance in most codes are designed for densely-packed bulk materials and can, thus, be too small for surfaces. Consequently, atom-centered codes tend to yield too small work functions and interface dipoles. As can be seen from Fig. 11, this can be solved by increasing (and converging) the truncation distance manually.

4.4.2 Influence on the adsorption energy. Adsorption energies are defined as the total energy of the combined system minus the substrate and the isolated molecule(s):

$$E_{ads} = E^{combined} - E^{substrate} - E^{molecules} \qquad (3)$$

When calculating the energy of the combined system, atom-centered basis functions are prone to basis set superposition errors (BSSE).[292] This well-documented effect[300] is a consequence of incomplete basis sets and arises from the fact that in the combined system the adsorbate will "borrow" basis functions from the substrate to improve the description of its own electron density, and *vice versa*. In other words, the combined system has a larger effective basis set than the individual subsystems, which artificially lowers the energy of the combined system and leads to an overestimation of the adsorption energy. The BSSE becomes

particularly relevant when using highly correlated methods. It can be mitigated by either using a sufficiently large basis, or by employing the counterpoise correction scheme (using the full basis set of the combined system also when calculating the subsystems using "ghost atoms"),[301] In this context it should, however, be mentioned that the suitability of this method has recently been controversially discussed.[302]

When employing plane-wave basis sets, the number of plane wave basis functions depends on the chosen cutoff value as well as on the size of the unit cell. The number of atoms, *per se*, does not play a role. Therefore, calculations with such basis sets are principally BSSE-free. For plane-wave basis sets, the challenge when calculating adsorption energies occurs for the isolated molecule as reference system. Since (with few exceptions) plane wave calculations require periodic boundary conditions, in practice, the isolated molecule is calculated as a single molecule in a large unit cell, electrostatically decoupled by dipole- and quadrupole corrections schemes in all spatial directions. However, since the cell needed to decouple periodic replicas might be different from that of a tightly packed combined system, the basis set used for the description of the isolated molecule is usually not the same as the basis set for the combined system and for the substrate. The impact of varying the size of the unit cell for a given cutoff energy is exemplarily illustrated for the case of the TCNE molecule in Fig. 12. The energy variation is rather small (amounting to a few meV), but in the same spirit as the BSSE, this inconsistency can lead to small errors for the adsorption energy. Conceptually, this problem can be mitigated by using an overconverged basis set, *i.e.* by increasing the basis set cutoff until the energy no longer depends on the unit cell size. However, that same cutoff must be used also for the combined system, making this strategy quite costly (and, given the small error, often not worth the effort).
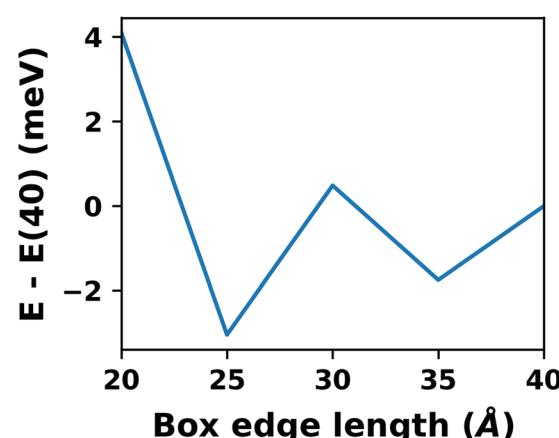
Fig. 12 Total energy of a TCNE molecule placed into a cubic box of increasing size, relative to the value obtained for the largest box (with an edge-length of 40 Å). The results were calculated with VASP[298,299] using the PBE functional and the default cutoff energy of ~ 290 eV for the plane waves. More details and the full calculations can be found at https://dx.doi.org/10.17172/NOMAD/2020.12.07-4.

8150 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

5 Algorithms, parameters and convergence: best practices for interface simulations

Once the major choices regarding the structural model, the employed xc-functional, and the basis set have been made, the simulation can be run. To solve the various parts of the Kohn-equations numerically, several different algorithms are used. A salient difference to the issues discussed in Sections 3 and 4 is that these algorithms and their settings are often somewhat hidden from the user, inasmuch as often default choices exist that are not mandatory to be specified in the input file.

As we have frequently mentioned in this work, the default choices are often sub-optimal for interfaces. In this section, we will, therefore, discuss some of the most common algorithms together with the usually applied settings. We will specifically focus on their major pitfalls when modelling hybrid inorganic–organic interfaces with band-structure based DFT, highlighting unfortunate instances where defaults can even lead to physically incorrect results. Although we use periodic boundary conditions for the examples in this section, we emphasize that the challenges demonstrated here can be analogously encountered in cluster calculations.

This section will be organized as follows: First, we will discuss the most critical settings within the self-consistent field (SCF) approach. These are (i) how to choose appropriate convergence settings, (ii) the importance of choosing the right initial guess for the electronic structure, (iii) different options to update the electronic structure (and the ideal parameters to achieve convergence fast), and (iv) the impact of level broadening and its dependence on the number of atoms in the system. We then proceed to discuss geometry optimizations, showing why sometimes completely different geometries are obtained, when using the same starting geometries, but different optimization strategies. As a final note, we demonstrate what can happen when a higher-level xc-functional is applied on top of a geometry obtained with a lower-level xc-functional.

5.1 SCF convergence

Within the self-consistent field (SCF) technique, the Kohn–Sham equations are solved iteratively until the results (*i.e.*, one or several properties of the system) no longer change between two subsequent iterations. When this is the case, *i.e.*, when the selected property no longer changes within a pre-defined threshold, the SCF is "converged". In principle, any property that can be computed can also be selected as convergence criterion, including in particular the total energy, the change of the electron density, the sum of eigenvalues, *etc.* In practice, many codes only consider the total energy (and, for geometry optimizations, the forces on the atoms). For details on the inner workings of SCF algorithms and the assessment of performance, we refer to the recent review by Woods *et al.*[303]

5.1.1 Spurious convergence. To illustrate two potential pitfalls, in Fig. 13 we show the evolution of the total energy

and the dipole moment for a system we recently investigated, a phenyl-piperazine based self-assembled monolayer bonded to the Au surface *via* a dithiocarbamate anchoring group.[304] For this example, the SCF was run with parameters that were "inherited" from a previous calculation (*i.e.*, performed well there). The details of the settings are not relevant here (they can be found at https://dx.doi.org/10.17172/NOMAD/2020.12.07-5), except that the SCF was set to converge when the energy change between subsequent iterations fell below 10^{-6} eV (a common, tight default) and to run for at least 60 steps.

Most importantly, we note that had we used only the change in energy as a threshold (orange line) as criterion, the calculation would have stopped at iteration 38. At this point, the SCF visited, by coincidence, two solutions to the Kohn–Sham equations that are almost identical in energy, leading to an apparent, spurious convergence. However, as can be seen from the blue line, the absolute energy (plotted relative to the energy at the last iteration) is still significantly above that of the fully converged energy (obtained at iteration 79 here). Such situations are rare, but not very. It is, therefore, generally advisable to inspect the SCF evaluation after each calculation. Sudden, surprising drops of the energy change are often an indication for spurious convergence and should be critically second guessed. In order to avoid this behavior in the first place, it is useful to employ additional convergence criteria (if the code allows), such as criteria for the change of the electron density or the sum of eigenvalues, as it becomes increasingly unlikely that multiple of these thresholds are met simultaneously unless the solution has approached the true minimum.

As second point, Fig. 13b illustrates that the convergence of the interface dipole does not directly correlate with the convergence of the energy. If our energy threshold for the energy change had been 10^{-4} eV (which is often sufficient to obtain reasonable energies), the calculation would have finished at SCF step 69 (the spurious convergence at iteration 38 notwithstanding). However, at this point the interface dipole would still be approximately 50 meV away from the value obtained when converging the calculation to an energy threshold of 10^{-6} eV (see blue curve). Even when continuing the SCF until the threshold of 10^{-6} eV is reached, the dipole still changes several meVs between the last iterations. This indicates that in this example, even using the relatively tight convergence criterion above may have been insufficient to obtain a fully converged value.

5.1.2 Convergence thresholds. We can now ask, which convergence threshold for the total energy would be ideal to obtain a numerically accurate result with as few iterations as possible. For this, we must first discuss how accurate our result needs to be. When comparing theory and experiment, it is rarely necessary to obtain energies with an accuracy better than 1 meV, so this is chosen as target accuracy. In fact, as we discuss below, for interfaces the limits imposed by the various numerical approximations (sampling of k-space, density of the integration grid, basis set, *etc.*) usually limits the numerical accuracy for the total energy to something on the order of 0.1–1 meV per atom, which translates to roughly 1–100 meV per unit cell, making
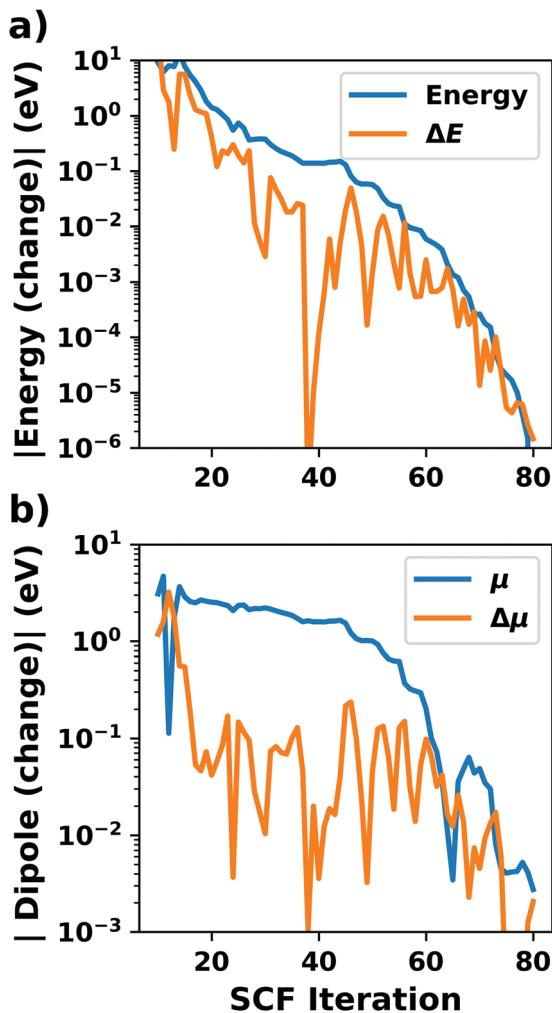
Fig. 13 Convergence behavior of phenyl-piperazinedithiocarbamate on Au(111) using reasonable but non-optimized settings. (a) Evolution of the energy (relative to the last energy given by the calculation) and the (absolute of) the change of the energy between subsequent iterations. (b) Evolution of the dipole moment (again relative to the last value given) and the change of the dipole moment between subsequent iterations. The original calculation and its settings can be found at https://dx.doi.org/10.17172/NOMAD/2020.12.07-5.

a much tighter target accuracy for energies obsolete. Once could, therefore, ask if this makes significantly higher convergence criteria obsolete.

To suggest guidelines for choosing an appropriate threshold, we exemplarily look at three different systems we have investigated in the past: (a) A tetracyanoethylene (TCNE) molecule lying flat on an Ag(100) slab,[163] (b) a phenyl-piperazine based self-assembled monolayer bonded to the Au surface via a dithiocarbamate anchoring group,[304] and (c) a graphene sheet with an adsorbed 4,4′-bis(phenylcarbonitrile) molecule.[305] The three systems are shown in the top panels of Fig. 14a–c. To illustrate how diverse the SCF procedure evolves, we re-calculated all of these systems using a plane-wave basis set,[298] an atom-centered basis,[306] and a regular real-space grid (using three different codes).[307] For the sake of illustration, we strictly adhered to default settings, rather than optimizing the

input file (except for switching on the dipole correction and enforcing an SCF threshold of 10^{-6} eV for the energy). Therefore, we emphasize that the results should not be mistaken as benchmarks, because (a) the time required per iteration differs between the different packages, and (b) because when choosing smarter algorithms and settings within each code, the number of iterations could be substantially optimized. Instead, the point of the comparisons in Fig. 14 is to show how different the convergence of different approaches can be.

When comparing the energy change between iterations ΔE (third row) with the value of the total energy (second row), it becomes clear that a given accuracy for E is typically obtained when ΔE is approximately one order of magnitude smaller (e.g., to obtain an energy accurate to within 1 meV, ΔE should be 10^{-4} eV or less). Notably, that statement only holds when the SCF has reached a state of "exponential convergence", where most iterations yield ever smaller values of ΔE (as opposed to "sideways movement" for the energy, as observed, e.g., for the blue line in the first 100 steps for system c). As above, we recommend to manually inspect the SCF behavior to assert that this exponential convergence region has been reached.

From the fourth row of Fig. 14, which shows the dipole moment, it becomes clear that the three approaches visit very different regions of the electronic potential energy surface before finding the minimum energy solution. In particular, the plane wave basis (blue line), has relatively large values for the dipole in the first steps. This behavior is a consequence of the basis set type (with a finite cutoff): The degree to which the total energy is converged is primarily determined by the electron density in the slab. As discussed in Section 4.3, in plane-wave calculations there are, however, also some small but non-zero amplitudes for the total wave function in the vacuum region. This spuriously places electron density into vacuum and, due to the overall charge neutrality of the unit cell, abstracts electron density from the slab. As the energy contribution of this dipole is very small (as discussed in Section 3), the SCF experiences only a weak "driving force" to remove this spurious dipole. We emphasize that this (only apparently inefficient) convergence behavior can be substantially improved by playing with the basis set cutoff and the size of the vacuum.

At the same time, the fourth row of Fig. 14 demonstrates that obtaining a converged energy does not necessarily result in robust and converged values for other quantities, e.g. a ΔE threshold of 10^{-4} eV does not mean that the interface dipole is converged to 1 meV. Because, in fact, most electronic properties do not directly correlate with the total energy, it is generally advisable to monitor not only this quantity, but also the quantity of interest during the SCF procedure. In fact, the best practice would be to include an explicit convergence criterion for the primary quantity of interest directly into the SCF cycle. I.e., when calculating adsorbate-induced changes of the sample work function, requiring the unit cell dipole to converge as one of the criteria to exit the SCF cycle is a very useful strategy. Although this can generally be implemented with only a few lines of code, unfortunately it seems that the convergence of the

8152 | Phys. Chem. Chem. Phys., 2021, 23, 8132–8180

This journal is © the Owner Societies 2021
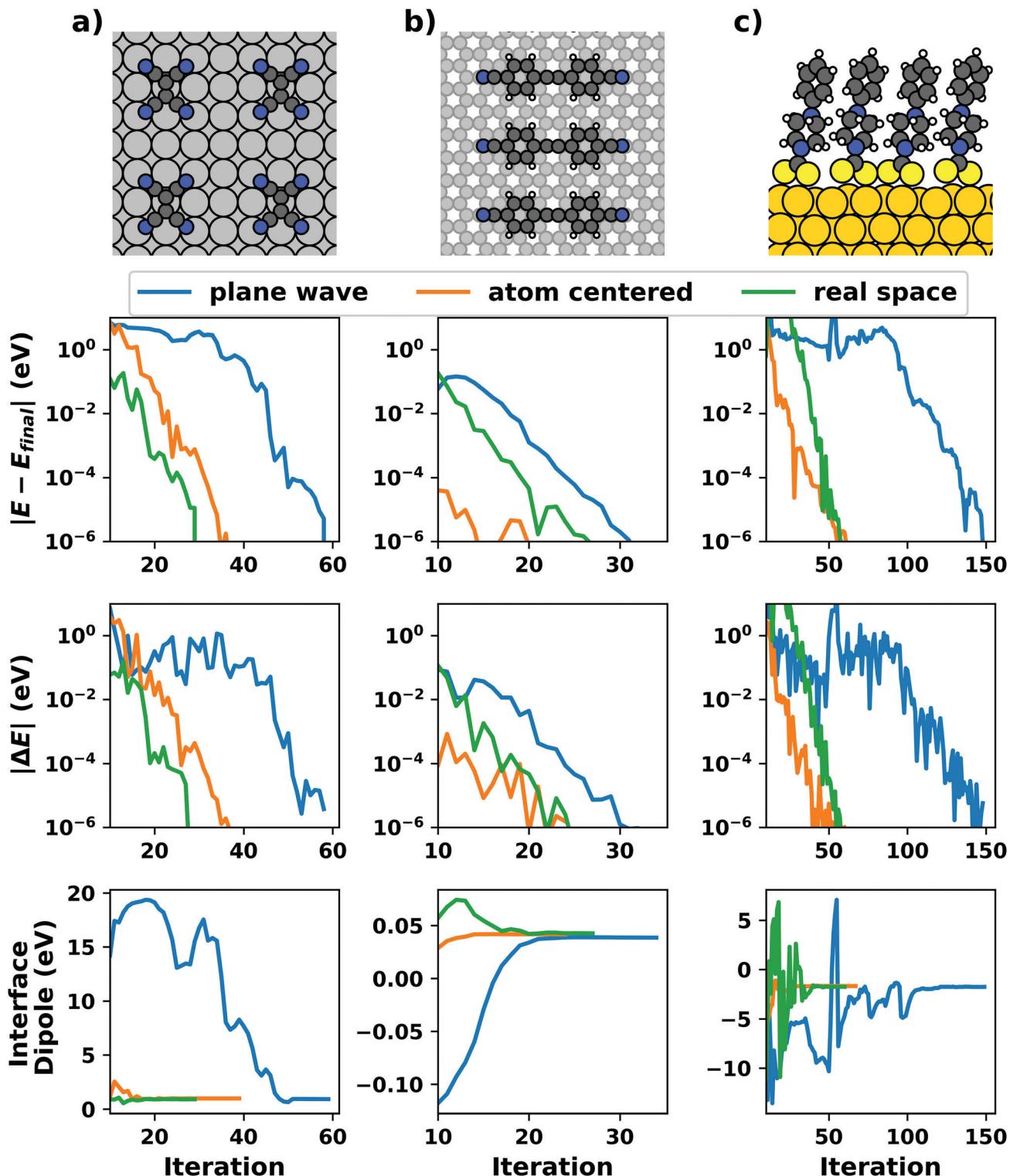
Fig. 14   Convergence behavior of the SCF for three model system using the plane-wave codes VASP, GPAW version 19.8.1b1 with a real-space grid, and the LCAO code FHI-aims. The different systems are (a) TCNE on Ag(100),[163] (b) phenyl-piperazinedithiocarbamate on Au(111)[304] and (c) bis-cyanophenyl on graphene.[305] From top to bottom, the panels show an atomistic representation of the simulated system, the evolution of the final energy (relative to the energy obtained after converging the SCF to a threshold of $\Delta E < 10^{-6}$ eV), the change of the total energy between two subsequent iterations, and the dipole moment per unit cell. For all numeric settings, the default values as supplied by the codes were used. We emphasize that the results should not be mistaken as benchmarks, because (a) the time required for the iterations differs between the different packages, and (b) because when choosing smarter algorithms and settings within each code, the number of iterations could be substantially optimized.

PCCP

View Article Online

Perspective

unit cell dipole is only rarely monitored or used as an SCF convergence criterion in practice.

In passing, we note that in the seminal work by Lejaeghere *et al.*, it has been established that all three codes used here give consistent results for energy-derived values[295] when using properly converged settings. Still, using default settings (and a ΔE threshold of 10^{-6} eV, as commonly done in literature), we find that the interface dipoles vary by approximately 5–10% between the three approaches/codes for the three studied systems. This certainly has to be kept in mind, when comparing results found in literature and when comparing computed to experimentally obtained values.

5.2 Initial guess for the SCF procedure

The initial guess for the SCF cycle provides the starting point for any calculation. Trivially, the closer the starting point is to the correct electron density, the fewer iterations will be required to complete the calculation. However, the starting point can also have a qualitative impact on the outcome of the calculation: In cases where multiple minima for the electronic structure exist, the initial guess for the SCF decides into which solution the SCF will converge. We found this to be of practical relevance for magnetic and open-shell systems. One anecdotal example is the adsorption of the (open-shell) molecule Cu-phtalocyanine on Ag(111).[308] There, when initializing the calculation with a spin of 1 μB or larger on the central ligand, it converges into a metal-centered state. Conversely, when non-zero initial spins are set on both the central atom and the four closest nitrogen atoms, the result is a more extended charge transfer, involving the ligand as well as the central atom. The two different states also show a different overall multiplicity: While the metal-centered state is a doublet, the state involving charge transfer to the ligand has triplet multiplicity.[308]

Typically, the initialization of the SCF procedure relies on one of two strategies: The method traditionally used by most quantum-chemistry codes, which are geared towards molecules and employ open boundary conditions, is to provide a guess of the wave function obtained by an approximate method that does not need to be calculated iteratively. Examples would be a guess derived from the extended Hückel method,[309] the Harris functional,[310] or the diagonalization of the core Hamiltonian. From the guess wave function, the electron density (and/or the density matrix) is constructed. The major advantage of this strategy is that it provides a high level of control over the initial guess, as the population of each orbital can be manually specified. The downside is that such methods are heavily parameterized and that suitable parameters often do not exist for heavy atoms or non-trivial bond topologies, as encountered at interfaces. Therefore, wave function-based guesses are rarely used in band-structure codes.

Instead, most band structure codes rely on an initial guess for the electron density, rather than the wave function, for which the (initial) basis set coefficients are initialized at random. The electron density guess is commonly obtained using the so-called superposition of atomic densities (SAD).[311,312] Here, for each atom in the system a hypothetical, non-interacting electron density is calculated by solving the corresponding radially-symmetric Kohn–Sham equations. The first guess of the electron density is then given by the sum of all individual electron densities. This method has the advantage that it provides a relatively good idea about the electron distribution of the core electrons and even a rough description of the valence electrons. However, it does not contain chemical insight, *i.e.* the initial guess is agnostic of the bonding environment. Consequently, for metals, where the bonding is (mostly) isotropic, SAD provides a reasonable first guess for the electron density, but it performs less well for covalently bonded systems, such as molecules.

We note in passing that a similar concept is applied when using the Harris functional, where the energy is calculated non-self consistently from an input density that is constructed as the sum of the densities of the system constituents.[310] However, the choice of constituents is *a priori* arbitrary, and should be done such that they only interact weakly with each other.[310] For molecular crystals, it has been shown to be prudent to use the pre-converged electron density of individual molecules (rather than atoms), to obtain good results.[313]

The SAD method has two challenges that should be kept in mind. First, by construction, the dipole moment of a SAD guess is zero, as for all atoms the charges in the nuclei are exactly compensated by the charges in the spherically symmetric electron shells. Especially for interfaces with substantial charge transfer the electrostatic potential of the initial guess can, therefore, differ significantly from the potential of the final solution. Consequently, because the electrostatic potential affects the relative ordering of the orbitals (and their occupation), the initial guess is often far away from the converged situation. This can lead to a slow convergence or even a divergence of the SCF procedure. In practice, this is usually indicated by strong oscillations of the electron density and dipole moment between subsequent iterations. As we show below, this can often be mitigated *via* a strong damping of the SCF employing small mixing coefficients.

