Ab initio study of a mechanically gated molecule: From weak to strong correlation

A. Greuling, 1 M. Rohlfing, 1, 2 R. Temirov, 2 F. S. Tautz, 2 and F. B. Anders 3

1Fachbereich Physik, Universität Osnabrück, Barbarastrasse 7, 49069 Osnabrück, Germany
2Peter Grünberg Institut (PGI-3), Forschungszentrum Jülich, 52425 Jülich, Germany and JARA-Fundamentals of Future Information Technology, 52425 Jülich, Germany
3Fakultät für Physik, TU Dortmund, Otto-Hahn-Straße 4, 44227 Dortmund, Germany

(Received 26 July 2011; published 6 September 2011)

DOI: 10.1103/PhysRevB.84.125413 PACS number(s): 73.20.Hb, 71.15.—m, 73.22.Dj

I. INTRODUCTION

The control of the geometry of a molecular junction and the manipulation of its electronic structure by an external parameter are of central importance for molecular electronics. So far, the best control over the junction structure is achieved by a scanning tunnelling microscope (STM), 1, 2 which allows to select an individual molecule in a specific environment and to contact it at a defined position within the molecule (see e.g., Ref. 2). On the other hand, the most common approach to tuning the electronic structure, electrical gating, is difficult to combine with STM. However, in a recent experimental study, 2 we contacted a surface-adsorbed molecule with an STM tip and peeled it off the surface by tip retraction, as shown schematically in Fig. 1. Spectroscopic data recorded during tip retraction revealed a mechanical gating effect, i.e., one of the molecular orbitals responds to the structural change and shifts with respect to the Fermi level (E F ) of the substrate, before becoming pinned at E F [see Fig. 2(d)].

In this paper, we use this experiment, which was carried out on 3,4,9,10-perylene-tetracarboxylic acid-dianhydride (PTCDA) adsorbed on the Ag(111) surface, as a motivation for a theoretical study of the interplay between the geometric structure and the electronic spectrum of a molecular junction. The mechanical gating is subject to subtle details of electronic correlation, e.g., screening of the intramolecular Coulomb repulsion by the electrodes, which is not described by conventional density-functional theory (DFT). We, therefore, propose the following strategy: DFT addresses the atomistic details of the junction structure but does not provide reliable, well founded spectral data. To evaluate the electronic spectrum, we combine DFT with many-body perturbation theory (MBPT) 7 for nonlocal correlation and with numerical renormalization-group (NRG) theory 8 for correlation dynamics beyond the mean-field level. Among these three approaches, MBPT constitutes the essential link between DFT and NRG. Most NRG studies so far have been based on empirical model parameters. MBPT spectra, on the other hand, provide better founded NRG input data.

The interrelation between the structure of the junction and its transport spectra was studied in Ref. 9 in terms of experiment and transport theory disregarding the correlation mechanisms (will be discussed below). One conclusion of Ref. 9 was that further many-body correlation effects are mandatory to arrive at a conclusive understanding of the electronic structure of the interface. This is the starting point for our present work.

One focus in the present paper is on the realization of MBPT for large and complex systems such as, for example, the tip/PTCDA/Ag junction, which combines metallic with nonmetallic screening and shows partial occupancy of the most important molecular level. For such systems, the commonly employed GW approximation of MBPT requires further simplification. Here, we suggest “LDA + GdW” (see below) as an efficient and systematic implementation of MBPT. The simplifications used in LDA + GdW are justified by a subsequent evaluation of NRG spectra, which are directly compared to the measurements in Secs. III and IV.

II. EXPERIMENTAL AND THEORETICAL METHODS

A. Theory

Our calculations of the structure are based on conventional density-functional theory (DFT) for the electronic ground state. Previous studies of PTCDA/Ag(111) 10–16 showed that DFT employing the local-density approximation (LDA) yields reliable structural data in good agreement with available experiments 11, 14 and with more elaborate many-body total-energy approaches, 16 although many details of long-range correlation are missing from the DFT-LDA. We thus employ the LDA for all DFT calculations throughout this work.