The second challenge arises for open-shell systems. Usually, by default the SAD density will be of singlet multiplicity with an equal amount of spin-up and spin-down electrons. *A priori*, this will also be the case for systems in which the real situation is different, like for magnetic systems or (weakly interacting) singly charged molecules, *i.e.* cations or anions, at interfaces. The main problem arising from that initialization is that for symmetry reason, singlets are always stationary points on the electronic potential energy surface (either true minima or saddle points).[314] Hence, different spin solution can only be obtained by preparing an initial guess with a net spin moment. Unfortunately, the SAD approach allows only limited control here. Since orbitals/bands cannot be directly addressed, the initial guess must be constructed by assigning an "initial" spin moment to each atom, which is not necessarily straightforward, even if the spin multiplicity of the unit cell is known. Thus, even when the initial guess has the correct multiplicity, it can be far from the correct solution. In fact, since semilocal

8154 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

Perspective

PCCP

functionals only have relatively small barriers between stable spin solutions (due to the lack of static correlation, as shown by Cohen *et al.*[314,315]), it is easily possible to converge to a solution with a total spin moment different from the desired one.

5.3 Density/wave function update algorithms

Once the initial wavefunction (and/or electron density) has been set up, the SCF algorithm solves the Kohn–Sham equations for the given charge distribution yielding a new wavefunction (and/or electron density). It turns out that if one uses the output directly as input for the next iteration, convergence is hard to reach, as SCF iterations can show highly oscillatory behavior or very slow convergence. Typically, the step size between consecutive SCF steps is restricted by only allowing the density to deviate by a certain amount from the previous step. To achieve faster convergence of the SCF, state-of-the-art programs sacrifice guaranteed convergence for performance using various acceleration algorithms. In this work, we will consider three of those, that are often used together: The density mixing method (with the specific example of the Pulay method),[316] Charge Preconditioning, and Level Broadening. All of these algorithms have the potential to yield efficient convergence, but require parameters that ought to be (somewhat) adapted for each system. Unfortunately, the optimal parameters differ significantly for metals, molecules, and interfaces, as will be described below.

5.3.1 Density mixing. The simplest of these strategies is linear mixing (also known as naïve mixing or underrelaxation), where a linear interpolation between the input guess and the output wavefunction/density is taken as new input:

$$n_{i+1} = n_i + \alpha \Delta n \tag{4}$$

Here n_i is the electron density of the *i*th iteration ($i + 1$ being the next iteration), Δn the difference between the input and the output of the present SCF step, and α the mixing parameter. This strategy bears the important advantage of guaranteed convergence, if very small mixing coefficients are employed. However, since small mixing parameters require a very large number of iterations, it is too costly to rely exclusively on this approach for interfaces. Still, it can be useful for systems that are hard to converge. If the initial guess is particularly bad (*e.g.*, when using SAD-densities for systems with very large interface dipoles), it is useful to employ linear mixing for the first few (5–10) iterations. This allows the SCF to reach a "harmonic" part of the solution space, where the more sophisticated approaches described below can take over.

To accelerate convergence, many codes replace linear mixing with various forms of nonlinear density mixing. Different approaches include the Broyden[317] and the Pulay scheme, which is also known as direct inversion in iterative subspace (DIIS) and which is the most prevalent convergence acceleration scheme.[316,318] While a full mathematical explanation of the DIIS scheme is beyond the scope of this contribution, the core idea is readily explained. In every iteration, the "residual", *i.e.* the deviation of the charge density from the (expected)

converged situation, is computed and stored. After a given number of iterations, the residuals from the previous iterations are combined in a coupled set of equations to determine the point on the electronic potential energy surface for which a zero "length" of the residual (the "null vector") is expected. This works best if the electronic structure is already close to the converged results, *i.e.* in the "harmonic" part of the energy surface. Conversely, when far away from the minimum, the error vectors may point to the wrong direction or change direction rapidly between consecutive steps making this approach potentially unstable.

Like in linear mixing, also in the Pulay scheme scaling is used to interpolate between the present wave function (or density) and the predicted optimal one. Pulay mixing, thus, contains two parameters, namely the number of iterations for which the residuals are stored, and the mixing parameter, in analogy to eqn (4). These define how many iterations should be stored and how much the difference between the present and the "expectedly correct" electron density should be damped.

As will be discussed in more detail below, for metals, where the SAD electron density is qualitatively close to the final electron density and the charge rearrangements within the SCF are mostly short-ranged, it is good to apply a small mixing parameter (< 0.2). For molecules, where the initial guess is worse, faster convergence is usually achieved by applying larger mixing parameters (> 0.4). Simulations of hybrid metal–organic interfaces, thus, face the challenge that the optimal values for the constituents differ significantly. This is amplified by the fact that many of these interfaces exhibit long-ranged charge-transfer between the substrate and the adsorbate. As this long-ranged charge transfer and its associated interface dipole shift the relative levels of the constituents, overshooting the charge-transfer in one iteration (which is then over-corrected in the next iteration) poses one of the major caveats for the SCF. Although this overshooting can be reduced using a charge-preconditioner (see below), in practice it is best to choose a relatively small mixing parameter (~ 0.2, close to metal value). Generally, this will make the calculations slow (*i.e.*, require many iterations), but decrease the likelihood that the SCF exhibits strong oscillations that prevent it from converging at all.

5.3.2 Charge preconditioning. Bulk metals are isotropically bonded with a highly homogenous electron density distribution. At the surface, however, the homogeneity is broken as electron density "spills out" into the vacuum, forming a surface dipole. In terms of SCF convergence for a metal slab, this means that electron density needs to be redistributed from the center of the slab to its surfaces. Despite the small values for charge mixing advised in the previous section, most charge mixing parameters tend to overshoot this long-ranged bulk-to-surface charge transfer in some SCF iteration, bringing too much electron density to the surface. In the next iteration, this is then (over)compensated, redistributing the electron density back into the bulk or possibly even to the opposite surface of the slab. The resulting oscillatory behavior of the SCF slows down or outright prevents convergence. This

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8155**
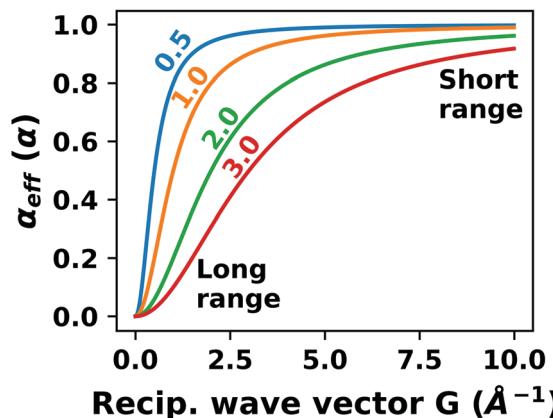
Fig. 15 Effective mixing parameter α_{eff} as a function of the reciprocal wave vector G for different values of q_0 when applying the Kerker preconditioner.

phenomenon is a variation of what is commonly known as "charge-sloshing".[319]

The problem of charge sloshing is routinely countered using so-called preconditioner algorithms. Multiple variants exist (*e.g.*, Kerker,[320] Broyden,[321] or Teter[322] preconditioners), that all rely on the same basic idea: The charge density difference (between the SCF input and output) is transformed into reciprocal space (*i.e.*, Fourier-transformed), and the low-frequency components, which correspond to long range charge density differences, are then damped or removed. This is most easily demonstrated on the Kerker preconditioner, which has the form

$$\alpha_{eff} = \hat{\alpha}\frac{G}{G + q_0} \tag{5}$$

Here, the effective mixing coefficient α_{eff} is given by α, the "global" mixing coefficient of the density mixing algorithm, G, the Fourier component (*i.e.*, wavelength of the charge-density difference in reciprocal space), and a parameter q_0. The impact of this parameter is illustrated in Fig. 15. Small values for q_0 result in a relatively unperturbed mixing parameter for short-distance charge rearrangements, and a rapid decay of the effective mixing for long-range rearrangements. In other words,

small values of q_0 efficiently suppress long-ranged charge-transfer compared to short-range changes. Larger values of q_0 lead to a smoother transition, *i.e.* they still emphasize short-ranged rearrangements over long-ranged changes, but not as pronounced. Note that at the same time, even in the short range the effective mixing parameter is reduced compared to its global value. Therefore, when choosing larger q_0, it is often effective to concomitantly increase α in order to retain an efficient convergence behavior.

Choosing or deriving an optimal value for q_0 is often challenging. For bulk materials, optimal values for the different preconditioner schemes can be deduced based on theoretical considerations. The situation is, however, more complicated when dealing with interfaces. The above-mentioned charge-sloshing results in charge rearrangements perpendicular to the surface (*i.e.*, in z-direction), while parallel to the surface (*i.e.*, in x- and y-direction), the slab behaves more like a bulk material. Clearly, the ideal situation would be to impose a different damping behavior in each direction. Unfortunately, most (if not all) implementations of preconditioners are isotropic. Moreover, the situation is further complicated by the fact that the magnitude of the bulk-to-surface charge transfer depends on the thickness of the slab. Depending on the implementation, even the size of the vacuum may play a numerical role. Finally, the ideal settings for the preconditioner are strongly, but non-trivially related to that of the Pulay mixer.

To demonstrate this issue – and to provide some general guidelines on how to choose parameters – we have tested the required iterations to converge the TCNE/Ag(100) interface (from the previous example, see Fig. 14) to $\Delta E = 10^{-4}$ eV using various combinations for the SCF settings (varying the Pulay scaling parameter between 0.05 and 0.6 and the settings for q_0 between 0.5 and 3.0 Å^{-1}). The results are shown in Fig. 16 for both the combined system and the isolated subsystems. As Fig. 16a shows, the molecule converges quite efficiently irrespective of the settings, but a large mixing parameter and a large value for the preconditioner are optimal. The Ag(100) metal slab, shown in Fig. 16b, is more sensitive to the chosen parameters (and requires more iterations) and shows the opposite behavior: Here, small values for mixer and preconditioner are beneficial.
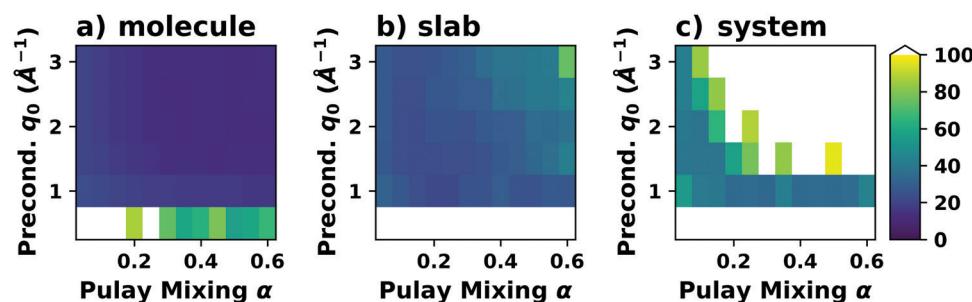
Fig. 16 Number of iterations required for (a) the TCNE molecule (nonperiodic) alone (b) the Ag(100) slab alone, and (c) the TCNE molecule on Ag(100) to converge to 10^{-4} eV using different settings for the Pulay mixer, α, and the Kerker preconditioner, q_0. White fields indicate that convergence was not reached within 100 iterations The calculations were done with FHI-aims using the PBE functional and a $3 \times 3 \times 1$ Γ-centered k-grid. Further details can be found at https://dx.doi.org/10.17172/NOMAD/2020.12.07-6.

8156 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

The joint TCNE/Ag(100) system is electronically the most complicated, and, as shown in Fig. 16c, also the one most sensitive to the chosen parameters. Indeed, for this system (which is rather prototypical for interfaces), we find that not all parameter combinations allow the system to converge at all (or at least within 100 iterations). Especially the combination of a large q_0 value for the Kerker preconditioner and large values for the Pulay mixing is very inefficient. Overall, we find that it follows the trends of the slab, *i.e.* mixer and preconditioner should have small values. Interestingly, combinations where either the preconditioner or the mixer have small values converge fast, and if either is small, the value of the other tends to matter less. However, (as also for the slab and the molecule), too small values for q_0 (here 0.5 Å^{-1}) prevent the system from converging at all, at least within the first 100 SCF steps.

In passing, we note that the energy obtained in all converged runs is highly reproducible. For the more than 200 different algorithmic settings for which the SCF converged, employing a convergence criterion of $\Delta E = 10^{-4}$ eV lead to a standard deviation in the final energies of only 2×10^{-5} eV.

5.3.3 Level broadening and occupation. Another class of commonly applied algorithms to facilitate convergence are so called "level broadening" or smearing schemes. Those pursue a two-fold purpose: All (periodic) observables need to be calculated by integrating over the 1st Brillouin zone. This can be either done using broadening methods (*i.e.*, replacing the step-function for the occupation with a different function) or by interpolation, *e.g.*, using the Tetrahedron method.[323] On the one hand, artificially broadening the energy levels allows for interpolations in k-space, thus reducing the number of k-points that need to be calculated to obtain the correct electronic structure (and, hence, the computational effort). On the other hand, for metallic systems, broadening schemes also reduce the so-called "level-switching" problem, which adversely affects the SCF convergence.

This problem occurs because at zero temperature, the Fermi function is reduced to a step function with the step at the Fermi-energy; *i.e.* all states below the Fermi-energy are fully occupied, while all states above are empty. For metallic systems, where the Fermi-energy cuts through a band, small changes of the electron density in one iteration can easily change the potential such that some eigenstates are shifted below the Fermi-energy, while others are shifted above it. In the next iteration, the electron density is adjusted accordingly, which again leads to shifts of some eigenstates from above to below the Fermi energy and *vice versa*. This can lead to notable charge rearrangements throughout the SCF, which impede convergence. The core concept of level broadening (or "smearing") schemes is to replace the zero-Kelvin Fermi-function with a different occupation function that exhibits a smoother transition between completely empty and fully occupied states and allows for fractionally occupied orbitals. All these replacement functions contain a free parameter σ, that governs how fast the occupation drops from full to empty around the Fermi energy.

In passing, we note that there are two different ways to look at these schemes: either, each state is at a discrete energy and (partially) occupied according to the replacement function (*e.g.*, a Fermi–Dirac-distribution function, see below), or each state is broadened in energy (using the derivative of the replacement function) and occupied using a step function. Both viewpoints are mathematically equivalent. For convenience, most codes work internally with the former, but plot the density of states using the latter approach.

The easiest – and most straightforward – way would be to employ the Fermi–Dirac occupation function with a finite temperature. Although this does solve the "level switching problem", it shows very unfavorable behavior for the interpolation within reciprocal space, *i.e.* it does not efficiently reduce the required number of k-points. Therefore, Fermi–Dirac distributions are usually replaced with a different functional form. Fig. 17 compares the shapes of some of these occupation functions for a given value of σ. For a given state it also illustrates the corresponding level broadenings, which yield identical occupations when combined with a step function. The most commonly used occupation functions are Methfessel-Paxton[324] functions, which have the following general form:

$$ f(x) = \frac{1}{2}\mathrm{erfc}\left(\frac{x}{\sigma}\right) + \sum_{m}^{n} A_m H_{2m+1} \times \frac{x}{\sigma}\exp\left(\frac{\sigma}{x}\right)^2 \qquad (6) $$
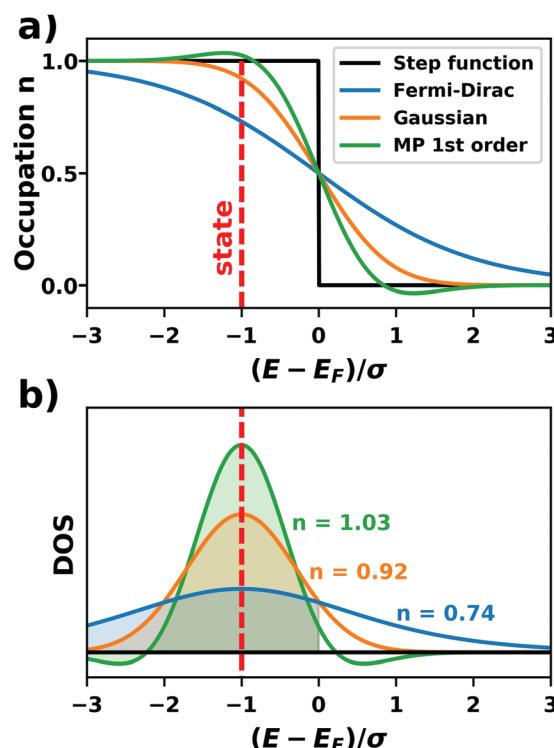
Fig. 17 Relationship between employing level broadening and using occupation functions. Panel (a) shows the occupation number for a state at an energy 1σ below the Fermi energy (red dashed line) when using different occupation functions. Panel b shows the same state broadened by a corresponding smearing function (using a broadening parameter of 1σ). The integral over the state up to the Fermi energy is the occupation, which is indicated in the plot. Note that the smearing function is the derivative of the occupation function.

Here, the first term is the error function, *i.e.* the integral over a Gaussian function, while the second term is a sum over Hermite polynomials up to *n*'th order. 1st-order Methfessel–Paxton functions are known to show favorable convergence (with respect to the number of *k*-points) for metals. However, because they allow for negative occupation numbers, they can easily lead to unphysical results for semiconductors and particularly for molecules. For interfaces that are not necessarily fully metallic, it is, thus, advisable to use a Gaussian smearing (which corresponds to a zeroth-order Methfessel–Paxton occupation function).

As mentioned above, the main advantage of using smearing methods is that the energy in the SCF procedure converges significantly faster with the number of *k*-points, *i.e.* it is possible to perform the same calculation with fewer *k*-points, and, thus, less computational effort. However, the disadvantage is that the total energy in the SCF no longer represents the "correct" total energy at $T = 0$ K (nor at any other temperature); even for a Fermi–Dirac type occupation, it would only be an ensemble-average approximation. Rather, the SCF now calculates an effective "free energy". As shown for the example of a 5-layer Cu(111) slab (primitive unit cell, calculated using VASP) in Fig. 18, the free energy depends on the chosen value of σ, even for a converged *k*-point grid. Moreover, the free energy is no longer variational (*i.e.*, it can also increase during the SCF procedure). Fortunately, once the SCF procedure is converged, it is still possible to obtain the correct electronic energy by extrapolating to $\sigma = 0$.[299,325,326] The corresponding energy is also the one that should be taken for all postprocessing. As shown in Fig. 18, the result after extrapolation as expected still varies significantly with the *k*-point grid but has become rather independent of σ.


Fig. 18 Convergence of the free energy (dashed lines, labeled as "not corrected") and the energy extrapolated to $\sigma = 0$ (solid lines, "corrected") as a function of the number of *k*-points for a 5 layer Cu(111) slab using an evenly spaced Γ-centered *k*-point grid and obtained using the Gaussian broadening scheme with different smearing parameters. Note that the dashed green line coincides with the full green line. All data were calculated with VASP using the PBE functional and the default value for the cutoff of 294.4 eV. Further computational details and the full results can be found in the NOMAD database https://dx.doi.org/10.17172/NOMAD/2020.12.07-7.

Still, one needs to be careful when performing geometry optimizations. During the SCF, the forces acting on the atoms, *i.e.* the derivative of the energy with respect to nuclear displacements, are obtained from the free-energy, not the back-extrapolated zero-K energy. This may result in the geometry optimization requiring more steps (*i.e.*, being slow), or even yielding incorrect results. For a given calculation, the best σ is, thus, a balance of being large enough to require as few *k* points as possible, while not introducing too large errors in the total energy. Generally, values between 0.01 and 0.2 eV are known to produce (relatively) sensible results for the energy and the geometry.

To benefit maximally from the smearing interpolation while avoiding artificial results, the most prudent approach is to perform initial calculations (including geometry optimizations) with a relatively large broadening and few *k*-points – although not too few, as this may lead to serious artefacts, as we show below. The final results should then be verified using a smaller broadening (0.01 eV), with an appropriately higher *k*-point density.

5.4 Local geometry optimization

In this section, we will focus on strategies for efficiently finding the structure of the (local) energetic minimum. Local in this context means that we will only consider "conventional" geometry optimization schemes that find the minimum structure in a specific search basin of the potential energy surface which is determined by the initial guess structure. For a discussion of global structure search algorithms the readers are referred to the recent book by Oganov.[327] At first, we will look at sensible thresholds for the geometry optimization for a simplified potential energy surface. We will then proceed to a more complex system. In particular, we will demonstrate why the choice of the initial guess for the Hesse matrix is imperative for the performance (and even the qualitative outcome) of a geometry optimization and look at error bars for adsorption heights. Finally, we will discuss strategies to make this typically rather time-consuming task as efficient as possible.

In a geometry optimization, atomic displacements are performed in consecutive iterations until the forces acting on the atoms fall below a certain threshold. Generally, it is difficult to recommend good, generally valid thresholds (since this depends on the system). Nevertheless, it is instructive to consider a prototypical example, namely the adsorption of PTCDA on Ag(111). To simplify the discussion, we will not describe an actual geometry optimization, but discuss a simplified potential energy surface where the molecule was kept planar.[282] Fig. 19 shows the adsorption energy as a function of the adsorption distance together with its derivative in *z*-direction, *i.e.* the force pulling/pushing the PTCDA molecules towards/away from the surface. The displayed data were calculated with PBE+vdW$^{\text{surf}}$[161] and were taken from Hörmann *et al.*[282] (where also further information on the calculations can be found). Notably, the forces are below 10^{-1} eV Å^{-1} for essentially all displayed adsorption heights, *i.e.* stopping the geometry optimization when the (maximum) force per atom falls

8158 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

Fig. 19 (a) Total energy (relative to the minimum) of PTCDA on Ag(111) (calculated with PBE+vdW^{surf} for a single molecule in a 6×6 unit cell, as described in detail in Hörmann et al.)[282] as a function of its distance (relative to the equilibrium distance). (b) Force per atom acting on the PTCDA molecule in z-direction.

below 10^{-1} eV Å^{-1} would yield an uncertainty of about 0.10 Å on the adsorption height. Optimizations done with a threshold of 10^{-2} eV Å^{-1} (which are probably the majority of calculations in the literature at present) would here result in an intrinsic uncertainty of approx. 0.06 Å. This means that it would still not be appropriate to report adsorption distances with two or more significant figures behind the decimal point when using this convergence criterion. To obtain an accuracy of 0.01 Å, which is at or below the lowest experimental accuracies that can be achieved with modern structure characterization methods such as X-ray standing wave measurements,[328] the remaining forces need to fall below 3×10^{-3} eV Å^{-1} (red region). Naturally, more accurate results can be obtained by choosing even tighter criteria, but this would also incur a dramatic increase in the number of geometry optimization steps required to reach this threshold.

It is important to emphasize that the connection between force threshold and accuracy of the geometry depends sensitively on the system and the method, *i.e.* the xc-functional and the employed van der Waals correction scheme,[176,282] as they affect the corrugation and steepness of the energy landscape. The example here, PTCDA/Ag(111), is primarily bonded through dispersion, which shows a relatively shallow minimum. For systems which exhibit stronger molecule-surface interactions, a deeper and steeper minimum would be expected, *i.e.* smaller displacements could lead to larger forces.

To get a first idea about the intrinsic uncertainty of the adsorption height for a given force threshold, one can, in

principle, start two different geometry optimizations with one starting guess notably above and the other notably below the expected adsorption height. However, when doing so one needs to be aware that between two geometry optimization steps, the molecule can move over substantial distances. As we discuss below for a different example, this is often > 0.1 Å, *i.e.* similarly large as the basin itself. Therefore, at which point in the basin the optimization ends when employing this approach is, at least to some extent, arbitrary (*i.e.*, both optimizations could coincidentally end at very similar geometries or at the opposite boundaries of the basin). Thus, to get a reasonable idea about the associated error bar, one either needs to perform multiple geometry optimizations (and analyze the results statistically), or reduce the maximum distance each atom can move in one step. Alternatively, one can manually scan particularly sensitive geometrical parameters (like the adsorption distance in the above example) and then determine the distance at which the energy becomes a minimum from a fit (in analogy to performing a Birch–Murnaghan fit,[329] that is used when calculating the optimum lattice constant in a bulk optimization).

Most practically relevant optimization schemes start by calculating the energy gradients with respect to displacements of the nuclei, *i.e.* the forces acting on each atom. The most straightforward method for geometry optimization would be to follow these gradients, *i.e.* move the atoms in the direction of these forces, until the gradients vanish. Such gradient-based methods, which indeed exist in multiple variants (steepest descent, conjugate gradient, *etc.*), are conceptually similar to linear mixing for the SCF inasmuch as they are guaranteed to converge, but are also computationally inefficient. Especially near the energetic minimum, gradient descent methods are prone to show oscillations, as the optimization will primarily follow "hard" degrees of freedom (where small displacements lead to large energy changes) rather than "soft" degrees of freedom (where large displacements have a modest impact on the energy). As a result, gradient-based methods produce good energies reasonably quickly, but may yield geometries where parts of the molecule are relatively far away from their correct positions.

Most modern codes instead rely on quasi-Newton optimization algorithms.[330–334] In quasi-Newton methods a harmonic potential energy surface is assumed. This implies that the energy as function of the nuclear coordinates $E(Q)$ can be described by a quadratic expansion around the energetic minimum (E_0):

$$E(Q) \approx E_0 + \sum_i \frac{\mathrm{d}E}{\mathrm{d}Q_i} Q_i + \sum_i \sum_j \frac{1}{2} \frac{\mathrm{d}^2 E}{\mathrm{d}Q_i \mathrm{d}Q_j} Q_i Q_j + \dots \quad (7)$$

The first derivative of the energy with respect to the nuclear coordinates are the forces acting on the nuclei, while the second derivative is commonly termed Hesse matrix (or Hessian). In principle, if the assumption of a harmonic potential energy is valid and if the Hessian is known, the equation above can be inverted, and the minimum geometry is obtained within a single update. In practice, usually neither of the two conditions actually applies. In passing, we note that quasi-Newton schemes only

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8159**

update the inverse Hessian, as this is mathematically more efficient.

The assumption of a harmonic potential energy surface is only valid, when the structure is already close to the energetic minimum. This encompasses a relatively small fraction of the potential energy surface: Typically, the quadratic expansion is a reasonable model while atoms are within 0.05–0.20 Å of their equilibrium positions. Because the performance of quasi-Newton methods quickly deteriorates outside this region, for interfaces it is generally a good idea to put some effort into generating the initial (guess) structure, and to consider carefully, what the preferred adsorption sites and the most likely orientations of the molecule would be.

As a related issue, the harmonicity of a potential energy surface strongly depends on the coordinate system that is chosen. As so often for interfaces, the "optimal" choice differs for its chemically different constituents. For isotropically bonded materials with many bonding partners (metals and ionic materials), the movements of individual atoms are only weakly coupled. There, a simple Cartesian coordinate system (moving each atom in x,y,z) is sensible and effective. Conversely, for covalently bonded systems, such as molecules, the coupling is strong. Here, coordinate transformations to "internal coordinates" (which consist of bond lengths, angles, and torsions) are known to lead to much faster convergence.[247,335,336] In practice, many band structure packages still perform their geometry optimization exclusively in Cartesian (or fractional) coordinates, but several "wrappers" exist that allow to circumvent this limitation.[337–340] Although many programs are not very explicit about the choice of the coordinate system, it is strongly advised that the user finds out and takes the necessary caution when interpreting the results of a geometry optimization.

The more relevant challenge for quasi-Newton optimization schemes is that the Hessian (or its inverse) is not known *a priori*. While it could be calculated for the guess geometry, the computational effort for this is enormous: If calculated numerically, it requires $6N$ evaluations of the forces on all atoms (where N is the number of unconstrained atoms, typically a few 10 to 100). This is as expensive as performing $6N$ geometry optimization steps. Even if "analytic" frequencies were used (employing density functional perturbation theory, which is available in a few code packages), this would still be computationally inefficient (despite a significant reduction of the necessary effort).

Therefore, the typically applied strategy is to start from an estimated Hessian as initial guess and to update it as the geometry optimization progresses. Following the common theme of interfaces, the optimal guess Hessians differs for different material classes. For the reasons discussed already in the context of choosing the coordinate system, isotropically bonded materials (metals/insulators) are commonly initialized using a scaled diagonal Hessian. This basically corresponds to assuming springs that bind each atom individually to its equilibrium position, neglecting any coupling between the atoms. Conversely, for covalently bonded systems, it is ideal to start from a Hessian in internal coordinates (bond lengths, angles, and dihedrals) and to choose the corresponding force

constants based on pre-tabulated values (*i.e.*, based on a force field).[341,342]

To illustrate the importance of the guess Hessian, we have calculated a low coverage monolayer of cyano-biphenylthiole on Au(111). As initial guess, we chose an almost upright, slightly tilted molecule, with plenty of space to each side, as shown in the central panel of Fig. 20. The optimization is performed in Cartesian coordinates. From a physics point of view we know that, triggered by van der Waals interactions, the molecule will "fall over",[147] *i.e.*, we expect a large change in the tilt angle during the geometry optimization. Thus, the initial structure is, geometrically speaking, quite far away from the actual minimum.

For our example we used two different initial guesses for the Hesse matrices, a scaled diagonal guess (which is the default in many codes) and a Lindh guess,[341] which generates the guess based on pre-determined parameters for bond-lengths, angles, and dihedrals. In both cases, the geometry optimization was stopped when the remaining force on each atom fell below 0.01 eV Å^{-1}, which is a typical, rather tight threshold employed in geometry optimizations. For the present example, the geometry optimization required more than 100 steps for both types of Hesse matrix initializations.

The resulting geometry for the diagonal Hessian is shown in the right panel of Fig. 20a. Only some small, short-ranged rearrangements of the atom positions have taken place (causing, *e.g.*, an increase of the inter-ring twist). However, the tilt angle of the molecular axis relative to the surface normal (identified before as the degree-of-freedom of primary importance) remains essentially unchanged. During the optimization, the total energy changes by slightly more than 1 eV relative to the energy of the initial geometry. Conversely, performing the same optimization using the Lindh-guess leads to much larger changes. Here, the energy decreased by almost 3 eV.

Geometrically, the most striking difference between the two geometry optimizations is the tilt angle relative to the surface. It is visibly different between the diagonal guess, where the molecule remains upright, and the Lindh guess (left part of Fig. 20a), where both rings are "in contact" with the Au surface. This is also reflected in the interface dipole $\Delta\Phi$, which, for the diagonal guess, remains at approx. 0.75 eV, but even goes to negative values for the optimization using the Lindh guess (see Fig. 20d).

Since the two geometry optimizations exited gracefully, the two resulting geometries could easily be interpreted as being two separate minima. However, it is unlikely that this is indeed the case. Had we continued the optimization to an even tighter threshold, even with the diagonal guess the optimization would eventually find the same minimum structure – however, very inefficiently, making the associated computational effort too large to be tractable. We can backtrack the Lindh optimization and ask what would have happened, had we used a lighter threshold: Had we chosen to run the optimization until the remaining forces fell below 5×10^{-2} eV Å^{-1} (which is, *per se*, not unreasonable for large systems), also the optimization with the Lindh Hesse matrix would have finished with a de-facto upright standing molecule.

8160 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021
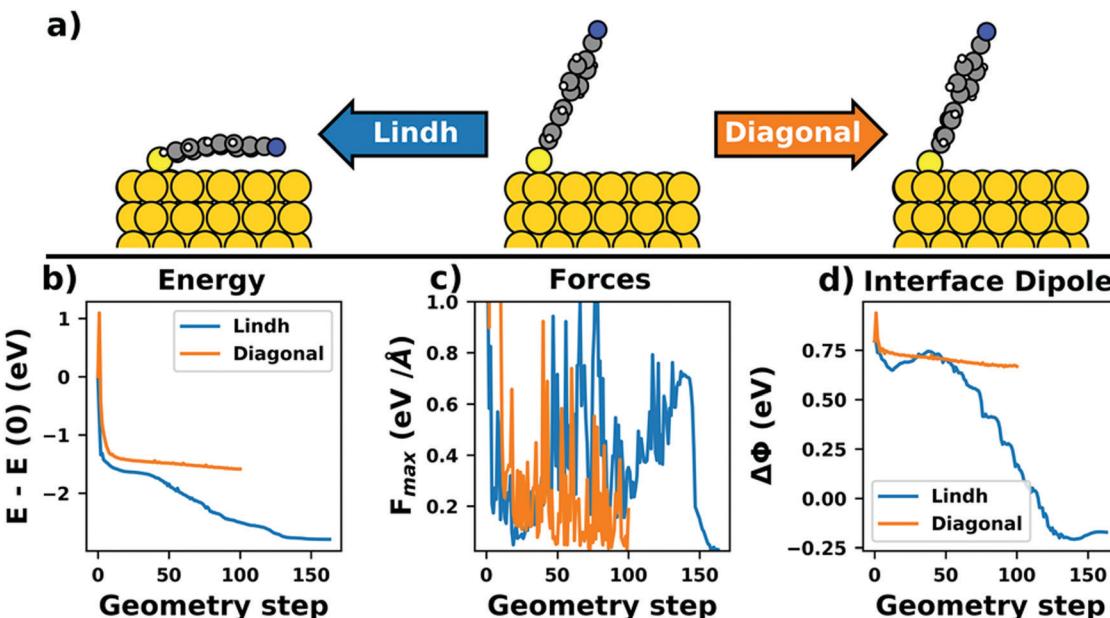
Fig. 20 Geometries of a cyano-biphenylthiole molecule on Au(111), calculated in a $5 \times 2\sqrt{(3)}$ unit cell with FHI-aims using the PBE+vdW^{surf} method and a $3 \times 2 \times 1$ Γ-centered k-point grid. A three-layer metal slab was used. All atoms of the molecule and the topmost layer of the slab were allowed to relax until the remaining forces on all atoms fell below 0.01 eV Å^{-1}. The calculations can be found on the NOMAD database. [https://dx.doi.org/10.17172/NOMAD/2020.12.07-9]. (a) Shows the starting geometry (center) and the final result after the optimization using a diagonal guess for the Hesse matrix (right) or using a Lindh guess (left). (b) Shows the evolution of the total energy relative to the energy of the starting point, for the two geometry optimizations. (c) Shows the remaining maximum force component F at each iteration step, (d) gives the interface dipole $\Delta\Phi$.

The present and the previous example illustrate two important messages: (1) Even when a calculation converges nicely, it is important to critically second-guess the resulting geometry. (2) It is important to use sufficiently tight thresholds. The "ideal" value of the threshold depends on the physics of the system, the required accuracy of the geometry, and also the employed geometry optimization scheme and the employed parameters (such as the values for the guess Hessian in the present example).

It is also instructive to briefly consider why the optimization proceeds so slowly here (convergence only after more than 100 steps), even when using the Lindh-matrix as initialization. Three factors play a major role: First, as mentioned above, the interactions that primarily cause the molecule to tilt are van der Waals interactions. These are, however, not contained in the Lindh parameterization. Secondly, because van der Waals forces decay very quickly with distance, they "pull" more efficiently on the bottom ring than on the top ring. This reduces the overall tendency of the molecule to change its tilt. The major factor is, however, the maximum step size. All geometry optimizations employ a maximum distance by which a given atom can move between subsequent steps (typically about a tenth of an Å; for Fig. 21, the default value of 0.2 Å was chosen), in order to avoid creating unphysical structures with broken bonds. Here, for the molecule to fall over, the nitrogen atom has to move by several Å. Even in the most ideal case, the step size limit (for which, again, the default in the present example was clearly suboptimal) requires several dozens of steps to allow reaching the equilibrium structure. This problem can
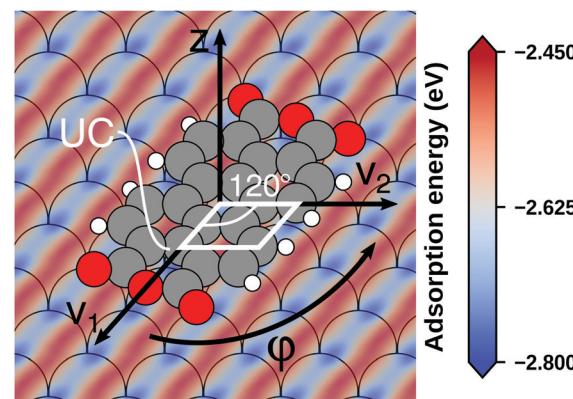
Fig. 21 Potential energy surface for the translation of a rigid PTCDA molecule 2.9 Å above the Ag(111) surface, calculated for a 6×6 supercell using PBE+vdW^{surf}, as obtained by Hörmann et al.[282] v1 and v2 denote the primitive lattice vectors of the substrate, z the adsorption height and φ the rotation angle of the molecule relative to the substrate. The potential energy surface shows a minimum on the Ag-bridge site, a low-energy saddle point along v1 and a high-energy saddle point along v2.

be mitigated by increasing the corresponding parameter in the input, or by using dedicated optimization tools (e.g., GADGET[337]).

The efficiency of quasi-Newton schemes for interfaces is, however, fundamentally limited by the physics of the interface. In quasi-Newton optimization schemes, the effective step size is always determined by the smallest eigenvalue of the inverse Hessian (the stiffest mode, since this leads to the largest energy change),

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8161**

while the maximum move occurs along the softest mode (described by the lowest eigenvalue of the Hessian). In other words, the convergence speed depends on the ratio of highest to lowest eigenvalue of the Hessian (the so-called "condition number"):[337] the higher the ratio, the slower the convergence. Although coordinate transformations can favorably shift the condition number[337] (*e.g.*, by going from Cartesian to internal coordinates), at interfaces, it will remain intrinsically high. The large substrate cell contains some very soft modes originating from a backfolding of the acoustic modes of the primitive substrate unit cell. Also intermolecular interactions are very soft, and so is the (often van der Waals driven) interaction between substrate and adsorbate. Therefore, the maximum move occurs along these modes. Conversely, covalent bonds are very stiff (particularly double and triple bonds with the CN-bond in the above example as the prototypical "worst case" scenario), determining the step size. Because this broad distribution of modes is a fundamental property of interfaces, which also does not change when using a different coordinate system, even optimizations that work in internal coordinates would be slow here.

It is important to stress that the described behavior is quite general for interfaces, and not limited to the somewhat extreme biphenylthiole example from above, which has been chosen also because it illustrates the problems particularly clearly. Other examples where convergence is slow include cases in which the molecule ought to rotate on the surface to reach a minimum structure. There, the necessary significant changes in the Hessian matrix are not easily captured by the update algorithm when working in Cartesian coordinates, resulting in very slow convergence. A similar problem occurs when forces on individual atoms are small, but all point in the same direction, *i.e.* when the molecule translates.

A further peculiarity of quasi-Newton optimizers to keep in mind is that, unlike most other geometry optimization schemes, they are not guaranteed to follow the direction of the gradient downward in energy. Rather, the optimization will proceed to the closest stationary point, which can also be a saddle point, if the present guess geometry is sufficiently close to it (*i.e.*, in its "harmonic" region).[292] For interfaces, this is mostly an issue for degrees of freedom that move unbounded between symmetry equivalent positions, such as translation across the surface or rotation of the adsorbate, as shown in Fig. 21. For these motions equivalent minima are separated by a saddle point. For reasonably complicated adsorbates, there is little chemical intuition where the minimum should be, making it likely for novice (and even veteran) computational scientists to start the geometry optimization near a saddle point.

To reduce the chances of landing on a saddle point, it is possible to pre-optimize the geometry with a different method, such as steepest descent. This is guaranteed to follow the direction of the gradient. It is furthermore a robust method for pre-optimizations when the geometry is far away from the harmonic region of the potential energy surface. Therefore, Conjugate Gradient preoptimizations are especially advisable

when there is little intuition about the potential energy surface and the positions of the energetic minima.

We note that also other geometry optimization algorithms exist that sometimes may be superior to quasi-Newton methods. Since these are not very widespread (yet), here we would like to mention only two: damped molecular dynamics (DMD) and Bayesian Optimization.[166,343] Both have the advantage of not relying on a harmonic expansion, *i.e.* they should perform well also far away from the minimum. In our experience, DMD at interfaces suffers from the simultaneous presence of soft and hard modes: Hard modes necessitate small timesteps, since otherwise the molecule breaks up (which happens frequently, *e.g.*, when CN groups are involved), but too small timesteps make it harder for the molecule to move efficiently along soft modes. Overall, while DMD with optimized parameters may show a faster convergence than quasi-Newton methods ever will, finding these parameters can be tricky and (unless good values are known beforehand) is, in our experience, hardly worth the effort.

Bayesian Optimization is a relatively new method, for which not much experience has been accumulated so far in the context of surface science. Generally, Bayesian Optimization techniques work well when there are only few degrees of freedom, but the computational effort becomes too large for larger systems.[344] Explained in a very simplified way, Bayesian Optimization relies on calculating a few selected geometries and constructing a (conditional) probability distribution for the energy of other geometries, *i.e.* effectively interpolating between them. In contrast to the other optimization methods, it can, therefore, provide a general (at least qualitative) overview of the potential energy surface, including the positions and locations of the minima and saddle points. We presently recommend to employ Bayesian Optimization following the strategy described by the Rinke group:[349] Keep the internal structure of the adsorbate (bond lengths, angles, *etc.*) and the substrate fixed, only moving and rotating the molecule as a whole across the surface. This procedure allows to quickly identify the attraction basins of different local minima on the surface. Moreover, it provides a good overview of (but certainly no exact result for) the correct values for these soft degrees of freedom. The obtained geometries are then used as starting points for quasi-Newton optimizations, which then find the minimum structures (usually) very efficiently. In this case, the Bayesian Optimization replaces the above-mentioned preoptimization with conjugate gradient and additionally allows a more systematic screening of the entire potential energy surface, revealing several local minimum structures.

Besides the different geometry optimization strategies mentioned here, also other settings in the calculation can influence the optimized geometry. We illustrate this for the example of TCNE on Cu(111). There, we performed various geometry optimizations starting with a flat lying molecule from the same, randomly (*i.e.*, deliberately, not intentionally) chosen position while varying (i) van der Waals corrections, (ii) *k*-points, (iii) substrate relaxation strategies (fixing the whole substrate *vs.* optimizing the two uppermost Cu layers). Fig. 22 illustrates the

8162 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

Initial geometry
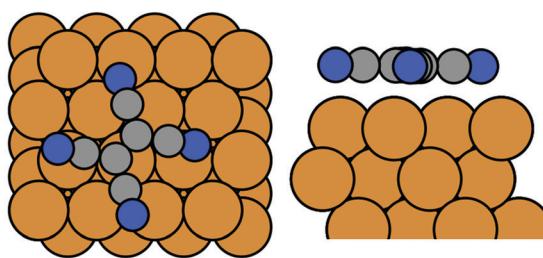
Optimization results
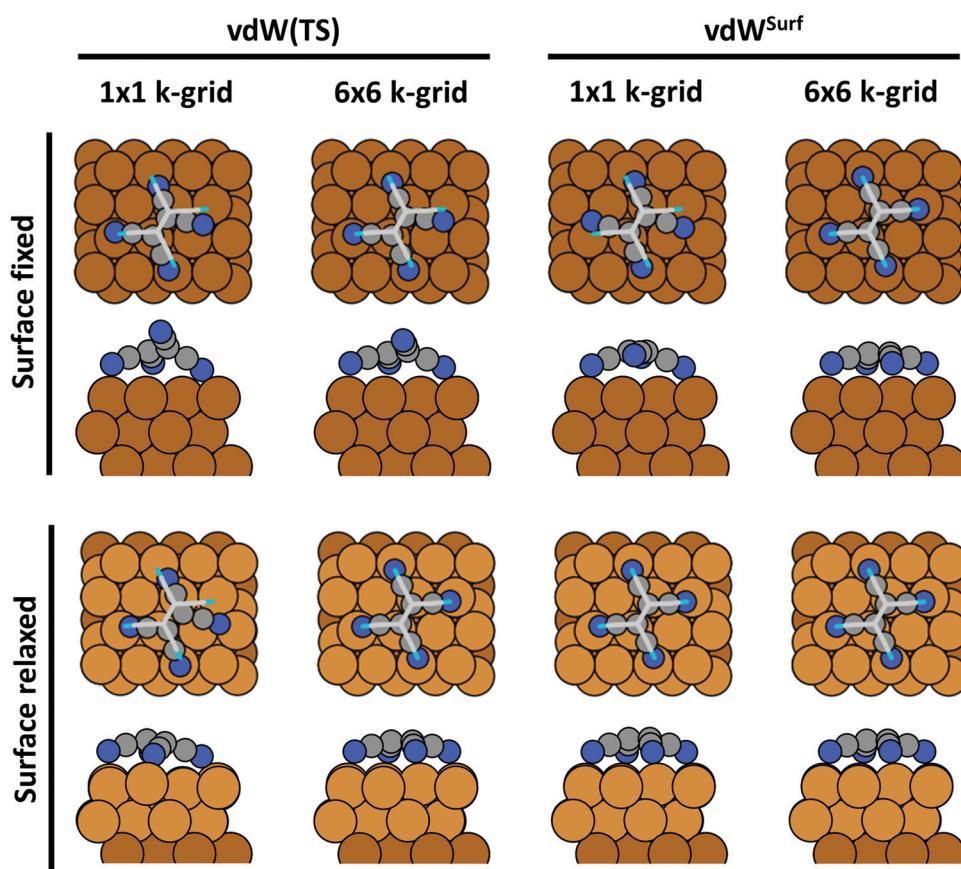
Fig. 22 Geometry optimizations of a flat lying TCNE molecule on Cu(111) with varying computational strategies: changing k-point densities, different van der Waals corrections and fixing the full substrate *vs.* relaxing the two uppermost layers. The grey wireframes indicate the adsorption position obtained with the most elaborate strategy (relaxed surface, 6×6 k-points and vdW^{surf} correction). Dark Cu atoms were fixed, lighter Cu atoms were allowed to relax during the geometry optimizations. All calculations were done with FHI-aims. Additional details can be found in https://dx.doi.org/10.17172/NOMAD/2020.12.10-1.

geometry optimization results for all combinations of settings. When fixing the substrate and using a vdW correction without surface reparameterization, the TCNE molecule only binds with two C≡N groups while the others are bent away from the surface. Introducing a surface-parametrized vdW correction partly prevents this behavior due to the now stronger adsorbate-surface interaction, but there are still differences between the optimization with a single k-point only and a well

converged k-grid (different adsorption site and bending). Optimizing the uppermost substrate layers makes the situation more reproducible, at least in the considered case. Here the surface parametrized vdW correction leads to identical results independent of k-point density and also without surface parametrization the ''correct'' adsorption geometry is found with the 6×6 k-point grid. This shows that for reliable results, not only the geometry optimization algorithm must be chosen

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8163**

suitably, but also the other methodological choices (discussed earlier in this manuscript) need to be chosen appropriately.