The DFT-LDA calculations are carried out within the SIESTA package. 19 We use a double-zeta basis set with two s, two p, and one d orbital for carbon as well as for oxygen, two s and one p orbital for hydrogen, and two s, one p, and two d orbitals for silver. By using a very small contraction energy shift of 0.002 Ry in the preparation of the orbitals, 19 we achieve a spatial truncation of the most extended orbitals at 3.57, 2.81, 3.45, and 4.54 Å for C, O, H, and Ag. Such large extent of the orbitals turned out to be crucial to obtain reliable, converged results. 16 The potential and the charge density are represented by a Fourier series with a cutoff energy of 200 Ry.

Our system setup consists of three layers of silver (six layers for the evaluation of the spectra), a single molecule, and a tip of ten silver atoms in the shape of a pyramid (attached to more
silver for the evaluation of the spectra). Choosing Ag as the tip material is motivated by the experimental tip preparation method (see below). Tungsten as the tip material leads to equivalent results as discussed here (apart from somewhat stronger tip-molecule bond energy for tungsten).5

The LDA spectrum of the LUMO in contact with the substrate results from projection of all states of the respective system (\(|\psi_n\rangle\) onto \(|\psi_{\text{LUMO}}\rangle\), i.e.,

\[
f^{\text{LDA}}(E) = \sum_n \langle \psi_n | \psi_{\text{LUMO}} \rangle^2 \delta(E - E^{\text{LDA}}_n)
\]

[projected density of states (PDOS)]. The occupancy of the state is \(n = f^{\text{LDA}}_n f(E) dE\).

A DFT-LDA spectrum should not be interpreted as a realistic mean-field spectrum since it lacks physical significance, essentially being an auxiliary quantity within a ground-state total-energy approach. A proper single-particle mean-field spectrum which requires much less numerical effort than a full GW calculation.22 This reduced numerical effort enables to investigate the current metal–molecule hybrid system within MBPT.

For PtCDA on Ag(111), our LDA + GdW method involves a further approximation, i.e., the evaluation of \(G\) from molecular states only (i.e., from a calculation without substrate and tip), disregarding the hybridisation with substrate states. This is justified by the fact that the molecular states are only slightly shifted and broadened due to the contact with the substrate.

For the subsequent use in the NRG theory, the bare LUMO level positions \(E_{\text{LUMO}}(z, E)\) and coupling functions \(\Gamma(E, z)\) from the LDA + GdW spectra at each \(z\) in the NRG are extracted from the mean-field spectrum \(f^{\text{GdW}}(E, z)\) of the molecule–metal
contact by equating

\[ f^{GW}(E, z) = \frac{1}{\pi} \frac{1}{E - \delta(z) - \text{e} U(z) - \Gamma(E, z)} \]  

with \( f^{GW}(E, z) = f^{LDA}(E - \Delta^{GW}(z)) \) being the PDOS of Eq. (1) including the QP correction \( \Delta^{GW} \).

\section*{B. Experiments}

The experiments are carried out at 5 K in a CRE-TEC low-temperature scanning tunneling microscope on the PTCDAGAg(111) system. The Ag(111) surface is prepared by cycles of ion bombardment (0.8 keV \text{Ar}^+\) ions) and annealing at 550 °C. PTCDAGAg(111) surface is deposited onto the room temperature Ag(111) surface from a self-built evaporator. Electrochemically etched tungsten tips are used. After standard tip cleaning procedures in vacuum, the tip is dipped repeatedly into the Ag surface, until its density of states appears featureless; this is controlled by recording spectra of the Ag(111) surface state. Most likely, the tip apex is therefore a silver cluster.