Before we conclude the section on geometry optimization, we want to briefly comment on the possibility of optimizing the geometry at a given level of theory, while then performing a single-point calculation on that geometry at a higher level of theory to obtain a supposedly more accurate electronic structure of the interface. Due to the high cost of geometry optimizations, this is a tempting scheme, which is frequently applied in quantum chemistry. For interfaces, this usually means that the geometry is obtained using a dispersion-corrected semilocal functional (e.g., PBE), and then the electronic structure is computed with a more expensive hybrid functional (e.g., PBE0). In most cases, this approach works well and is very useful when trying to gain a more quantitative insight. However, because electronic and geometric structure are strongly coupled, this approach bears the risk of running into serious artefacts, especially when the two methods predict qualitatively different electronic structures of the interface. Obvious manifestations of qualitatively different solutions occur, when the application of hybrid functionals leads to a strong localization of the electrons,[13] when the order of occupied and unoccupied states swap,[43,103] or when different spin-configurations are obtained by the two methods.[308] A less obvious, but nonetheless salient difference can occur when modelling the adsorption of flat-lying conjugated organic molecules undergoing significant charge transfer with the substrate. This is demonstrated in Fig. 23 for the adsorption of tetrafluorobenzoquinone (TFBQ) on Cu(111).[237]

Here, calculating the projected density of states for the adsorbed molecule using PBE+vdW^{surf} for both the geometry and the electronic structure (shown in Fig. 23b) shows a partially filled molecular feature at the Fermi energy, i.e., a metallic character of the adsorbate.[345] Re-calculating the electronic structure at a higher rung, i.e. with the hybrid functional PBE0 (Fig. 23d) yields an electronic structure that is similar inasmuch as it still contains a partially filled adsorbate-derived feature at the Fermi-energy. The situation is still qualitatively different, since now the projected density of states exhibits a clear double peak structure (indicated by arrows). Following the Newns–Anderson model[346] (and, in fact, much of established chemistry), such double-peak structures are characteristic for the formation of bonding and antibonding states due to strong chemisorption, i.e. the formation of a new, partially covalent bond.

However, when also optimizing the geometry at the PBE0+vdW^{surf} level, the density of states near the Fermi-energy vanishes, as does the double peak structure (see Fig. 23c). In other words, the interface chemistry is completely different: The adsorbate is no longer metallic, i.e. no longer exhibits a fractionally filled state, the LUMO-peak moves to lower energies and loses its double-peak structure, indicating that the bond between substrate and adsorbate is predominantly of ionic, not covalent, character. This behavior is not a peculiarity of the TFBQ molecule. Rather, it originates from the fact that semilocal and hybrid functionals yield notably different bond lengths for single and double bonds.[347] Consequently, when performing a hybrid functional
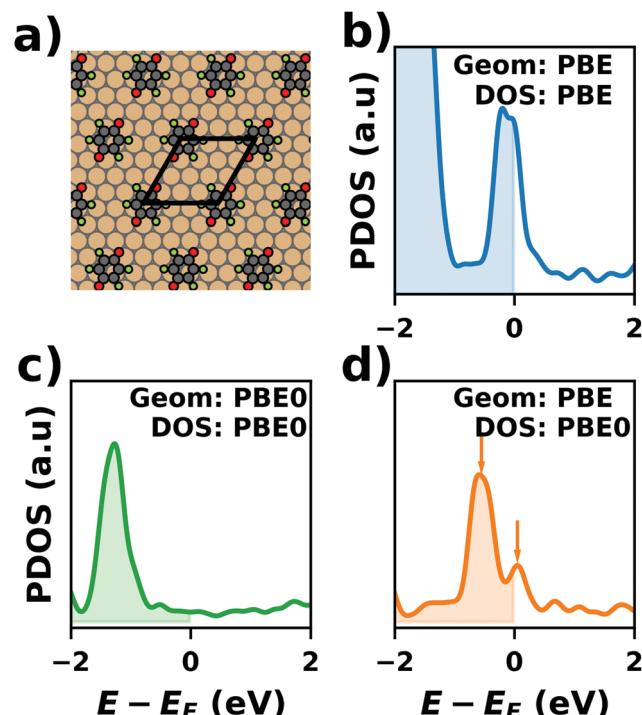
Fig. 23 Projected densities of states (PDOS) of tetrafluorobenzoquinone in a $p(4 \times 4)$Cu(111) cell projected onto the adsorbate layer. (a) Is a representation of the geometry in a 2×2 supercell. (b) Shows the result for the projected DOS obtained with PBE after also optimizing the geometry with PBE. For the results in (d) the same geometry was used, but the electronic structure of the interface was recalculated using PBE0. For (c) PBE0 was used for re-optimizing the geometry and for calculating the electronic structure of the interface. All calculations were done using FHI-aims and a $6 \times 6 \, \Gamma$-centered k-point grid. All geometry optimizations were performed including the vdW^{surf} correction. For further details see Wruss et al.[348] and the calculations in the NOMOAD database at 10.17172/NOMAD/2020.12.11-2.

calculation with a semilocal geometry, the individual bonds are out of their equilibrium, which makes the system too reactive, triggering a tendency to form new, covalent bonds.

6 Summary

In this work, we discuss various relevant technical, mathematical, physical, and chemical aspects that need to be considered when performing state-of-the-art first-principles simulations on hybrid inorganic–organic interfaces. Those interfaces pose a significant challenge, since they display strongly varied physics: they contain covalent, metallic, ionic, hydrogen and van der Waals bonds, often all at the same time. Charge transfer, localization, and the emergence of collective electrostatic effects that shift the relative level alignment of substrate and adsorbate complicate the electronic structure. Due to the different bond types, interfaces contain very hard and very soft degrees of freedom. All of these effects need to be accounted for accurately. This is often a considerable challenge due to the large size of the systems used to model such interfaces, which makes the simulations computationally very demanding. Therefore, we discuss different methods to model hybrid

8164 | Phys. Chem. Chem. Phys., 2021, 23, 8132–8180

This journal is © the Owner Societies 2021

Table 1 Suggested "best practice" rules for interface simulations

---

Best practice rules for interface simulations

The structural model
• Open *versus* periodic boundary conditions: cluster calculations with proper embedding and slab calculations (with appropriate supercells) give conceptually equivalent results when properly converged. Clusters can be more efficient when modelling inherently non-periodic or diluted effects (*e.g.*, defects or excitations). Slab calculations are typically easier to converge.
• When using periodic boundary conditions, employ a dipole correction or a Poisson solver for mixed boundary conditions.
• Make sure the position of the vacuum level is further away from the ends of the slab than the separation between the dipoles (*i.e.*, more than the length of the lattice vectors in x and y)
• When calculating charged unit cells, use a charge compensation scheme (CREST, metallic boundary conditions, generalized dipole correction, *etc.*)
• Periodic boundary conditions often require stretching or compressing the unit cells of adsorbate or substrate to obtain a common (commensurate) unit cell. Test carefully that this scaling does not affect the electronic properties.
• There are often many local minima for the adsorption of an organic molecule on a surface. Explore the potential energy surface sensibly, if necessary, by using multiple starting points for geometry optimizations, preferably by a global structure search algorithm, such as Gaussian Process Regression or Genetic Algorithms.
• When simulating an interface, use at less than 4 layers for the substrate only in well-justified circumstances. If feasible, converge the number of layers for the property of interest.
• When optimizing adsorbate geometries, relax at least the two topmost substrate layers; keep half of the slab or more constrained to speed up the geometry optimization and reproduce bulk-like behavior for the slab.

The electronic structure method
• Van der Waals forces are very important for interface systems. Always use a van der Waals-inclusive method (either couple with a correction scheme or use a non-local functional)!
• Different functionals will yield different results. Resist the urge to vary functionals just to obtain the desired results. Try to understand why one functional provides the desired result and another one doesn't.
• Understand numerical limitations of the electronic structure method. Metal–organic interfaces are typically well described with semilocal functionals. Semiconductor–organic interfaces require (tuned) hybrid functionals and a method to account for doping.
• Obtain geometry and electronic structure at the same level of theory if possible. If not, be aware that even small changes of the geometry may affect the interface physics/chemistry.
• For systems with large interface dipoles, use a linear mixer with a small mixing coefficient before switching to the Pulay mixer.
• Use relatively small coefficients for preconditioner and mixing. Very small preconditioner values reduce the impact of the mixing parameter.
• Start with a reasonably large broadening parameter (~ 0.1–0.2 eV), verify results with a smaller value (~ 0.01) and a denser k-point mesh.

Numerical settings and algorithms
• Be aware which numerical accuracy for which property is desired. Choose numerical settings tight enough to converge a property to a meaningful numerical accuracy. Any tighter and you waste computational time. Any looser and you generate underconverged results.
• Always converge the settings of the calculations for the property of interest. This is not always the total energy! Monitor convergence for each desired property explicitly.
• Check for spurious convergence by inspecting the evolution of the SCF. Suddenly reaching the requested threshold is suspicious.
• Converge, converge, converge! Converge your numerical settings, but also converge the thresholds for the SCF. Convergence tests are time-consuming and boring, but they are the only way to reliably obtain accurate and robust results.
Understand the defaults of the software package you are using.
• The convergence behavior differs between similar systems – do not lightheadedly adopt settings developed for another system.

Geometry
• Get an overview of the potential energy surface, *e.g.* by using Gaussian Process or Bayes Linear regression. Start geometry optimizations near expected minima.
• Create good guess structures before starting the geometry optimization; ideally, multiple guess structures should be used to obtain an idea about the reproducibility.
• Optimize in internal coordinates if possible; be aware of the initialization of the guess for the Hesse matrix; avoid diagonal matrices if possible.
• For systems that are (partially) open-shell, break the spin-symmetry by assigning non-zero initial spin to the atoms.
• When optimizing geometries, be aware of the limits of the force thresholds. If possible, map out important degrees of freedom explicitly to determine the minimum accurately. Alternatively, start multiple geometry optimizations from different starting points.

---

inorganic–organic interfaces, study the choices and reasoning behind different ways to construct atomistic models of the interfaces, and debate the impact of a variety of choices concerning the numerical algorithms and parameters employed in calculating the electronic structure and in performing geometry optimizations. Throughout this review, we highlight the advantages and disadvantages of different approaches in order to provide the reader with the basis to make educated and informed choices. We provide some general "best practice rules" throughout the work, which are summarized in Table 1. We emphasize that these rules reflect our current knowledge – they are certainly not complete, nor are they necessarily completely universal.

It is just as important to be aware of the many possible algorithms and options that can be chosen to calculate the electronic structure. We lay the focus of this work on illustrating some of the pitfalls that exist when using the default values and algorithms that pertinent codes use. Based on several examples – some from our previous work, but most specifically calculated for this work – we show how the application of inappropriate defaults or other sub-optimal settings can easily lead to quantitatively or also qualitatively incorrect results. Our account is based on personal experience and the current state-of-the-art of available methods. We expect (and indeed hope) that some of the best practice suggestions might become obsolete in the future as more advanced methods become available. We hope

that the lessons contained in this work, at the very least, will help to spot spurious results in literature and to avoid falling for artefacts in one's own work.

At the same time, we try to suggest sensible values and default values for interface simulations, that will not only provide correct results, but will also allow to develop faster and more efficient computational workflows. Of course, every interface is special, and those defaults will not always be optimal. We do our best to explain the reasoning for those values and emphasize the correlation with the underlying physics of the interface, to allow the reader to tailor these suggestions to their own system.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 *The Molecule-Metal Interface: KOCH:INTERFACES O-BK*, ed. Koch, N., Ueno, N. and Wee, A. T. S., Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2013, DOI: 10.1002/9783527653171.

2 *Organic Electronics*, ed. Cicoira, F. and Santato, C., Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2013, DOI: 10.1002/9783527650965.

3 *Organic Electronics Materials and Devices*, ed. Ogawa, S., Springer Japan, Tokyo, 2015, DOI: 10.1007/978-4-431-55654-1.

4 H. Ishii, K. Sugiyama, E. Ito and K. Seki, Energy Level Alignment and Interfacial Electronic Structures at Organic/Metal and Organic/Organic Interfaces, *Adv. Mater.*, 1999, **11**(8), 605–625, DOI: 10.1002/(SICI)1521-4095(199906)11:8<605::AID-ADMA605>3.0.CO;2-Q.

5 N. Koch, Organic Electronic Devices and Their Functional Interfaces, *ChemPhysChem*, 2007, **8**(10), 1438–1455, DOI: 10.1002/cphc.200700177.

6 M. Fahlman, S. Fabiano, V. Gueskine, D. Simon, M. Berggren and X. Crispin, Interfaces in Organic Electronics, *Nat. Rev. Mater.*, 2019, **4**(10), 627–650, DOI: 10.1038/s41578-019-0127-y.

7 A. Tkatchenko, L. Romaner, O. T. Hofmann, E. Zojer, C. Ambrosch-Draxl and M. Scheffler, Van Der Waals Interactions Between Organic Adsorbates and at Organic/Inorganic Interfaces, *MRS Bull.*, 2010, **35**(6), 435–442, DOI: 10.1557/mrs2010.581.

8 R. Otero, Vázquez de Parga and A. L. Gallego, J. M. Electronic, Structural and Chemical Effects of Charge-Transfer at Organic/Inorganic Interfaces, *Surf. Sci. Rep.*, 2017, **72**(3), 105–145, DOI: 10.1016/j.surfrep.2017.03.001.

9 J. Nitta, K. Miwa, N. Komiya, E. Annese, J. Fujii, S. Ono and K. Sakamoto, The Actual Electronic Band Structure of a Rubrene Single Crystal, *Sci. Rep.*, 2019, **9**(1), 9645, DOI: 10.1038/s41598-019-46080-4.

10 E. Zojer, T. C. Taucher and O. T. Hofmann, The Impact of Dipolar Layers on the Electronic Properties of Organic/Inorganic Hybrid Interfaces, *Adv. Mater. Interfaces*, 2019, 1900581, DOI: 10.1002/admi.201900581.

11 A. Nilsson; L. Pettersson and J. K. Nørskov, *Chemical Bonding at Surfaces and Interfaces*, Elsevier, Amsterdam, Oxford, 2008.

12 V. Obersteiner, D. A. Egger, G. Heimel and E. Zojer, Impact of Collective Electrostatic Effects on Charge Transport through Molecular Monolayers, *J. Phys. Chem. C*, 2014, **118**(38), 22395–22401, DOI: 10.1021/jp5084955.

13 O. T. Hofmann, P. Rinke, M. Scheffler and G. Heimel, Integer *versus* Fractional Charge Transfer at Metal(/Insulator)/Organic Interfaces: Cu(/NaCl)/TCNE, *ACS Nano*, 2015, **9**(5), 5391–5404, DOI: 10.1021/acsnano.5b01164.

14 F. Rissner, D. A. Egger, A. Natan, T. Körzdörfer, S. Kümmel, L. Kronik and E. Zojer, Collectively Induced Quantum-Confined Stark Effect in Monolayers of Molecules Consisting of Polar Repeating Units, *J. Am. Chem. Soc.*, 2011, **133**(46), 18634–18645, DOI: 10.1021/ja203579c.

15 S. Braun, W. Osikowicz, Y. Wang and W. R. Salaneck, Energy Level Alignment Regimes at Hybrid Organic-Organic and Inorganic-Organic Interfaces, *Org. Electron.*, 2007, **8**(1), 14–20.

16 A. J. Fisher, Methods of Embedding for Defect and Surface Problems, *J. Phys. C Solid State Phys.*, 1988, **21**(17), 3229–3249, DOI: 10.1088/0022-3719/21/17/013.

17 J. C. Cuevas and E. Scheer, *Molecular Electronics: An Introduction to Theory and Experiment*, World Scientific series in nanoscience and nanotechnology, World Scientific, New Jersey, 2nd edn, 2017.

18 N. A. Richter, S. Sicolo, S. V. Levchenko, J. Sauer and M. Scheffler, Concentration of Vacancies at Metal-Oxide Surfaces: Case Study of MgO(100), *Phys. Rev. Lett.*, 2013, **111**(4), 045502, DOI: 10.1103/PhysRevLett.111.045502.

19 C. G. Van de Walle and J. Neugebauer, First-Principles Calculations for Defects and Impurities: Applications to III-Nitrides, *J. Appl. Phys.*, 2004, **95**(8), 3851–3879, DOI: 10.1063/1.1682673.

20 H.-P. Komsa and A. Pasquarello, Finite-Size Supercell Correction for Charged Defects at Surfaces and Interfaces, *Phys. Rev. Lett.*, 2013, **110**(9), 095505, DOI: 10.1103/PhysRevLett.110.095505.

8166 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

21 T. C. Taucher, O. T. Hofmann and E. Zojer, Final-State Simulations of Core-Level Binding Energies at Metal-Organic Hybrid Interfaces: Artifacts Caused by Spurious Collective Electrostatic Effects, *ACS Omega*, 2020, acsomega.0c03209, DOI: 10.1021/acsomega.0c03209.

22 R. A. Beck, L. Lu, A. Petrone, A. C. Ong, P. J. Pauzauskie and X. Li, Spectroscopic Signatures of the B and H$_4$ Polyatomic Nitrogen Aggregates in Nanodiamond, *J. Phys. Chem. C*, 2020, **124**(33), 18275–18283, DOI: 10.1021/acs.jpcc.0c03106.

23 J. Aarons, L. Jones, A. Varambhia, K. E. MacArthur, D. Ozkaya, M. Sarwar, C.-K. Skylaris and P. D. Nellist, Predicting the Oxygen-Binding Properties of Platinum Nanoparticle Ensembles by Combining High-Precision Electron Microscopy and Density Functional Theory, *Nano Lett.*, 2017, **17**(7), 4003–4012, DOI: 10.1021/acs.nanolett.6b04799.

24 A. Petrone, J. J. Goings and X. Li, Quantum Confinement Effects on Optical Transitions in Nanodiamonds Containing Nitrogen Vacancies, *Phys. Rev. B*, 2016, **94**(16), 165402, DOI: 10.1103/PhysRevB.94.165402.

25 A. J. Logsdail, C. A. Downing, T. W. Keal, P. Sherwood, A. A. Sokol and C. R. A. Catlow, Hybrid-DFT Modeling of Lattice and Surface Vacancies in MnO, *J. Phys. Chem. C*, 2019, **123**(13), 8133–8144, DOI: 10.1021/acs.jpcc.8b07846.

26 A. D. Boese and J. Sauer, Accurate Adsorption Energies of Small Molecules on Oxide Surfaces: CO–MgO(001), *Phys. Chem. Chem. Phys.*, 2013, **15**(39), 16481, DOI: 10.1039/c3cp52321g.

27 D. Berger, H. Oberhofer and K. Reuter, First-Principles Embedded-Cluster Calculations of the Neutral and Charged Oxygen Vacancy at the Rutile TiO$_2$ (110) Surface, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2015, **92**(7), 075308, DOI: 10.1103/PhysRevB.92.075308.

28 F. Libisch, C. Huang and E. A. Carter, Embedded Correlated Wavefunction Schemes: Theory and Applications, *Acc. Chem. Res.*, 2014, **47**(9), 2768–2775, DOI: 10.1021/ar500086h.

29 X. Crispin, C. Bureau, V. Geskin, R. Lazzaroni and J.-L. Brédas, Local Density Functional Study of Copper Clusters: A Comparison between Real Clusters, Model Surface Clusters, and the Actual Metal Surface, *Eur. J. Inorg. Chem.*, 1999, 349–360, DOI: 10.1002/(SICI)1099-0682(19990202)1999:2<349::AID-EJIC349>3.0.CO;2-#.

30 P. S. Bagus, C. J. Nelin, S. V. Levchenko, X. Zhao, E. M. Davis, H. Kuhlenbeck and H.-J. Freund, Surface Core Level BE Shifts for CaO(100): Insights into Physical Origins, *Phys. Chem. Chem. Phys.*, 2019, **21**(45), 25431–25438, DOI: 10.1039/C9CP04847B.

31 D. Berger, A. J. Logsdail, H. Oberhofer, M. R. Farrow, C. R. A. Catlow, P. Sherwood, A. A. Sokol, V. Blum and K. Reuter, Embedded-Cluster Calculations in a Numeric Atomic Orbital Density-Functional Theory Framework, *J. Chem. Phys.*, 2014, **141**(2), 024105, DOI: 10.1063/1.4885816.

32 P. Álvarez-Zapatero and A. Aguado, Computational Characterisation of Structure and Metallicity in Small Neutral and Singly-Charged Cadmium Clusters, *Phys. Chem. Chem. Phys.*, 2019, **21**(23), 12321–12334, DOI: 10.1039/C9CP01814J.

33 *Cluster Models for Surface and Bulk Phenomena*, ed. G. Pacchioni, P. S. Bagus and F. Parmigiani, NATO ASI Series, Springer US, Boston, MA, 1992, vol. 283, DOI: 10.1007/978-1-4684-6021-6.

34 F. Weigend and R. Ahlrichs, Quantum Chemical Treatments of Metal Clusters, *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, 2010, **368**(1915), 1245–1263, DOI: 10.1098/rsta.2009.0268.

35 K. Hermann, P. S. Bagus and C. J. Nelin, Size Dependence of Surface Cluster Models: CO Adsorbed on Cu(100), *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1987, **35**(18), 9467–9473, DOI: 10.1103/PhysRevB.35.9467.

36 I. Panas, J. Schüle, P. Siegbahn and U. Wahlgren, On the Cluster Convergence of Chemisorption Energies, *Chem. Phys. Lett.*, 1988, **149**(3), 265–272, DOI: 10.1016/0009-2614(88)85024-3.

37 R. A. Evarestov and V. A. Veryazov, The Electronic Structure of Copper Oxide Crystalline Compounds. I. LUC-CNDO Approach to the Electronic Structure of Cu$_2$O and CuO Crystals, *Phys. Status Solidi B*, 1990, **157**(1), 281–291, DOI: 10.1002/pssb.2221570129.

38 L.-W. Wang and J. Li, First-Principles Thousand-Atom Quantum Dot Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2004, **69**(15), 153302, DOI: 10.1103/PhysRevB.69.153302.

39 G. Kresse, O. Dulub and U. Diebold, Competing Stabilization Mechanism for the Polar ZnO(0001)-Zn Surface, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2003, **68**(24), 245409, DOI: 10.1103/PhysRevB.68.245409.

40 X. Huang, E. Lindgren and J. R. Chelikowsky, Surface Passivation Method for Semiconductor Nanostructures, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2005, **71**(16), 165328, DOI: 10.1103/PhysRevB.71.165328.

41 M. Argeri, A. Fraccarollo, F. Grassi, L. Marchese and M. Cossi, Density Functional Theory Modeling of PbSe Nanoclusters: Effect of Surface Passivation on Shape and Composition, *J. Phys. Chem. C*, 2011, **115**(23), 11382–11389, DOI: 10.1021/jp201112x.

42 Y. Zhang, J. Zhang, K. Tse, L. Wong, C. Chan, B. Deng and J. Zhu, Pseudo-Hydrogen Passivation: A Novel Way to Calculate Absolute Surface Energy of Zinc Blende (111)/($\bar{1}\bar{1}\bar{1}$) Surface, *Sci. Rep.*, 2016, **6**(1), 20055, DOI: 10.1038/srep20055.

43 Y. Xu, O. T. Hofmann, R. Schlesinger, S. Winkler, J. Frisch, J. Niederhausen, A. Vollmer, S. Blumstengel, F. Henneberger, N. Koch, P. Rinke and M. Scheffler, Space-Charge Transfer in Hybrid Inorganic-Organic Systems, *Phys. Rev. Lett.*, 2013, **111**(22), 226802, DOI: 10.1103/PhysRevLett.111.226802.

44 E. Q. Chong, D. B. Lingerfelt, A. Petrone and X. Li, Classical or Quantum? A Computational Study of Small Ion Diffusion in II–VI Semiconductor Quantum Dots, *J. Phys. Chem. C*, 2016, **120**(34), 19434–19441, DOI: 10.1021/acs.jpcc.6b05883.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | **8167**

45 A. J. Fisher, A Comparison of Techniques for Embedding Defect Cluster Calculations, *Theor. Chim. Acta*, 1987, **72**(4), 319–324, DOI: 10.1007/BF00529036.

46 Y. Mishin, M. Asta and J. Li, Atomistic Modeling of Interfaces and Their Impact on Microstructure and Properties, *Acta Mater.*, 2010, **58**(4), 1117–1151, DOI: 10.1016/j.actamat.2009.10.049.

47 P. Huang and E. A. Carter, Advances in Correlated Electronic Structure Methods for Solids, Surfaces, and Nanostructures, *Annu. Rev. Phys. Chem.*, 2008, **59**(1), 261–290, DOI: 10.1146/annurev.physchem.59.032607.093528.

48 Q. Sun and G. K.-L. Chan, Quantum Embedding Theories, *Acc. Chem. Res.*, 2016, **49**(12), 2705–2712, DOI: 10.1021/acs.accounts.6b00356.

49 C. R. Jacob and J. Neugebauer, Subsystem Density-Functional Theory, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**(4), 325–362, DOI: 10.1002/wcms.1175.

50 T. Klüner, N. Govind, Y. A. Wang and E. A. Carter, Periodic Density Functional Embedding Theory for Complete Active Space Self-Consistent Field and Configuration Interaction Calculations: Ground and Excited States., *J. Chem. Phys.*, 2002, **116**(1), 42–54, DOI: 10.1063/1.1420748.

51 P. Huang and E. A. Carter, Self-Consistent Embedding Theory for Locally Correlated Configuration Interaction Wave Functions in Condensed Matter, *J. Chem. Phys.*, 2006, **125**(8), 084102, DOI: 10.1063/1.2336428.

52 F. Rosei, M. Schunack, Y. Naitoh, P. Jiang, A. Gourdon, E. Laegsgaard, I. Stensgaard, C. Joachim and F. Besenbacher, Properties of Large Organic Molecules on Metal Surfaces, *Prog. Surf. Sci.*, 2003, **71**(5–8), 95–146, DOI: 10.1016/S0079-6816(03)00004-2.

53 G. Nenchev, B. Diaconescu, F. Hagelberg and K. Pohl, Self-Assembly of Methanethiol on the Reconstructed Au(111) Surface, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **80**(8), 081401, DOI: 10.1103/PhysRevB.80.081401.

54 E. I. Altman and R. J. Colton, The Interaction of C60 with Noble Metal Surfaces, *Surf. Sci.*, 1993, **295**(1–2), 13–33, DOI: 10.1016/0039-6028(93)90181-I.

55 Th Wandlowski, B. M. Ocko, O. M. Magnussen, S. Wu and J. Lipkowski, The Surface Structure of Au(111) in the Presence of Organic Adlayers: A Combined Electrochemical and Surface X-Ray Scattering Study, *J. Electroanal. Chem.*, 1996, **409**(1–2), 155–164, DOI: 10.1016/0022-0728(95)04479-5.

56 C. Prado, F. Prieto, M. Rueda, J. Feliu and A. Aldaz, Adenine Adsorption on Au(111) and Au(100) Electrodes: Characterisation, Surface Reconstruction Effects and Thermodynamic Study, *Electrochim. Acta*, 2007, **52**(9), 3168–3180, DOI: 10.1016/j.electacta.2006.09.062.

57 Q. Chen, D. J. Frankel and N. V. Richardson, Organic Adsorbate Induced Surface Reconstruction: *P*-Aminobenzoic Acid on Cu{110}, *Langmuir*, 2001, **17**(26), 8276–8280, DOI: 10.1021/la0108712.

58 W. W. Pai, C. L. Hsu, K. C. Lin, L. Y. Sin and T. B. Tang, Characterization and Control of Molecular Ordering on Adsorbate-Induced Reconstructed Surfaces, *Appl. Surf. Sci.*, 2005, **241**(1–2), 194–198, DOI: 10.1016/j.apsusc.2004.09.089.

59 A. Ulman, Formation and Structure of Self-Assembled Monolayers, *Chem. Rev.*, 1996, **96**(4), 1533–1554, DOI: 10.1021/cr9502357.

60 J. C. Love, L. A. Estroff, J. K. Kriebel, R. G. Nuzzo and G. M. Whitesides, Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology, *Chem. Rev.*, 2005, **105**(4), 1103–1170, DOI: 10.1021/cr0300789.

61 C. Vericat, M. E. Vela, G. Benitez, P. Carro and R. C. Salvarezza, Self-Assembled Monolayers of Thiols and Dithiols on Gold: New Challenges for a Well-Known System, *Chem. Soc. Rev.*, 2010, **39**(5), 1805, DOI: 10.1039/b907301a.

62 F. Schreiber, Structure and Growth of Self-Assembling Monolayers, *Prog. Surf. Sci.*, 2000, **65**(5–8), 151–257, DOI: 10.1016/S0079-6816(00)00024-1.

63 X. Crispin, V. Geskin, A. Crispin, J. Cornil, R. Lazzaroni, W. R. Salaneck and J.-L. Brédas, Characterization of the Interface Dipole at Organic/Metal Interfaces, *J. Am. Chem. Soc.*, 2002, **124**(27), 8131–8141, DOI: 10.1021/ja025673r.

64 E. Fabiano, M. Piacenza and S. D'Agostino, Della Sala, F. Towards an Accurate Description of the Electronic Properties of the Biphenylthiol/Gold Interface: The Role of Exact Exchange, *J. Chem. Phys.*, 2009, **131**(23), 234101, DOI: 10.1063/1.3271393.

65 M. L. Sushko, P. V. Sushko, I. V. Abarenkov and A. L. Shluger, QM/MM Method for Metal-Organic Interfaces, *J. Comput. Chem.*, 2010, **31**(16), 2955–2966, DOI: 10.1002/jcc.21591.

66 E. Verwüster, E. Wruss, E. Zojer and O. T. Hofmann, Exploring the Driving Forces behind the Structural Assembly of Biphenylthiolates on Au(111), *J. Chem. Phys.*, 2017, **147**(2), 024706, DOI: 10.1063/1.4991344.

67 E. D. Głowacki, M. Irimia-Vladu, S. Bauer and N. S. Sariciftci, Hydrogen-Bonds in Molecular Solids – from Biological Systems to Organic Electronics, *J. Mater. Chem. B*, 2013, **1**(31), 3742, DOI: 10.1039/c3tb20193g.

68 A. Natan, L. Kronik, H. Haick and R. T. Tung, Electrostatic Properties of Ideal and Non-Ideal Polar Organic Monolayers: Implications for Electronic Devices, *Adv. Mater.*, 2007, **19**(23), 4103–4117, DOI: 10.1002/adma.200701681.

69 D. Cahen, R. Naaman and Z. Vager, The Cooperative Molecular Field Effect, *Adv. Funct. Mater.*, 2005, **15**(10), 1571–1578, DOI: 10.1002/adfm.200500187.

70 A. Natan, Y. Zidon, Y. Shapira and L. Kronik, Cooperative Effects and Dipole Formation at Semiconductor and Self-Assembled-Monolayer Interfaces, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2006, **73**(19), 193310, DOI: 10.1103/PhysRevB.73.193310.

71 G. Heimel, F. Rissner and E. Zojer, Modeling the Electronic Properties of π-Conjugated Self-Assembled Monolayers, *Adv. Mater.*, 2010, **22**(23), 2494–2513, DOI: 10.1002/adma.200903855.

72 O. L. A. Monti, Understanding Interfacial Electronic Structure and Charge Transfer: An Electrostatic Perspective,

8168 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

J. Phys. Chem. Lett., 2012, **3**(17), 2342–2351, DOI: 10.1021/jz300850x.

73 G. Heimel, L. Romaner, J.-L. Brédas and E. Zojer, Interface Energetics and Level Alignment at Covalent Metal-Molecule Junctions: π-Conjugated Thiols on Gold, *Phys. Rev. Lett.*, 2006, **96**(19), 196806, DOI: 10.1103/PhysRevLett.96.196806.

74 D. Deutsch, A. Natan, Y. Shapira and L. Kronik, Electrostatic Properties of Adsorbed Polar Molecules: Opposite Behavior of a Single Molecule and a Molecular Monolayer, *J. Am. Chem. Soc.*, 2007, **129**(10), 2989–2997, DOI: 10.1021/ja068417d.

75 G. Heimel, F. Rissner and E. Zojer, Modeling the Electronic Properties of π-Conjugated Self-Assembled Monolayers, *Adv. Mater.*, 2010, **22**(23), 2494–2513, DOI: 10.1002/adma.200903855.

76 T. C. Taucher, I. Hehn, O. T. Hofmann, M. Zharnikov and E. Zojer, Understanding Chemical versus Electrostatic Shifts in X-Ray Photoelectron Spectra of Organic Self-Assembled Monolayers, *J. Phys. Chem. C*, 2016, **120**(6), 3428–3437, DOI: 10.1021/acs.jpcc.5b12387.

77 D. A. Egger, F. Rissner, E. Zojer and G. Heimel, Polarity Switching of Charge Transport and Thermoelectricity in Self-Assembled Monolayer Devices, *Adv. Mater.*, 2012, **24**(32), 4403–4407, DOI: 10.1002/adma.201200872.

78 H. Edlbauer, E. Zojer and O. T. Hofmann, Postadsorption Work Function Tuning *via* Hydrogen Pressure Control, *J. Phys. Chem. C*, 2015, **119**(48), 27162–27172, DOI: 10.1021/acs.jpcc.5b08827.

79 V. Obersteiner, G. Huhs, N. Papior and E. Zojer, Unconventional Current Scaling and Edge Effects for Charge Transport through Molecular Clusters, *Nano Lett.*, 2017, **17**(12), 7350–7357, DOI: 10.1021/acs.nanolett.7b03066.

80 A. J. Karttunen, T. Tynell and M. Karppinen, Atomic-Level Structural and Electronic Properties of Hybrid Inorganic–Organic ZnO:Hydroquinone Superlattices Fabricated by ALD/MLD, *J. Phys. Chem. C*, 2015, **119**(23), 13105–13114, DOI: 10.1021/acs.jpcc.5b03433.

81 S. Bedwani, D. Wegner, M. F. Crommie and A. Rochefort, Strongly Reshaped Organic-Metal Interfaces: Tetracyanoethylene on Cu(100), *Phys. Rev. Lett.*, 2008, **101**(21), 216105, DOI: 10.1103/PhysRevLett.101.216105.

82 P. J. Blowey, S. Velari, L. A. Rochford, D. A. Duncan, D. A. Warr, T.-L. Lee, A. De Vita, G. Costantini and D. P. Woodruff, Re-Evaluating How Charge Transfer Modifies the Conformation of Adsorbed Molecules, *Nanoscale*, 2018, **10**(31), 14984–14992, DOI: 10.1039/C8NR02237B.

83 J. Neugebauer and M. Scheffler, Adsorbate-Substrate and Adsorbate-Adsorbate Interactions of Na and K Adlayers on Al(111), *Phys. Rev. B: Condens. Matter Mater. Phys*, 1992, **46**(24), 16067–16080, DOI: 10.1103/PhysRevB.46.16067.

84 L. Bengtsson, Dipole Correction for Surface Supercell Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys*, 1999, **59**(19), 12301–12304, DOI: 10.1103/PhysRevB.59.12301.

85 J. Nyman and G. M. Day, Static and Lattice Vibrational Energy Differences between Polymorphs, *CrystEngComm*, 2015, **17**(28), 5154–5165, DOI: 10.1039/C5CE00045A.

86 C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross and A. Rubio, Exact Coulomb Cutoff Technique for Supercell Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys*, 2006, **73**(20), 205119, DOI: 10.1103/PhysRevB.73.205119.

87 L. Yu, V. Ranjan, W. Lu, J. Bernholc and M. B. Nardelli, Equivalence of Dipole Correction and Coulomb Cutoff Techniques in Supercell Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys*, 2008, **77**(24), 245102, DOI: 10.1103/PhysRevB.77.245102.

88 G. J. Martyna and M. E. Tuckerman, A Reciprocal Space Based Method for Treating Long Range Interactions in *Ab Initio* and Force-Field-Based Calculations in Clusters, *J. Chem. Phys.*, 1999, **110**(6), 2810–2821, DOI: 10.1063/1.477923.

89 A. Cerioni, L. Genovese, A. Mirone and V. A. Sole, Efficient and Accurate Solver of the Three-Dimensional Screened and Unscreened Poisson's Equation with Generic Boundary Conditions, *J. Chem. Phys.*, 2012, **137**(13), 134108, DOI: 10.1063/1.4755349.

90 A. Natan, A. Benjamini, D. Naveh, L. Kronik, M. L. Tiago, S. P. Beckman and J. R. Chelikowsky, Real-Space Pseudopotential Method for First Principles Calculations of General Periodic and Partially Periodic Systems, *Phys. Rev. B: Condens. Matter Mater. Phys*, 2008, **78**(7), 075109, DOI: 10.1103/PhysRevB.78.075109.

91 J. C. Womack, L. Anton, J. Dziedzic, P. J. Hasnip, M. I. J. Probert and C.-K. Skylaris, DL_MG: A Parallel Multigrid Poisson and Poisson–Boltzmann Solver for Electronic Structure Calculations in Vacuum and Solution, *J. Chem. Theory Comput.*, 2018, **14**(3), 1412–1432, DOI: 10.1021/acs.jctc.7b01274.

92 N. D. M. Hine, J. Dziedzic, P. D. Haynes and C.-K. Skylaris, Electrostatic Interactions in Finite Systems Treated with Periodic Boundary Conditions: Application to Linear-Scaling Density Functional Theory, *J. Chem. Phys.*, 2011, **135**(20), 204103, DOI: 10.1063/1.3662863.

93 M. M. Melander, M. J. Kuisma, T. E. K. Christensen and K. Honkala, Grand-Canonical Approach to Density Functional Theory of Electrocatalytic Systems: Thermodynamics of Solid-Liquid Interfaces at Constant Ion and Electrode Potentials, *J. Chem. Phys.*, 2019, **150**(4), 041706, DOI: 10.1063/1.5047829.

94 F. Rissner, A. Natan, D. A. Egger, O. T. Hofmann, L. Kronik and E. Zojer, Dimensionality Effects in the Electronic Structure of Organic Semiconductors Consisting of Polar Repeat Units, *Org. Electron.*, 2012, **13**(12), 3165–3176, DOI: 10.1016/j.orgel.2012.09.003.

95 P. S. Bagus, C. J. Nelin, X. Zhao, S. V. Levchenko, E. Davis, X. Weng, F. Späth, C. Papp, H. Kuhlenbeck and H.-J. Freund, Revisiting Surface Core-Level Shifts for Ionic Compounds, *Phys. Rev. B*, 2019, **100**(11), 115419, DOI: 10.1103/PhysRevB.100.115419.

96 S. Lany and A. Zunger, Assessment of Correction Methods for the Band-Gap Problem and for Finite-Size Effects in Supercell Defect Calculations: Case Studies for ZnO and GaAs, *Phys. Rev. B: Condens. Matter Mater. Phys*, 2008, **78**(23), 235104, DOI: 10.1103/PhysRevB.78.235104.

97 G. Makov and M. C. Payne, Periodic Boundary Conditions in *Ab Initio* Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1995, **51**(7), 4014–4022, DOI: 10.1103/PhysRevB.51.4014.

98 C. Freysoldt, J. Neugebauer and C. G. Van de Walle, Fully *Ab Initio* Finite-Size Corrections for Charged-Defect Supercell Calculations, *Phys. Rev. Lett.*, 2009, **102**(1), 016402, DOI: 10.1103/PhysRevLett.102.016402.

99 H.-P. Komsa, T. T. Rantala and A. Pasquarello, Finite-Size Supercell Correction Schemes for Charged Defect Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2012, **86**(4), 045112, DOI: 10.1103/PhysRevB.86.045112.

100 S. E. Taylor and F. Bruneval, Understanding and Correcting the Spurious Interactions in Charged Supercells, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **84**(7), 075155, DOI: 10.1103/PhysRevB.84.075155.

101 N. Moll, Y. Xu, O. T. Hofmann and P. Rinke, Stabilization of Semiconductor Surfaces through Bulk Dopants, *New J. Phys.*, 2013, **15**(8), 083009, DOI: 10.1088/1367-2630/15/8/083009.

102 O. Sinai, O. T. Hofmann, P. Rinke, M. Scheffler, G. Heimel and L. Kronik, Multiscale Approach to the Electronic Structure of Doped Semiconductor Surfaces, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2015, **91**(7), 075311, DOI: 10.1103/PhysRevB.91.075311.

103 S. Erker, P. Rinke, N. Moll and O. T. Hofmann, Doping Dependence of the Surface Phase Stability of Polar O-Terminated (000$\bar{1}$) ZnO, *New J. Phys.*, 2017, **19**(8)), 083012, DOI: 10.1088/1367-2630/aa79e7.

104 M. Otani and O. Sugino, First-Principles Calculations of Charged Surfaces and Interfaces: A Plane-Wave Nonrepeated Slab Approach., *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2006, **73**(11), 115407, DOI: 10.1103/PhysRevB.73.115407.

105 I. Dabo, B. Kozinsky, N. E. Singh-Miller and N. Marzari, Electrostatics in Periodic Boundary Conditions and Real-Space Corrections, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **77**(11), 115139, DOI: 10.1103/PhysRevB.77.115139.

106 T.-L. Chan, S. B. Zhang and J. R. Chelikowsky, Charged Dopants in Semiconductor Nanowires under Partially Periodic Boundary Conditions, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **83**(24), 245440, DOI: 10.1103/PhysRevB.83.245440.

107 *Electrified Interfaces in Physics, Chemistry and Biology*, ed. R. Guidelli, 1992, ISBN 978-94-011-2566-6.

108 G. Kastlunger, P. Lindgren and A. A. Peterson, Controlled-Potential Simulation of Elementary Electrochemical Reactions: Proton Discharge on Metal Surfaces, *J. Phys. Chem. C*, 2018, **122**(24), 12771–12781, DOI: 10.1021/acs.jpcc.8b02465.

109 M. Scheffler, Lattice Relaxations at Substitutional Impurities in Semiconductors, *Phys. BC*, 1987, **146**(1–2), 176–186, DOI: 10.1016/0378-4363(87)90060-X.

110 R. Schlesinger, Y. Xu, O. T. Hofmann, S. Winkler, J. Frisch, J. Niederhausen, A. Vollmer, S. Blumstengel, F. Henneberger, P. Rinke, M. Scheffler and N. Koch, Controlling the Work Function of ZnO and the Energy-Level Alignment at the Interface to Organic Semiconductors with a Molecular Electron Acceptor, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**(15), 155311, DOI: 10.1103/PhysRevB.87.155311.

111 M. Gruenewald, L. K. Schirra, P. Winget, M. Kozlik, P. F. Ndione, A. K. Sigdel, J. J. Berry, R. Forker, J.-L. Brédas, T. Fritz and O. L. A. Monti, Integer Charge Transfer and Hybridization at an Organic Semiconductor/Conductive Oxide Interface, *J. Phys. Chem. C*, 2015, **119**(9), 4865–4873, DOI: 10.1021/jp512153b.

112 C. Freysoldt, A. Mishra, M. Ashton and J. Neugebauer, Generalized Dipole Correction for Charged Surfaces in the Repeated-Slab Approach, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2020, **102**(4), 045403, DOI: 10.1103/PhysRevB.102.045403.

113 K. Hermann, *Crystallography and Surface Structure: An Introduction for Surface Scientists and Nanoscientists*, Wiley-VCH, Weinheim, 2nd, revised and expanded edition, 2016.

114 R. Forker, M. Meissner and T. Fritz, Classification of Epitaxy in Reciprocal and Real Space: Rigid versus Flexible Lattices, *Soft Matter*, 2017, **13**(9), 1748–1758, DOI: 10.1039/C6SM02688E.

115 O. M. Braun and Y. S. Kivshar, Nonlinear Dynamics of the Frenkel–Kontorova Model, *Phys. Rep.*, 1998, **306**(1–2), 1–108, DOI: 10.1016/S0370-1573(98)00029-5.

116 M. Meissner, F. Sojka, L. Matthes, F. Bechstedt, X. Feng, K. Müllen, S. C. B. Mannsfeld, R. Forker and T. Fritz, Flexible 2D Crystals of Polycyclic Aromatics Stabilized by Static Distortion Waves, *ACS Nano*, 2016, **10**(7), 6474–6483, DOI: 10.1021/acsnano.6b00935.

117 J. R. Reimers, M. J. Ford and L. Goerigk, Problems, Successes and Challenges for the Application of Dispersion-Corrected Density-Functional Theory Combined with Dispersion-Based Implicit Solvent Models to Large-Scale Hydrophobic Self-Assembly and Polymorphism, *Mol. Simul.*, 2016, **42**(6–7), 494–510, DOI: 10.1080/08927022.2015.1066504.

118 J. Behler, Perspective: Machine Learning Potentials for Atomistic Simulations, *J. Chem. Phys.*, 2016, **145**(17), 170901, DOI: 10.1063/1.4966192.

119 N. Bernstein, J. R. Kermode and G. Csányi, Hybrid Atomistic Simulation Methods for Materials Systems, *Rep. Prog. Phys.*, 2009, **72**(2), 026501, DOI: 10.1088/0034-4885/72/2/026501.

120 K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O. A. von Lilienfeld, A. Tkatchenko and K.-R. Müller, Assessment and Validation of Machine Learning Methods for Predicting Molecular Atomization Energies, *J. Chem. Theory Comput.*, 2013, **9**(8), 3404–3419, DOI: 10.1021/ct400195d.

121 M. Rupp, A. Tkatchenko, K.-R. Müller and O. A. von Lilienfeld, Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning, *Phys. Rev. Lett.*, 2012, **108**(5), 058301, DOI: 10.1103/PhysRevLett.108.058301.

122 F. Noé, A. Tkatchenko, K.-R. Müller and C. Clementi, Machine Learning for Molecular Simulation, *Annu. Rev.*

8170 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

Phys. Chem., 2020, **71**(1), 361–390, DOI: 10.1146/annurev-physchem-042018-052331.

123 T. Mueller, A. Hernandez and C. Wang, Machine Learning for Interatomic Potential Models, *J. Chem. Phys.*, 2020, **152**(5), 050902, DOI: 10.1063/1.5126336.

124 A. M. Track, F. Rissner, G. Heimel, L. Romaner, D. Käfer, A. Bashir, G. M. Rangger, O. T. Hofmann, T. Bučko, G. Witte and E. Zojer, Simultaneously Understanding the Geometric and Electronic Structure of Anthraceneselenolate on Au(111): A Combined Theoretical and Experimental Study, *J. Phys. Chem. C*, 2010, **114**(6), 2677–2684, DOI: 10.1021/jp9102756.

125 C. Zwick, A. Baby, M. Gruenewald, E. Verwüster, O. T. Hofmann, R. Forker, G. Fratesi, G. P. Brivio, E. Zojer and T. Fritz, Complex Stoichiometry-Dependent Reordering of 3,4,9,10-Perylenetetracarboxylic Dianhydride on Ag(111) upon K Intercalation, *ACS Nano*, 2016, **10**(2), 2365–2374, DOI: 10.1021/acsnano.5b07145.

126 G. Koller, S. Surnev, M. G. Ramsey and F. P. Netzer, Sexiphenyl on a Ni(110)(2 × 1)-O Surface: A Single-Molecule STM Study, *Surf. Sci.*, 2004, **559**(2–3), L187–L193, DOI: 10.1016/j.susc.2004.04.019.

127 M. Wagner, P. Puschnig, S. Berkebile, F. P. Netzer and M. G. Ramsey, Alternating Chirality in the Monolayer H2TPP on Cu(110)–(2 × 1)O, *Phys. Chem. Chem. Phys.*, 2013, **15**(13), 4691, DOI: 10.1039/c3cp44239j.

128 G. Meyer, B. Neu and K.-H. Rieder, Controlled Lateral Manipulation of Single Molecules with the Scanning Tunneling Microscope, *Appl. Phys. Mater. Sci. Process.*, 1995, **60**(3), 343–345, DOI: 10.1007/BF01538415.

129 STM and AFM Studies on (Bio)Molecular Systems: Unravelling the Nanoworld, ed. Samorì, P., *Topics in Current Chemistry*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, vol. 285, DOI: 10.1007/978-3-540-78395-4.

130 P. Kocán, Y. Yoshimoto, K. Yagyu, H. Tochihara and T. Suzuki, Adsorption of PTCDA on Ge(001), *J. Phys. Chem. C*, 2017, **121**(6), 3320–3326, DOI: 10.1021/acs.jpcc.6b09793.

131 P. S. Weiss and D. M. Eigler, Site Dependence of the Apparent Shape of a Molecule in Scanning Tunneling Microscope Images: Benzene on Pt{111}, *Phys. Rev. Lett.*, 1993, **71**(19), 3139–3142, DOI: 10.1103/PhysRevLett.71.3139.

132 M. N. Faraggi, N. Jiang, N. Gonzalez-Lakunza, A. Langner, S. Stepanow, K. Kern and A. Arnau, Bonding and Charge Transfer in Metal–Organic Coordination Networks on Au(111) with Strong Acceptor Molecules, *J. Phys. Chem. C*, 2012, **116**(46), 24558–24565, DOI: 10.1021/jp306780n.

133 P. Maksymovych, O. Voznyy, D. B. Dougherty, D. C. Sorescu and J. T. Yates, Gold Adatom as a Key Structural Component in Self-Assembled Monolayers of Organosulfur Molecules on Au(111), *Prog. Surf. Sci.*, 2010, **85**(5–8), 206–240, DOI: 10.1016/j.progsurf.2010.05.001.

134 D. Wegner, R. Yamachika, Y. Wang, V. W. Brar, B. M. Bartlett, J. R. Long and M. F. Crommie, Single-Molecule Charge Transfer and Bonding at an Organic/Inorganic Interface: Tetracyanoethylene on Noble Metals, *Nano Lett.*, 2008, **8**(1), 131–135, DOI: 10.1021/nl072217y.

135 P. Maksymovych, D. C. Sorescu and J. T. Yates, Gold-Adatom-Mediated Bonding in Self-Assembled Short-Chain Alkanethiolate Species on the Au(111) Surface, *Phys. Rev. Lett.*, 2006, **97**(14), 146103, DOI: 10.1103/PhysRevLett.97.146103.

136 G. Yang and G. Liu, New Insights for Self-Assembled Monolayers of Organothiols on Au(111) Revealed by Scanning Tunneling Microscopy, *J. Phys. Chem. B*, 2003, **107**(34), 8746–8759, DOI: 10.1021/jp0219810.

137 R. Mazzarello, A. Cossaro, A. Verdini, R. Rousseau, L. Casalis, M. F. Danisman, L. Floreano, S. Scandolo, A. Morgante and G. Scoles, Structure of a CH3S Monolayer on Au(111) Solved by the Interplay between Molecular Dynamics Calculations and Diffraction Measurements, *Phys. Rev. Lett.*, 2007, **98**(1), 016102, DOI: 10.1103/PhysRevLett.98.016102.

138 O. Voznyy, J. J. Dubowski, J. T. Yates and P. Maksymovych, The Role of Gold Adatoms and Stereochemistry in Self-Assembly of Methylthiolate on Au(111), *J. Am. Chem. Soc.*, 2009, **131**(36), 12989–12993, DOI: 10.1021/ja902629y.

139 A. Della-Pia, M. Riello, J. Lawrence, D. Stassen, T. S. Jones, D. Bonifazi, A. De-Vita and G. Costantini, Two-Dimensional Ketone-Driven Metal-Organic Coordination on Cu(111), *Chem. – Eur. J*, 2016, **22**(24), 8105–8112, DOI: 10.1002/chem.201600368.

140 G. Wang, A. Rühling, S. Amirjalayer, M. Knor, J. B. Ernst, C. Richter, H.-J. Gao, A. Timmer, H.-Y. Gao, N. L. Doltsinis, F. Glorius and H. Fuchs, Ballbot-Type Motion of N-Heterocyclic Carbenes on Gold Surfaces, *Nat. Chem.*, 2017, **9**(2), 152–156, DOI: 10.1038/nchem.2622.

141 J. V. Barth, G. Costantini and K. Kern, Engineering Atomic and Molecular Nanostructures at Surfaces, *Nature*, 2005, **437**(7059), 671–679, DOI: 10.1038/nature04166.

142 O. Dulub, U. Diebold and G. Kresse, Novel Stabilization Mechanism on Polar Surfaces: ZnO(0001)-Zn, *Phys. Rev. Lett.*, 2003, **90**(1), 016102, DOI: 10.1103/PhysRevLett.90.016102.

143 J. V. Lauritsen, S. Porsgaard, M. K. Rasmussen, M. C. R. Jensen, R. Bechstein, K. Meinander, B. S. Clausen, S. Helveg, R. Wahl, G. Kresse and F. Besenbacher, Stabilization Principles for Polar Surfaces of ZnO, *ACS Nano*, 2011, **5**(7), 5987–5994, DOI: 10.1021/nn2017606.

144 C. Woll, The Chemistry and Physics of Zinc Oxide Surfaces, *Prog. Surf. Sci.*, 2007, **82**(2–3), 55–120, DOI: 10.1016/j.progsurf.2006.12.002.

145 J. Niederhausen, A. Franco-Cañellas, S. Erker, T. Schultz, K. Broch, A. Hinderhofer, S. Duhm, P. K. Thakur, D. A. Duncan, A. Gerlach, T.-L. Lee, O. T. Hofmann, F. Schreiber and N. Koch, X-Ray Standing Waves Reveal Lack of OH Termination at Hydroxylated ZnO(0001) Surfaces, *Phys. Rev. Mater.*, 2020, **4**(2), 020602, DOI: 10.1103/PhysRevMaterials.4.020602.

146 H. Wang, S. V. Levchenko, T. Schultz, N. Koch, M. Scheffler and M. Rossi, Modulation of the Work Function by the Atomic Structure of Strong Organic Electron Acceptors on

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | 8171

H-Si(111), *Adv. Electron. Mater.*, 2019, **5**(5), 1800891, DOI: 10.1002/aelm.201800891.

147 E. Verwüster, O. T. Hofmann, D. A. Egger and E. Zojer, Electronic Properties of Biphenylthiolates on Au(111): The Impact of Coverage Revisited, *J. Phys. Chem. C*, 2015, **119**(14), 7817–7825, DOI: 10.1021/acs.jpcc.5b00992.

148 S. Duhm, C. Bürker, J. Niederhausen, I. Salzmann, T. Hosokai, J. Duvernay, S. Kera, F. Schreiber, N. Koch, N. Ueno and A. Gerlach, Pentacene on Ag(111): Correlation of Bonding Distance with Intermolecular Interaction and Order, *ACS Appl. Mater. Interfaces*, 2013, **5**(19), 9377–9381, DOI: 10.1021/am402778u.

149 B. P. Klein, J. M. Morbec, M. Franke, K. K. Greulich, M. Sachs, S. Parhizkar, F. C. Bocquet, M. Schmid, S. J. Hall, R. J. Maurer, B. Meyer, R. Tonner, C. Kumpf, P. Kratzer and J. M. Gottfried, Molecule–Metal Bond of Alternant versus Nonalternant Aromatic Systems on Coinage Metal Surfaces: Naphthalene versus Azulene on Ag(111) and Cu(111), *J. Phys. Chem. C*, 2019, **123**(48), 29219–29230, DOI: 10.1021/acs.jpcc.9b08824.

150 B. Bröker, O. T. Hofmann, G. M. Rangger, P. Frank, R.-P. Blum, R. Rieger, L. Venema, A. Vollmer, K. Müllen, J. P. Rabe, A. Winkler, P. Rudolf, E. Zojer and N. Koch, Density-Dependent Reorientation and Rehybridization of Chemisorbed Conjugated Molecules for Controlling Interface Electronic Structure, *Phys. Rev. Lett.*, 2010, **104**(24), 246805, DOI: 10.1103/PhysRevLett.104.246805.

151 O. T. Hofmann, H. Glowatzki, C. Bürker, G. M. Rangger, B. Bröker, J. Niederhausen, T. Hosokai, I. Salzmann, R.-P. Blum, R. Rieger, A. Vollmer, P. Rajput, A. Gerlach, K. Müllen, F. Schreiber, E. Zojer, N. Koch and S. Duhm, Orientation-Dependent Work-Function Modification Using Substituted Pyrene-Based Acceptors, *J. Phys. Chem. C*, 2017, **121**(44), 24657–24668, DOI: 10.1021/acs.jpcc.7b08451.

152 M.-T. Chen, O. T. Hofmann, A. Gerlach, B. Bröker, C. Bürker, J. Niederhausen, T. Hosokai, J. Zegenhagen, A. Vollmer, R. Rieger, K. Müllen, F. Schreiber, I. Salzmann, N. Koch, E. Zojer and S. Duhm, Energy-Level Alignment at Strongly Coupled Organic–Metal Interfaces, *J. Phys.: Condens. Matter*, 2019, **31**(19), 194002, DOI: 10.1088/1361-648X/ab0171.

153 S. Lach, A. Altenhof, S. Shi, M. Fahlman and C. Ziegler, Electronic and Magnetic Properties of a Ferromagnetic Cobalt Surface by Adsorbing Ultrathin Films of Tetracyanoethylene, *Phys. Chem. Chem. Phys.*, 2019, **21**(28), 15833–15844, DOI: 10.1039/C9CP02205H.

154 F. Sojka, M. Meissner, T. Yamada, T. Munakata, R. Forker and T. Fritz, Naphthalene's Six Shades on Graphite: A Detailed Study on the Polymorphism of an Apparently Simple System, *J. Phys. Chem. C*, 2016, **120**(40), 22972–22978, DOI: 10.1021/acs.jpcc.6b06702.

155 M. Lackinger, S. Griessl, W. M. Heckl and M. Hietschold, Coronene on Ag(111) Investigated by LEED and STM in UHV, *J. Phys. Chem. B*, 2002, **106**(17), 4482–4485, DOI: 10.1021/jp014275s.

156 G. Mercurio, R. J. Maurer, W. Liu, S. Hagen, F. Leyssner, P. Tegeder, J. Meyer, A. Tkatchenko, S. Soubatch, K. Reuter and F. S. Tautz, Quantification of Finite-Temperature Effects on Adsorption Geometries of π-Conjugated Molecules: Azobenzene/Ag(111), *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **88**(3), 035421, DOI: 10.1103/PhysRevB.88.035421.

157 M. Müller, K. Diller, R. J. Maurer and K. Reuter, Interfacial Charge Rearrangement and Intermolecular Interactions: Density-Functional Theory Study of Free-Base Porphine Adsorbed on Ag(111) and Cu(111), *J. Chem. Phys.*, 2016, **144**(2), 024701, DOI: 10.1063/1.4938259.

158 A. T. Egger, L. Hörmann, A. Jeindl, M. Scherbela, V. Obersteiner, M. Todorović, P. Rinke and O. T. Hofmann, Charge Transfer into Organic Thin Films: A Deeper Insight through Machine-Learning-Assisted Structure Search, *Adv. Sci.*, 2020, 2000992, DOI: 10.1002/advs.202000992.

159 F. Blobner, P. N. Abufager, R. Han, J. Bauer, D. A. Duncan, R. J. Maurer, K. Reuter, P. Feulner and F. Allegretti, Thiolate-Bonded Self-Assembled Monolayers on Ni(111): Bonding Strength, Structure, and Stability, *J. Phys. Chem. C*, 2015, **119**(27), 15455–15468, DOI: 10.1021/acs.jpcc.5b04351.

160 M. Willenbockel, R. J. Maurer, C. Bronner, M. Schulze, B. Stadtmüller, S. Soubatch, P. Tegeder, K. Reuter and F. Stefan Tautz, Coverage-Driven Dissociation of Azobenzene on Cu(111): A Route towards Defined Surface Functionalization, *Chem. Commun.*, 2015, **51**(83), 15324–15327, DOI: 10.1039/C5CC05003K.

161 V. Ruiz, W. Liu, E. Zojer, M. Scheffler and A. Tkatchenko, Density-Functional Theory with Screened van Der Waals Interactions for the Modeling of Hybrid Inorganic-Organic Systems, *Phys. Rev. Lett.*, 2012, **108**(14), 146103, DOI: 10.1103/PhysRevLett.108.146103.

162 A. Jeindl, J. Domke, L. Hörmann, F. Sojka, R. Forker, T. Fritz and O. T. Hofmann, Nonintuitive Surface Self-Assembly of Functionalized Molecules on Ag(111), *ACS Nano*, 2021, DOI: 10.1021/acsnano.0c10065.

163 M. Scherbela, L. Hörmann, A. Jeindl, V. Obersteiner and O. T. Hofmann, Charting the Energy Landscape of Metal/Organic Interfaces via Machine Learning, *Phys. Rev. Mater.*, 2018, **2**(4), 043803, DOI: 10.1103/PhysRevMaterials.2.043803.

164 T. Deilmann, P. Krüger, M. Rohlfing and D. Wegner, Adsorption and STM Imaging of Tetracyanoethylene on Ag(001): An Ab Initio Study, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2014, **89**(4), 045405, DOI: 10.1103/PhysRevB.89.045405.

165 L. Hörmann, A. Jeindl, A. T. Egger, M. Scherbela and O. T. Hofmann, SAMPLE: Surface Structure Search Enabled by Coarse Graining and Statistical Learning, *Comput. Phys. Commun.*, 2019, S0010465519301973, DOI: 10.1016/j.cpc.2019.06.010.

166 M. Todorović, M. U. Gutmann, J. Corander and P. Rinke, Bayesian Inference of Atomistic Structure in Functional Materials, *Npj Comput. Mater.*, 2019, **5**(1), 35, DOI: 10.1038/s41524-019-0175-2.

167 J. A. Lloyd, A. C. Papageorgiou, S. Fischer, S. C. Oh, Ö. Sağlam, K. Diller, D. A. Duncan, F. Allegretti,

8172 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

F. Klappenberger, M. Stöhr, R. J. Maurer, K. Reuter, J. Reichert and J. V. Barth, Dynamics of Spatially Confined Bisphenol A Trimers in a Unimolecular Network on Ag(111), *Nano Lett.*, 2016, **16**(3), 1884–1889, DOI: 10.1021/acs.nanolett.5b05026.

168 K. Krautgasser, C. Panosetti, D. Palagin, K. Reuter and R. J. Maurer, Global Structure Search for Molecules on Surfaces: Efficient Sampling with Curvilinear Coordinates., *J. Chem. Phys.*, 2016, **145**(8), 084117, DOI: 10.1063/1.4961259.

169 R. A. Puglisi, S. Caccamo, C. Bongiorno, G. Fisicaro, L. Genovese, S. Goedecker, G. Mannino and A. La Magna, Direct Observation of Single Organic Molecules Grafted on the Surface of a Silicon Nanowire, *Sci. Rep.*, 2019, **9**(1), 5647, DOI: 10.1038/s41598-019-42073-5.

170 P. Seema, J. Behler and D. Marx, Peeling by Nanomechanical Forces: A Route to Selective Creation of Surface Structures, *Phys. Rev. Lett.*, 2015, **115**(3)), 036102, DOI: 10.1103/PhysRevLett.115.036102.

171 D. M. Packwood, P. Han and T. Hitosugi, Chemical and Entropic Control on the Molecular Self-Assembly Process, *Nat. Commun.*, 2017, **8**(1), 14463, DOI: 10.1038/ncomms14463.

172 G. Copie, F. Cleri, Y. Makoudi, C. Krzeminski, M. Berthe, F. Cherioux, F. Palmino and B. Grandidier, Surface-Induced Optimal Packing of Two-Dimensional Molecular Networks, *Phys. Rev. Lett.*, 2015, **114**(6), 066101, DOI: 10.1103/PhysRevLett.114.066101.

173 T. J. Roussel, E. Barrena, C. Ocal and J. Faraudo, Predicting Supramolecular Self-Assembly on Reconstructed Metal Surfaces, *Nanoscale*, 2014, **6**(14), 7991–8001, DOI: 10.1039/C4NR01987C.

174 K. Burke, Perspective on Density Functional Theory, *J. Chem. Phys.*, 2012, **136**(15), 150901, DOI: 10.1063/1.4704546.

175 R. J. Maurer, C. Freysoldt, A. M. Reilly, J. G. Brandenburg, O. T. Hofmann, T. Björkman, S. Lebègue and A. Tkatchenko, Advances in Density-Functional Calculations for Materials Modeling, *Annu. Rev. Mater. Res.*, 2019, **49**(1), 1–30, DOI: 10.1146/annurev-matsci-070218-010143.

176 R. J. Maurer, V. G. Ruiz, J. Camarillo-Cisneros, W. Liu, N. Ferri, K. Reuter and A. Tkatchenko, Adsorption Structures and Energetics of Molecules on Metal Surfaces: Bridging Experiment and Theory, *Prog. Surf. Sci.*, 2016, **91**(2), 72–100, DOI: 10.1016/j.progsurf.2016.05.001.

177 J. G. Lee, Trends in Computational Materials Science Based on Density Functional Theory, *J. Korean Ceram. Soc.*, 2016, **53**(2), 184–193, DOI: 10.4191/kcers.2016.53.2.184.

178 W. Liu, A. Tkatchenko and M. Scheffler, Modeling Adsorption and Reactions of Organic Molecules at Metal Surfaces, *Acc. Chem. Res.*, 2014, **47**(11), 3369–3377, DOI: 10.1021/ar500118y.

179 J. P. P. Ramalho, J. R. B. Gomes and F. Illas, Accounting for van Der Waals Interactions between Adsorbates and Surfaces in Density Functional Theory Based Calculations: Selected Examples, *RSC Adv.*, 2013, **3**(32), 13085, DOI: 10.1039/c3ra40713f.

180 C. Draxl, D. Nabok and K. Hannewald, Organic/Inorganic Hybrid Materials: Challenges for *Ab Initio* Methodology, *Acc. Chem. Res.*, 2014, **47**(11), 3225–3232, DOI: 10.1021/ar500096q.

181 F. Flores, J. Ortega and H. Vázquez, Modelling Energy Level Alignment at Organic Interfaces and Density Functional Theory, *Phys. Chem. Chem. Phys.*, 2009, **11**(39), 8658, DOI: 10.1039/b902492c.

182 D. Rappoport; N. R. M. Crawford; F. Furche and K. Burke, Approximate Density Functionals: Which Should I Choose? in *Encyclopedia of Inorganic Chemistry*, ed. King, R. B., Crabtree, R. H., Lukehart, C. M., Atwood, D. A. and Scott, R. A., John Wiley & Sons, Ltd, Chichester, UK, 2009; p. ia615, DOI: 10.1002/0470862106.ia615.

183 R. Peverati and D. G. Truhlar, Quest for a Universal Density Functional: The Accuracy of Density Functionals across a Broad Spectrum of Databases in Chemistry and Physics, *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, 2014, **372**(2011), 20120476, DOI: 10.1098/rsta.2012.0476.

184 P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.*, 1964, **136**(3B), B864–B871, DOI: 10.1103/PhysRev.136.B864.

185 W. Kohn and L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.*, 1965, **140**(4A), A1133–A1138, DOI: 10.1103/PhysRev.140.A1133.

186 J. P. Perdew, A. Ruzsinszky, L. A. Constantin, J. Sun and G. I. Csonka, Some Fundamental Issues in Ground-State Density Functional Theory: A Guide for the Perplexed, *J. Chem. Theory Comput.*, 2009, **5**(4), 902–908, DOI: 10.1021/ct800531s.

187 N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler and L. Kronik, Dispersion Interactions with Density-Functional Theory: Benchmarking Semiempirical and Interatomic Pairwise Corrected Density Functionals, *J. Chem. Theory Comput.*, 2011, **7**(12), 3944–3951, DOI: 10.1021/ct2005616.

188 Z. Wang, S. M. Selbach and T. Grande, Van Der Waals Density Functional Study of the Energetics of Alkali Metal Intercalation in Graphite, *RSC Adv.*, 2014, **4**(8), 3973–3983, DOI: 10.1039/C3RA47187J.

189 J. Klimeš and A. Michaelides, Perspective: Advances and Challenges in Treating van Der Waals Dispersion Forces in Density Functional Theory, *J. Chem. Phys.*, 2012, **137**(12), 120901, DOI: 10.1063/1.4754130.

190 M. Rohlfing, R. Temirov and F. S. Tautz, Adsorption Structure and Scanning Tunneling Data of a Prototype Organic-Inorganic Interface: PTCDA on Ag(111), *Phys. Rev. B: Condens. Matter Mater. Phys*, 2007, **76**(11), 115421, DOI: 10.1103/PhysRevB.76.115421.

191 L. Romaner, D. Nabok, P. Puschnig, E. Zojer and C. Ambrosch-Draxl, Theoretical Study of PTCDA Adsorbed on the Coinage Metal Surfaces, Ag (111), Au (111) and Cu (111), *New J. Phys.*, 2009, **11**, 053010, DOI: 10.1088/1367-2630/11/5/053010.

192 G. Saranya, S. Nair, V. Natarajan and K. Senthilkumar, Adsorption of Perfluoropentacene on Aluminum (100)

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | 8173

Surface: Structural and Electronic Properties from First Principle Study, *Comput. Mater. Sci.*, 2014, **89**, 216–223, DOI: 10.1016/j.commatsci.2014.03.049.

193 J. I. Martínez, E. Abad, F. Flores and J. Ortega, Simulating the Organic-Molecule/Metal Interface TCNQ/Au(111), *Phys. Status Solidi B*, 2011, 2044–2049, DOI: 10.1002/pssb.201147136.

194 W. Liu, J. Carrasco, B. Santra, A. Michaelides, M. Scheffler and A. Tkatchenko, Benzene Adsorbed on Metals: Concerted Effect of Covalency and van Der Waals Bonding, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2012, **86**(24), 245405, DOI: 10.1103/PhysRevB.86.245405.

195 K. E. Riley, E. N. Brothers, K. B. Ayers and K. M. Merz, Accurate Atomic and Molecular Calculations without Gradient Corrections: Scaled SVWNV Density Functional, *J. Chem. Theory Comput.*, 2005, **1**(4), 546–553, DOI: 10.1021/ct050007c.

196 Y. Mo, G. Tian, R. Car, V. N. Staroverov, G. E. Scuseria and J. Tao, Performance of a Nonempirical Density Functional on Molecules and Hydrogen-Bonded Complexes, *J. Chem. Phys.*, 2016, **145**(23), 234306, DOI: 10.1063/1.4971853.

197 P. H. T. Philipsen and E. J. Baerends, Cohesive Energy of 3 *d* Transition Metals: Density Functional Theory Atomic and Bulk Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**(8), 5326–5333, DOI: 10.1103/PhysRevB.54.5326.

198 G.-X. Zhang, A. M. Reilly, A. Tkatchenko and M. Scheffler, Performance of Various Density-Functional Approximations for Cohesive Properties of 64 Bulk Solids, *New J. Phys.*, 2018, **20**(6), 063020, DOI: 10.1088/1367-2630/aac7f0.

199 J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, Atoms, Molecules, Solids, and Surfaces: Applications of the Generalized Gradient Approximation for Exchange and Correlation, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1992, **46**(11), 6671–6687, DOI: 10.1103/PhysRevB.46.6671.

200 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.*, 1996, **77**(18), 3865–3868, DOI: 10.1103/PhysRevLett.77.3865.

201 Y. Zhang and W. Yang, Comment on "Generalized Gradient Approximation Made Simple.", *Phys. Rev. Lett.*, 1998, **80**(4), 890, DOI: 10.1103/PhysRevLett.80.890.

202 J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces, *Phys. Rev. Lett.*, 2008, **100**(13), 136406, DOI: 10.1103/PhysRevLett.100.136406.

203 B. Hammer, L. B. Hansen and J. K. Nørskov, Improved Adsorption Energetics within Density-Functional Theory Using Revised Perdew-Burke-Ernzerhof Functionals, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**(11), 7413–7421, DOI: 10.1103/PhysRevB.59.7413.

204 R. Rurali, N. Lorente and P. Ordejón, Comment on "Molecular Distortions and Chemical Bonding of a Large π-Conjugated Molecule on a Metal Surface, *Phys. Rev. Lett.*, 2005, **95**(20), 209601, DOI: 10.1103/PhysRevB.95.209601.

205 V. N. Staroverov, G. E. Scuseria, J. Tao and J. P. Perdew, Tests of a Ladder of Density Functionals for Bulk Solids and Surfaces, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2004, **69**(7), 075102, DOI: 10.1103/PhysRevB.69.075102.

206 P. Haas, F. Tran and P. Blaha, Calculation of the Lattice Constant of Solids with Semilocal Functionals, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **79**(8), 085104, DOI: 10.1103/PhysRevB.79.085104.

207 F. Tran, R. Laskowski, P. Blaha and K. Schwarz, Performance on Molecules, Surfaces, and Solids of the Wu-Cohen GGA Exchange-Correlation Energy Functional, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **75**(11), 115131, DOI: 10.1103/PhysRevB.75.115131.

208 A. Stroppa and G. Kresse, The Shortcomings of Semi-Local and Hybrid Functionals: What We Can Learn from Surface Science Studies, *New J. Phys.*, 2008, **10**(6), 063020, DOI: 10.1088/1367-2630/10/6/063020.

209 S. De Waele, K. Lejaeghere, M. Sluydts and S. Cottenier, Error Estimates for Density-Functional Theory Predictions of Surface Energy and Work Function, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2016, **94**(23), 235418, DOI: 10.1103/PhysRevB.94.235418.

210 A. L. Hickey and C. N. Rowley, Benchmarking Quantum Chemical Methods for the Calculation of Molecular Dipole Moments and Polarizabilities, *J. Phys. Chem. A*, 2014, **118**(20), 3678–3687, DOI: 10.1021/jp502475e.

211 L. Romaner, G. Heimel, J.-L. Brédas, A. Gerlach, F. Schreiber, R. L. Johnson, J. Zegenhagen, S. Duhm, N. Koch and E. Zojer, Impact of Bidirectional Charge Transfer and Molecular Distortions on the Electronic Structure of a Metal-Organic Interface, *Phys. Rev. Lett.*, 2007, **99**(25), 256801, DOI: 10.1103/PhysRevLett.99.256801.

212 O. T. Hofmann, V. Atalla, N. Moll, P. Rinke and M. Scheffler, Interface Dipoles of Organic Molecules on Ag(111) in Hybrid Density-Functional Theory, *New J. Phys.*, 2013, **15**(12), 123028, DOI: 10.1088/1367-2630/15/12/123028.

213 S. Duhm, G. Heimel, I. Salzmann, H. Glowatzki, R. L. Johnson, A. Vollmer, J. P. Rabe and N. Koch, Orientation-Dependent Ionization Energies and Interface Dipoles in Ordered Molecular Assemblies, *Nat. Mater.*, 2008, **7**(4), 326–332, DOI: 10.1038/nmat2119.

214 G. Rangger, O. Hofmann, L. Romaner, G. Heimel, B. Bröker, R.-P. Blum, R. Johnson, N. Koch and E. Zojer, F4TCNQ on Cu, Ag, and Au as Prototypical Example for a Strong Organic Acceptor on Coinage Metals, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **79**(16), 165306, DOI: 10.1103/PhysRevB.79.165306.

215 T. Abu-Husein, S. Schuster, D. A. Egger, M. Kind, T. Santowski, A. Wiesner, R. Chiechi, E. Zojer, A. Terfort and M. Zharnikov, The Effects of Embedded Dipoles in Aromatic Self-Assembled Monolayers, *Adv. Funct. Mater.*, 2015, **25**(25), 3943–3957, DOI: 10.1002/adfm.201500899.

216 H. Vazquez, Barrier Formation at Metal/Organic Interfaces: Dipole Formation and the Charge Neutrality Level, *Appl. Surf. Sci.*, 2004, **234**(1–4), 107–112, DOI: 10.1016/j.apsusc.2004.05.084.

8174 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

217 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids, *Phys. Rev. Lett.*, 2003, **91**(14), 146401, DOI: 10.1103/PhysRevLett.91.146401.

218 J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin and J. Sun, Workhorse Semilocal Density Functional for Condensed Matter Physics and Quantum Chemistry, *Phys. Rev. Lett.*, 2009, **103**(2), 026403, DOI: 10.1103/PhysRevLett.103.026403.

219 J. Sun, A. Ruzsinszky and J. P. Perdew, Strongly Constrained and Appropriately Normed Semilocal Density Functional, *Phys. Rev. Lett.*, 2015, **115**(3), 036402, DOI: 10.1103/PhysRevLett.115.036402.

220 A. P. Bartók and J. R. Yates, Regularized SCAN Functional, *J. Chem. Phys.*, 2019, **150**(16), 161101, DOI: 10.1063/1.5094646.

221 D. Mejia-Rodriguez and S. B. Trickey, Deorbitalization Strategies for Meta-Generalized-Gradient-Approximation Exchange-Correlation Functionals, *Phys. Rev. A*, 2017, **96**(5), 052512, DOI: 10.1103/PhysRevA.96.052512.

222 H. Peng, Z.-H. Yang, J. P. Perdew and J. Sun, Versatile van Der Waals Density Functional Based on a Meta-Generalized Gradient Approximation, *Phys. Rev. X*, 2016, **6**(4), 041005, DOI: 10.1103/PhysRevX.6.041005.

223 A. Patra and A. Ruzsinszky, Revealing the Role of van Der Waals Interactions in Thiophene Adsorption on Copper Surfaces, 2019, ArXiv190710028 Cond-Mat Physicsphysics.

224 M. Marsman, J. Paier, A. Stroppa and G. Kresse, Hybrid Functionals Applied to Extended Systems, *J. Phys.: Condens. Matter*, 2008, **20**(6), 064201, DOI: 10.1088/0953-8984/20/6/064201.

225 J. P. Perdew, M. Ernzerhof and K. Burke, Rationale for Mixing Exact Exchange with Density Functional Approximations, *J. Chem. Phys.*, 1996, **105**(22), 9982–9985, DOI: 10.1063/1.472933.

226 J. Paier, M. Marsman and G. Kresse, Why Does the B3LYP Hybrid Functional Fail for Metals?, *J. Chem. Phys.*, 2007, **127**(2), 024103, DOI: 10.1063/1.2747249.

227 N. Mardirossian and M. Head-Gordon, Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals, *Mol. Phys.*, 2017, **115**(19), 2315–2372, DOI: 10.1080/00268976.2017.1333644.

228 A. V. Krukau, O. A. Vydrov, A. F. Izmaylov and G. E. Scuseria, Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals, *J. Chem. Phys.*, 2006, **125**(22), 224106, DOI: 10.1063/1.2404663.

229 J. Heyd, G. E. Scuseria and M. Ernzerhof, Hybrid Functionals Based on a Screened Coulomb Potential, *J. Chem. Phys.*, 2003, **118**(18), 8207, DOI: 10.1063/1.1564060.

230 O. A. Vydrov, J. Heyd, A. V. Krukau and G. E. Scuseria, Importance of Short-Range versus Long-Range Hartree-Fock Exchange for the Performance of Hybrid Density Functionals, *J. Chem. Phys.*, 2006, **125**(7), 074106, DOI: 10.1063/1.2244560.

231 L. Kronik, T. Stein, S. Refaely-Abramson and R. Baer, Excitation Gaps of Finite-Sized Systems from Optimally-Tuned Range-Separated Hybrid Functionals, *J. Chem. Theory Comput.*, 2012, **8**, 1515–1531, DOI: 10.1021/ct2009363.

232 S. Refaely-Abramson, S. Sharifzadeh, N. Govind, J. Autschbach, J. Neaton, R. Baer and L. Kronik, Quasiparticle Spectra from a Nonempirical Optimally Tuned Range-Separated Hybrid Density Functional, *Phys. Rev. Lett.*, 2012, **109**(22), 226405, DOI: 10.1103/PhysRevLett.109.226405.

233 D. Jacquemin, B. Moore, A. Planchat, C. Adamo and J. Autschbach, Performance of an Optimally Tuned Range-Separated Hybrid Functional for 0–0 Electronic Excitation Energies, *J. Chem. Theory Comput.*, 2014, **10**(4), 1677–1685, DOI: 10.1021/ct5000617.

234 A. Karolewski, L. Kronik and S. Kümmel, Using Optimally Tuned Range Separated Hybrid Functionals in Ground-State Calculations: Consequences and Caveats, *J. Chem. Phys.*, 2013, **138**(20), 204115, DOI: 10.1063/1.4807325.

235 J. F. Janak, Proof That $\partial E\, \partial n\, i = \varepsilon$ in Density-Functional Theory, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1978, **18**(12), 7165–7168, DOI: 10.1103/PhysRevB.18.7165.

236 V. Atalla, M. Yoon, F. Caruso, P. Rinke and M. Scheffler, Hybrid Density Functional Theory Meets Quasiparticle Calculations: A Consistent Electronic Structure Approach., *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **88**(16), 165122, DOI: 10.1103/PhysRevB.88.165122.

237 E. Wruss, E. Zojer and O. T. Hofmann, Distinguishing between Charge-Transfer Mechanisms at Organic/Inorganic Interfaces Employing Hybrid Functionals, *J. Phys. Chem. C*, 2018, **122**(26), 14640–14653, DOI: 10.1021/acs.jpcc.8b03699.

238 A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio Jr, A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H.-Y. Ko, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang, R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhu and C. R. Groom, Report on the Sixth Blind Test of Organic Crystal Structure Prediction Methods, *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.*, 2016, **72**(4), 439–459, DOI: 10.1107/S2052520616007447.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | 8175

239 M. Rohlfing and T. Bredow, Binding Energy of Adsorbates on a Noble-Metal Surface: Exchange and Correlation Effects, *Phys. Rev. Lett.*, 2008, **101**(26), 266106, DOI: 10.1103/PhysRevLett.101.266106.

240 T. Olsen and K. S. Thygesen, Random Phase Approximation Applied to Solids, Molecules, and Graphene-Metal Interfaces: From van Der Waals to Covalent Bonding, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**(7), 075111, DOI: 10.1103/PhysRevB.87.075111.

241 J. Hermann, R. A. DiStasio and A. Tkatchenko, First-Principles Models for van Der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications, *Chem. Rev.*, 2017, **117**(6), 4714–4758, DOI: 10.1021/acs.chemrev.6b00446.

242 K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard and B. I. Lundqvist, 10.1021/Ar500118y, *Rep. Prog. Phys.*, 2015, **78**(6), 066501, DOI: 10.1088/0034-4885/78/6/066501.

243 J. G. Brandenburg and S. Grimme, Dispersion Corrected Hartree–Fock and Density Functional Theory for Organic Crystal Structure Prediction, in *Prediction and Calculation of Crystal Structures*, ed. Atahan-Evrenk, S. and Aspuru-Guzik, A., Springer International Publishing, Cham, 2013, vol. 345, pp. 1–23, DOI: 10.1007/128_2013_488.

244 S. Grimme, Density Functional Theory with London Dispersion Corrections: Density Functional Theory with London Dispersion Corrections, *Wiley Interdisc. Rev. Comput. Mol. Sci.*, 2011, **1**(2), 211–228, DOI: 10.1002/wcms.30.

245 J. Klimeš, D. R. Bowler and A. Michaelides, Van Der Waals Density Functionals Applied to Solids, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **83**(19), 195131, DOI: 10.1103/PhysRevB.83.195131.

246 K. E. Riley, M. Pitoňák, P. Jurečka and P. Hobza, Stabilization and Structure Calculations for Noncovalent Interactions in Extended Molecular Systems Based on Wave Function and Density Functional Theories, *Chem. Rev.*, 2010, **110**(9), 5023–5063, DOI: 10.1021/cr1000173.

247 T. Bučko, J. Hafner, S. Lebègue and J. G. Ángyán, Improved Description of the Structure of Molecular and Layered Crystals: Ab Initio DFT Calculations with van Der Waals Corrections, *J. Phys. Chem. A*, 2010, **114**(43), 11814–11824, DOI: 10.1021/jp106469x.

248 M. Stöhr, T. Van Voorhis and A. Tkatchenko, Theory and Practice of Modeling van Der Waals Interactions in Electronic-Structure Calculations, *Chem. Soc. Rev.*, 2019, **48**(15), 4118–4154, DOI: 10.1039/C9CS00060G.

249 S. Grimme, Accurate Description of van Der Waals Complexes by Density Functional Theory Including Empirical Corrections, *J. Comput. Chem.*, 2004, **25**(12), 1463–1473, DOI: 10.1002/jcc.20078.

250 S. Grimme, Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction, *J. Comput. Chem.*, 2006, **27**(15), 1787–1799, DOI: 10.1002/jcc.20495.

251 A. Tkatchenko and M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, *Phys. Rev. Lett.*, 2009, **102**(7), 073005, DOI: 10.1103/PhysRevLett.102.073005.

252 A. D. Becke and E. R. Johnson, A Density-Functional Model of the Dispersion Interaction, *J. Chem. Phys.*, 2005, **123**(15), 154101, DOI: 10.1063/1.2065267.

253 T. Sato and H. Nakai, Density Functional Method Including Weak Interactions: Dispersion Coefficients Based on the Local Response Approximation, *J. Chem. Phys.*, 2009, **131**(22), 224104, DOI: 10.1063/1.3269802.

254 N. Ferri, R. A. DiStasio, A. Ambrosetti, R. Car and A. Tkatchenko, Electronic Properties of Molecules and Surfaces with a Self-Consistent Interatomic van Der Waals Density Functional, *Phys. Rev. Lett.*, 2015, **114**(17), 176802, DOI: 10.1103/PhysRevLett.114.176802.

255 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu, *J. Chem. Phys.*, 2010, **132**(15), 154104, DOI: 10.1063/1.3382344.

256 E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, A Generally Applicable Atomic-Charge Dependent London Dispersion Correction, *J. Chem. Phys.*, 2019, **150**(15), 154122, DOI: 10.1063/1.5090222.

257 A. D. Becke and E. R. Johnson, Exchange-Hole Dipole Moment and the Dispersion Interaction, *J. Chem. Phys.*, 2005, **122**(15), 154104, DOI: 10.1063/1.1884601.

258 J. Hermann and A. Tkatchenko, Density Functional Model for van Der Waals Interactions: Unifying Many-Body Atomic Approaches with Nonlocal Functionals, *Phys. Rev. Lett.*, 2020, **124**(14), 146401, DOI: 10.1103/PhysRevLett.124.146401.

259 J. Řezáč, Y. Huang, P. Hobza and G. J. O. Beran, Benchmark Calculations of Three-Body Intermolecular Interactions and the Performance of Low-Cost Electronic Structure Methods, *J. Chem. Theory Comput.*, 2015, **11**(7), 3065–3079, DOI: 10.1021/acs.jctc.5b00281.

260 S. R. Kachel, B. P. Klein, J. M. Morbec, M. Schöniger, M. Hutter, M. Schmid, P. Kratzer, B. Meyer, R. Tonner and J. M. Gottfried, Chemisorption and Physisorption at the Metal/Organic Interface: Bond Energies of Naphthalene and Azulene on Coinage Metal Surfaces, *J. Phys. Chem. C*, 2020, **124**(15), 8257–8268, DOI: 10.1021/acs.jpcc.0c00915.

261 E. Schiavo, A. B. Muñoz-García, V. Barone, A. Vittadini, M. Casarin, D. Forrer and M. Pavone, Tuning Dispersion Correction in DFT-D2 for Metal-Molecule Interactions: A Tailored Reparameterization Strategy for the Adsorption of Aromatic Systems on Ag(1 1 1), *Chem. Phys. Lett.*, 2018, **693**, 28–33, DOI: 10.1016/j.cplett.2018.01.005.

262 S. Ehrlich, J. Moellmann, W. Reckien, T. Bredow and S. Grimme, System-Dependent Dispersion Coefficients for the DFT-D3 Treatment of Adsorption Processes on Ionic Surfaces, *ChemPhysChem*, 2011, **12**(17), 3414–3420, DOI: 10.1002/cphc.201100521.

263 G.-X. Zhang, A. Tkatchenko, J. Paier, H. Appel and M. Scheffler, Van Der Waals Interactions in Ionic and

8176 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

Semiconductor Solids, *Phys. Rev. Lett.*, 2011, **107**(24), 245501, DOI: 10.1103/PhysRevLett.107.245501.

264 R. J. Maurer, W. Liu, I. Poltavsky, T. Stecher, H. Oberhofer, K. Reuter and A. Tkatchenko, Thermal and Electronic Fluctuations of Flexible Adsorbed Molecules: Azobenzene on Ag(111), *Phys. Rev. Lett.*, 2016, **116**(14), 146101, DOI: 10.1103/PhysRevLett.116.146101.

265 V. G. Ruiz, W. Liu and A. Tkatchenko, Density-Functional Theory with Screened van Der Waals Interactions Applied to Atomic and Molecular Adsorbates on Close-Packed and Non-Close-Packed Surfaces, *Phys. Rev. B*, 2016, **93**(3), 035118, DOI: 10.1103/PhysRevB.93.035118.

266 E. Caldeweyher, C. Bannwarth and S. Grimme, Extension of the D3 Dispersion Coefficient Model, *J. Chem. Phys.*, 2017, **147**(3), 034112, DOI: 10.1063/1.4993215.

267 P. Bultinck, C. Van Alsenoy, P. W. Ayers and R. Carbó-Dorca, Critical Analysis and Extension of the Hirshfeld Atoms in Molecules, *J. Chem. Phys.*, 2007, **126**(14), 144111, DOI: 10.1063/1.2715563.

268 T. Bučko, S. Lebègue, J. G. Ángyán and J. Hafner, Extending the Applicability of the Tkatchenko–Scheffler Dispersion Correction via Iterative Hirshfeld Partitioning, *J. Chem. Phys.*, 2014, **141**(3), 034114, DOI: 10.1063/1.4890003.

269 T. Gould, S. Lebègue, J. G. Ángyán and T. Bučko, A Fractionally Ionic Approach to Polarizability and van Der Waals Many-Body Dispersion Calculations, *J. Chem. Theory Comput.*, 2016, **12**(12), 5920–5930, DOI: 10.1021/acs.jctc.6b00925.

270 P. J. Blowey, B. Sohail, L. A. Rochford, T. Lafosse, D. A. Duncan, P. T. P. Ryan, D. A. Warr, T.-L. Lee, G. Costantini, R. J. Maurer and D. P. Woodruff, Alkali Doping Leads to Charge-Transfer Salt Formation in a Two-Dimensional Metal–Organic Framework, *ACS Nano*, 2020, acsnano.0c03133, DOI: 10.1021/acsnano.0c03133.

271 A. Tkatchenko, R. A. DiStasio, R. Car and M. Scheffler, Accurate and Efficient Method for Many-Body van Der Waals Interactions, *Phys. Rev. Lett.*, 2012, **108**(23), 236402, DOI: 10.1103/PhysRevLett.108.236402.

272 A. Tkatchenko, A. Ambrosetti and R. A. DiStasio, Interatomic Methods for the Dispersion Energy Derived from the Adiabatic Connection Fluctuation-Dissipation Theorem, *J. Chem. Phys.*, 2013, **138**(7), 074106, DOI: 10.1063/1.4789814.

273 R. A. DiStasio, V. V. Gobre and A. Tkatchenko, Many-Body van Der Waals Interactions in Molecules and Condensed Matter, *J. Phys.: Condens. Matter*, 2014, **26**(21), 213202, DOI: 10.1088/0953-8984/26/21/213202.

274 R. J. Maurer, V. G. Ruiz and A. Tkatchenko, Many-Body Dispersion Effects in the Binding of Adsorbates on Metal Surfaces, *J. Chem. Phys.*, 2015, **143**(10), 102808, DOI: 10.1063/1.4922688.

275 O. A. Vydrov and T. Van Voorhis, Dispersion Interactions from a Local Polarizability Model, *Phys. Rev. A*, 2010, **81**(6), 062708, DOI: 10.1103/PhysRevA.81.062708.

276 H. Rydberg, M. Dion, N. Jacobson, E. Schröder, P. Hyldgaard, S. I. Simak, D. C. Langreth and B. I. Lundqvist, Van Der Waals Density Functional for Layered Structures, *Phys. Rev. Lett.*, 2003, **91**(12), 126402, DOI: 10.1103/PhysRevLett.91.126402.

277 M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist, Van Der Waals Density Functional for General Geometries, *Phys. Rev. Lett.*, 2004, **92**(24), 246401, DOI: 10.1103/PhysRevLett.92.246401.

278 O. A. Vydrov and T. Van Voorhis, Nonlocal van Der Waals Density Functional Made Simple, *Phys. Rev. Lett.*, 2009, **103**(6), 063004, DOI: 10.1103/PhysRevLett.103.063004.

279 O. A. Vydrov and T. Van Voorhis, Nonlocal van Der Waals Density Functional: The Simpler the Better, *J. Chem. Phys.*, 2010, **133**(24), 244103, DOI: 10.1063/1.3521275.

280 K. Berland and P. Hyldgaard, Exchange Functional That Tests the Robustness of the Plasmon Description of the van Der Waals Density Functional, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2014, **89**(3), 035412, DOI: 10.1103/PhysRevB.89.035412.

281 K. Berland, C. A. Arter, V. R. Cooper, K. Lee, B. I. Lundqvist, E. Schröder, T. Thonhauser and P. Hyldgaard, van der Waals Density Functionals Built upon the Electron-Gas Tradition: Facing the Challenge of Competing Interactions, *J. Chem. Phys.*, 2014, **140**(18), 18A539, DOI: 10.1063/1.4871731.

282 L. Hörmann, A. Jeindl and O. T. Hofmann, Reproducibility of Potential Energy Surfaces of Organic/Metal Interfaces on the Example of PTCDA on Ag(111), *J. Chem. Phys.*, 2020, **153**(10), 104701, DOI: 10.1063/5.0020736.

283 F. S. Tautz, Structure and Bonding of Large Aromatic Molecules on Noble Metal Surfaces: The Example of PTCDA, *Prog. Surf. Sci.*, 2007, **82**(9–12), 479–520, DOI: 10.1016/j.progsurf.2007.09.001.

284 A. Hauschild, K. Karki, B. C. C. Cowie, M. Rohlfing, F. S. Tautz and M. Sokolowski, Molecular Distortions and Chemical Bonding of a Large π-Conjugated Molecule on a Metal Surface, *Phys. Rev. Lett.*, 2005, **94**(3), 036106, DOI: 10.1103/PhysRevLett.94.036106.

285 S. K. M. Henze, O. Bauer, T.-L. Lee, M. Sokolowski and F. S. Tautz, Vertical Bonding Distances of PTCDA on Au(111) and Ag(111): Relation to the Bonding Type, *Surf. Sci.*, 2007, **601**(6), 1566–1573, DOI: 10.1016/j.susc.2007.01.020.

286 R. Tonner, P. Rosenow and P. Jakob, Molecular Structure and Vibrations of NTCDA Monolayers on Ag(111) from Density-Functional Theory and Infrared Absorption Spectroscopy, *Phys. Chem. Chem. Phys.*, 2016, **18**(8), 6316–6328, DOI: 10.1039/C5CP06619K.

287 H. Hellmann, A New Approximation Method in the Problem of Many Electrons, *J. Chem. Phys.*, 1935, **3**(1), 61, DOI: 10.1063/1.1749559.

288 W. E. Pickett, Pseudopotential Methods in Condensed Matter Applications, *Comput. Phys. Rep.*, 1989, **9**(3), 115–197, DOI: 10.1016/0167-7977(89)90002-6.

289 P. E. Blöchl, Projector Augmented-Wave Method, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **50**(24), 17953–17979, DOI: 10.1103/PhysRevB.50.17953.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | 8177

290 D. J. Singh and L. Nordström, *Planewaves, Pseudopotentials, and the LAPW Method*, Springer, New York, NY, 2nd edn, 2006.

291 J. J. Mortensen, L. B. Hansen and K. W. Jacobsen, Real-Space Grid Implementation of the Projector Augmented Wave Method, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2005, **71**(3), 035109, DOI: 10.1103/PhysRevB.71.035109.

292 F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester, England, Hoboken, NJ, 2nd edn, 2007.

293 X. Andrade, D. Strubbe, U. De Giovannini, A. H. Larsen, M. J. T. Oliveira, J. Alberdi-Rodriguez, A. Varas, I. Theophilou, N. Helbig, M. J. Verstraete, L. Stella, F. Nogueira, A. Aspuru-Guzik, A. Castro, M. A. L. Marques and A. Rubio, Real-Space Grids and the Octopus Code as Tools for the Development of New Simulation Approaches for Electronic Systems, *Phys. Chem. Chem. Phys.*, 2015, **17**(47), 31371–31396, DOI: 10.1039/C5CP00351B.

294 J. R. Chelikowsky, N. Troullier and Y. Saad, Finite-Difference-Pseudopotential Method: Electronic Structure Calculations without a Basis, *Phys. Rev. Lett.*, 1994, **72**(8), 1240–1243, DOI: 10.1103/PhysRevLett.72.1240.

295 K. Lejaeghere, G. Bihlmayer, T. Bjorkman, P. Blaha, S. Blugel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, S. de Gironcoli, T. Deutsch, J. K. Dewhurst, I. Di Marco, C. Draxl, M. Du ak, O. Eriksson, J. A. Flores-Livas, K. F. Garrity, L. Genovese, P. Giannozzi, M. Giantomassi, S. Goedecker, X. Gonze, O. Granas, E. K. U. Gross, A. Gulans, F. Gygi, D. R. Hamann, P. J. Hasnip, N. A. W. Holzwarth, D. Iu an, D. B. Jochym, F. Jollet, D. Jones, G. Kresse, K. Koepernik, E. Kucukbenli, Y. O. Kvashnin, I. L. M. Locht, S. Lubeck, M. Marsman, N. Marzari, U. Nitzsche, L. Nordstrom, T. Ozaki, L. Paulatto, C. J. Pickard, W. Poelmans, M. I. J. Probert, K. Refson, M. Richter, G.-M. Rignanese, S. Saha, M. Scheffler, M. Schlipf, K. Schwarz, S. Sharma, F. Tavazza, P. Thunstrom, A. Tkatchenko, M. Torrent, D. Vanderbilt, M. J. van Setten, V. Van Speybroeck, J. M. Wills, J. R. Yates, G.-X. Zhang and S. Cottenier, Reproducibility in Density Functional Theory Calculations of Solids, *Science*, 2016, **351**(6280), aad3000, DOI: 10.1126/science.aad3000.

296 J. Tersoff and D. R. Hamann, Theory and Application for the Scanning Tunneling Microscope, *Phys. Rev. Lett.*, 1983, **50**(25), 1998–2001, DOI: 10.1103/PhysRevLett.50.1998.

297 J. Tersoff and D. R. Hamann, Theory of the Scanning Tunneling Microscope, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1985, **31**(2), 805–813, DOI: 10.1103/PhysRevB.31.805.

298 G. Kresse and D. Joubert, From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**(3), 1758–1775, DOI: 10.1103/PhysRevB.59.1758.

299 G. Kresse and J. Furthmüller, Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**(16), 11169–11186, DOI: 10.1103/PhysRevB.54.11169.

300 F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt and J. H. van Lenthe, State of the Art in Counterpoise Theory, *Chem. Rev.*, 1994, **94**(7), 1873–1885, DOI: 10.1021/cr00031a007.

301 S. F. Boys and F. Bernardi, The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors, *Mol. Phys.*, 1970, **19**(4), 553–566, DOI: 10.1080/00268977000101561.

302 Ł. M. Mentel and E. J. Baerends, Can the Counterpoise Correction for Basis Set Superposition Effect Be Justified?, *J. Chem. Theory Comput.*, 2014, **10**(1), 252–267, DOI: 10.1021/ct400990u.

303 N. D. Woods, M. C. Payne and P. J. Hasnip, Computing the Self-Consistent Field in Kohn–Sham Density Functional Theory, *J. Phys.: Condens. Matter*, 2019, **31**(45), 453001, DOI: 10.1088/1361-648X/ab31c0.

304 E. Sauter, G. Nascimbeni, D. Trefz, S. Ludwigs, E. Zojer, F. von Wrochem and M. Zharnikov, A Dithiocarbamate Anchoring Group as a Flexible Platform for Interface Engineering, *Phys. Chem. Chem. Phys.*, 2019, **21**, 22511–22525, DOI: 10.1039/C9CP03306H.

305 G. J. Kraberger, D. A. Egger and E. Zojer, Tuning the Electronic Structure of Graphene through Collective Electrostatic Effects, *Adv. Mater. Interfaces*, 2015, **2**(17), 1500323, DOI: 10.1002/admi.201500323.

306 V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler, Ab Initio Molecular Simulations with Numeric Atom-Centered Orbitals, *Comput. Phys. Commun.*, 2009, **180**(11), 2175–2196, DOI: 10.1016/j.cpc.2009.06.022.

307 J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dułak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen, H. H. Kristoffersen, M. Kuisma, A. H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P. G. Moses, J. Ojanen, T. Olsen, V. Petzold, N. A. Romero, J. Stausholm-Møller, M. Strange, G. A. Tritsaris, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G. K. H. Madsen, R. M. Nieminen, J. K. Nørskov, M. Puska, T. T. Rantala, J. Schiøtz, K. S. Thygesen and K. W. Jacobsen, Electronic Structure Calculations with GPAW: A Real-Space Implementation of the Projector Augmented-Wave Method., *J. Phys.: Condens. Matter*, 2010, **22**(25), 253202, DOI: 10.1088/0953-8984/22/25/253202.

308 E. Wruss, G. Prokopiou, L. Kronik, E. Zojer, O. T. Hofmann and D. A. Egger, Magnetic Configurations of Open-Shell Molecules on Metals: The Case of CuPc and CoPc on Silver, *Phys. Rev. Mater.*, 2019, **3**(8), 086002, DOI: 10.1103/PhysRevMaterials.3.086002.

309 R. Hoffmann, An Extended Hückel Theory. I. Hydrocarbons, *J. Chem. Phys.*, 1963, **39**(6), 1397–1412, DOI: 10.1063/1.1734456.

310 J. Harris, Simplified Method for Calculating the Energy of Weakly Interacting Fragments, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1985, **31**(4), 1770–1779, DOI: 10.1103/PhysRevB.31.1770.

8178 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 8132–8180

This journal is © the Owner Societies 2021

311 J. H. Van Lenthe, R. Zwaans, H. J. J. Van Dam and M. F. Guest, Starting SCF Calculations by Superposition of Atomic Densities, *J. Comput. Chem.*, 2006, **27**(8), 926–932, DOI: 10.1002/jcc.20393.

312 J. Almlöf, K. Faegri and K. Korsell, Principles for a Direct SCF Approach to LICAO-MO Ab-Initio Calculations: Direct SCF Approach, *J. Comput. Chem.*, 1982, **3**(3), 385–399, DOI: 10.1002/jcc.540030314.

313 X. Li, F. S. Curtis, T. Rose, C. Schober, A. Vazquez-Mayagoitia, K. Reuter, H. Oberhofer and N. Marom, Genarris: Random Generation of Molecular Crystal Structures and Fast Screening with a Harris Approximation., *J. Chem. Phys.*, 2018, **148**(24), 241701, DOI: 10.1063/1.5014038.

314 A. J. Cohen, P. Mori-Sánchez and W. Yang, Fractional Spins and Static Correlation Error in Density Functional Theory, *J. Chem. Phys.*, 2008, **129**(12), 121104, DOI: 10.1063/1.2987202.

315 A. J. Cohen, P. Mori-Sanchez and W. Yang, Insights into Current Limitations of Density Functional Theory, *Science*, 2008, **321**(5890), 792–794, DOI: 10.1126/science.1158722.

316 P. Pulay, ImprovedSCF Convergence Acceleration, *J. Comput. Chem.*, 1982, **3**(4), 556–560, DOI: 10.1002/jcc.540030413.

317 C. G. Broyden, A Class of Methods for Solving Nonlinear Simultaneous Equations, *Math. Comput.*, 1965, **19**(92), 577, DOI: 10.1090/S0025-5718-1965-0198670-6.

318 P. Pulay, Convergence Acceleration of Iterative Sequences. the Case of Scf Iteration, *Chem. Phys. Lett.*, 1980, **73**(2), 393–398, DOI: 10.1016/0009-2614(80)80396-4.

319 E. B. Stechel, A. R. Williams and P. J. Feibelman, *N*-Scaling Algorithm for Density-Functional Calculations of Metals and Insulators, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **49**(15), 10088–10101, DOI: 10.1103/PhysRevB.49.10088.

320 G. P. Kerker, Efficient Iteration Scheme for Self-Consistent Pseudopotential Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1981, **23**(6), 3082–3084, DOI: 10.1103/PhysRevB.23.3082.

321 D. D. Johnson, Modified Broyden's Method for Accelerating Convergence in Self-Consistent Calculations, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **38**(18), 12807–12813, DOI: 10.1103/PhysRevB.38.12807.

322 M. P. Teter, M. C. Payne and D. C. Allan, Solution of Schrödinger's Equation for Large Systems, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1989, **40**(18), 12255–12263, DOI: 10.1103/PhysRevB.40.12255.

323 P. E. Blöchl, O. Jepsen and O. K. Andersen, Improved Tetrahedron Method for Brillouin-Zone Integrations, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **49**(23), 16223–16233, DOI: 10.1103/PhysRevB.49.16223.

324 M. Methfessel and A. T. Paxton, High-Precision Sampling for Brillouin-Zone Integration in Metals, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1989, **40**(6), 3616–3621, DOI: 10.1103/PhysRevB.40.3616.

325 M. J. Gillan, Calculation of the Vacancy Formation Energy in Aluminium, *J. Phys.: Condens. Matter*, 1989, **1**(4), 689–711, DOI: 10.1088/0953-8984/1/4/005.

326 A. D. Vita and M. J. Gillan, The Energetics of Hydrogen in Aluminium Calculated from First Principles, *J. Phys.: Condens. Matter*, 1992, **4**(2), 599–611, DOI: 10.1088/0953-8984/4/2/028.

327 *Modern Methods of Crystal Structure Prediction*, ed. Oganov, A. R., Wiley-VCH, Weinheim, Germany, 2011.

328 D. P. Woodruff and D. A. Duncan, X-Ray Standing Wave Studies of Molecular Adsorption: Why Coherent Fractions Matter, *New J. Phys.*, 2020, **22**(11), 113012, DOI: 10.1088/1367-2630/abc63a.

329 F. Birch, Finite Elastic Strain of Cubic Crystals, *Phys. Rev.*, 1947, **71**(11), 809–824, DOI: 10.1103/PhysRev.71.809.

330 C. G. Broyden, The Convergence of a Class of Double-Rank Minimization Algorithms 1. General Considerations, *IMA J. Appl. Math.*, 1970, **6**(1), 76–90, DOI: 10.1093/imamat/6.1.76.

331 R. Fletcher, A New Approach to Variable Metric Algorithms, *Comput. J.*, 1970, **13**(3), 317–322, DOI: 10.1093/comjnl/13.3.317.

332 D. Goldfarb, A Family of Variable-Metric Methods Derived by Variational Means, *Math. Comput.*, 1970, **24**(109), 23, DOI: 10.1090/S0025-5718-1970-0258249-6.

333 D. F. Shanno, Conditioning of Quasi-Newton Methods for Function Minimization, *Math. Comput.*, 1970, **24**(111), 647, DOI: 10.1090/S0025-5718-1970-0274029-X.

334 P. Császár and P. Pulay, Geometry Optimization by Direct Inversion in the Iterative Subspace, *J. Mol. Struct.*, 1984, **114**, 31–34, DOI: 10.1016/S0022-2860(84)87198-7.

335 G. Fogarasi, X. Zhou, P. W. Taylor and P. Pulay, The Calculation of Ab Initio Molecular Geometries: Efficient Optimization by Natural Internal Coordinates and Empirical Correction by Offset Forces, *J. Am. Chem. Soc.*, 1992, **114**(21), 8191–8201, DOI: 10.1021/ja00047a032.

336 J. Baker, Techniques for Geometry Optimization: A Comparison of Cartesian and Natural Internal Coordinates, *J. Comput. Chem.*, 1993, **14**(9), 1085–1100, DOI: 10.1002/jcc.540140910.

337 T. Bučko, J. Hafner and J. G. Ángyán, Geometry Optimization of Periodic Systems Using Internal Coordinates, *J. Chem. Phys.*, 2005, **122**(12), 124508, DOI: 10.1063/1.1864932.

338 K. N. Kudin, G. E. Scuseria and H. B. Schlegel, A Redundant Internal Coordinate Algorithm for Optimization of Periodic Systems, *J. Chem. Phys.*, 2001, **114**(7), 2919–2923, DOI: 10.1063/1.1340578.

339 J. Andzelm, R. D. King-Smith and G. Fitzgerald, Geometry Optimization of Solids Using Delocalized Internal Coordinates, *Chem. Phys. Lett.*, 2001, **335**(3–4), 321–326, DOI: 10.1016/S0009-2614(01)00030-6.

340 A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson,

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 8132–8180 | 8179

C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, The Atomic Simulation Environment—a Python Library for Working with Atoms, *J. Phys.: Condens. Matter*, 2017, **29**(27), 273002, DOI: 10.1088/1361-648X/aa680e.

341 R. Lindh, A. Bernhardsson, G. Karlström and P.-Å. Malmqvist, On the Use of a Hessian Model Function in Molecular Geometry Optimizations, *Chem. Phys. Lett.*, 1995, **241**(4), 423–428, DOI: 10.1016/0009-2614(95)00646-L.

342 T. H. Fischer and J. Almlof, General Methods for Geometry and Wave Function Optimization, *J. Phys. Chem.*, 1992, **96**(24), 9768–9774, DOI: 10.1021/j100203a036.

343 E. Garijo del Río, J. J. Mortensen and K. W. Jacobsen, Local Bayesian Optimizer for Atomic Structures, *Phys. Rev. B*, 2019, **100**(10), 104103, DOI: 10.1103/PhysRevB.100.104103.

344 J. Järvi, P. Rinke and M. Todorović, Detecting Stable Adsorbates of (1*S*)-Camphor on Cu(111) with Bayesian Optimization, *Beilstein J. Nanotechnol.*, 2020, **11**, 1577–1589, DOI: 10.3762/bjnano.11.140.

345 H. Glowatzki, B. Bröker, R.-P. Blum, O. T. Hofmann, A. Vollmer, R. Rieger, K. Müllen, E. Zojer, J. P. Rabe and N. Koch, "Soft" Metallic Contact to Isolated C$_{60}$ Molecules, *Nano Lett.*, 2008, **8**(11), 3825–3829, DOI: 10.1021/nl8021797.

346 J. Muscat and D. Newns, Simple Model of Hydrogen and Lithium Chemisorption on Jellium Substrates, *Phys. Rev. B*, 1979, **19**(2), 1270–1282, DOI: 10.1103/PhysRevB.19.1270.

347 T. Körzdörfer, R. M. Parrish, J. S. Sears, C. D. Sherrill and J.-L. Brédas, On the Relationship between Bond-Length Alternation and Many-Electron Self-Interaction Error, *J. Chem. Phys.*, 2012, **137**(12), 124305, DOI: 10.1063/1.4752431.

348 E. Wruss, L. Hörmann and O. T. Hofmann, Impact of Surface Defects on the Charge Transfer at Inorganic/Organic Interfaces, *J. Phys. Chem. C*, 2019, **123**, 7118–7124, DOI: 10.1021/acs.jpcc.8b11403.

349 J. Järvi, P. Rinke and M. Todorović, Detecting stable adsorbates of (1 *S*) -camphor on Cu(111) with Bayesian optimization, *Beilstein J. Nanotechnol.*, 2020, **11**, 1577–1589, DOI: 10.3762/bjnano.11.140.