The calibration of the absolute \( z \) scale in our experiments proceeds as follows: \( (1) \) the prepared tip is stabilized above the center of a PTCDAG molecule at a set point with a current of \( I = 0.1 \text{ nA} \) at a bias voltage of \( V_b = 0.34 \text{ V} \). \( (2) \) The tip is moved into contact with the molecule, discernible by a sudden deviation from the exponential approach curve that is characteristic of vacuum tunneling. \( (24) \) Because the tip is approached in the center of the molecule above the carbon skeleton, no jump into contact and no chemical bonding take place. We find that contact appears after approaching by \( \Delta z = 4.3 \text{ Å} \) from the set point. \( (3) \) The \( z \) coordinate of the tip at the set point is calculated by \( z_{\text{set point}} = h_{\text{PTCDAG}} + r_{\text{carbon vdW}} + r_{\text{silver vdW}} + \Delta z \). Here, \( h_{\text{PTCDAG}} = 2.86 \text{ Å} \) is the height of the plane of the PTCDAG molecule above the outermost lattice plane of the Ag(111) surface, \( (11) \) and \( r_{\text{carbon vdW}} = 1.70 \text{ Å} \) and \( r_{\text{silver vdW}} = 1.72 \text{ Å} \) are the van der Waals radii of carbon and silver atoms, respectively. \( (25) \) From the above equation we find that the vertical distance between the center of the tip apex atom at the set point and the outermost lattice plane of the Ag(111) surface is \( z_{\text{set point}} = 10.6 \text{ Å} \). \( (4) \) All experimental tip-surface distances \( z \) in Fig. 2 are referenced to \( z_{\text{set point}} \), using the known and calibrated \( z \)-piezo constants of our microscope.

Experimental data in this paper is presented as a two-dimensional map in which \( dI/dV \) is recorded nearly continuously as a function of energy \( E - E_F = e V_b \) (bias voltage step is 1 meV) and tip-surface distance \( z \) (step is 0.0076 Å). The map is recorded by rapid approach-retraction cycles of the tip (time for one cycle is approximately 2.4 s), i.e., the fast scanning variable in the \( (E - E_F, z) \) plane is the tip-surface distance \( z \). The closest point of the approach-retraction cycle is chosen such that at this distance the tip-molecule bond forms spontaneously. \( dI/dV \) was detected with a lock-in amplifier, with modulation amplitude 4 mV and modulation frequency 2.9 kHz. For bias voltages exceeding \( \pm 100 \text{ meV} \), the tip-PTCDAG contact becomes unstable.

\section*{III. MEAN-FIELD SPECTRUM FROM MBPT}

We first discuss the geometric structure of the tip/PTCDAGAg(111) junction to provide a realistic basis for studying spectra. We address the structure within DFT-LDA. PTCDAG adsorbs on the Ag(111) surface in a flat-lying configuration. \( (11, 14, 15, 17, 18, 26) \) Several DFT studies of the adsorption have been published (see, e.g., Refs. 16, 18, and references therein). Following experiment, \( (2) \) we place the tip, which in our simulation consists of ten Ag atoms in a pyramidal shape (see Sec. II A), above one of the carboxylic oxygen atoms and approach it to the surface. The oxygen atom forms a covalent bond with the tip apex atom (with bond length of 2.16 Å). The tip can then be moved up and down reversibly, forcing the oxygen atom and the attached section of the molecule to follow, resulting in a peeling-like motion [see Figs. 1(a) and 1(b)]. Throughout the paper, \( z \) specifies the vertical distance between the tip apex atom and the uppermost surface layer. Some representative structures of the junction are displayed in Fig. 1. At a tip-surface distance of \( z = 7 \text{ Å} \) only a small section of the molecule is detached from the surface, while at \( z = 10 \text{ Å} \) about half of the molecule has lost contact with the substrate. The structural configurations are the same as were used in Ref. 9 for transport calculations.

The most interesting feature of the electronic structure of the junction is the PTCDAG LUMO (lowest unoccupied molecular orbital) because this is the orbital for which the gating effect as a function of the external parameter \( z \) is observed. \( (2, 9) \) In the gas phase, the LUMO is found \( \sim 2 \text{ eV} \) above the Ag(111) Fermi level [see Fig. 2(a)]. Upon adsorption, the LUMO is lowered in energy below the Ag(111) Fermi level due to the metallic polarizability of the substrate, \( (22, 29, 30) \) and thus becomes partially occupied with about 1.8 electrons (partially back-balanced by back-donation from other molecular states to the surface).

The partial occupation of the LUMO (which changes as a function of \( z \); see below) causes the high sensitivity of the LUMO spectrum to the junction structure: in experiment, for \( z \) below 8.2 Å the LUMO approaches the Fermi energy at a steep slope of \( \sim 0.2-0.3 \text{ eV/Å} \) and turns into a sharp resonance as soon as it reaches \( E_F \) at \( z = 8.2 \text{ Å} \). At \( z = 9.7 \text{ Å} \), the FWHM of the LUMO peak is 14 meV. We note that although a mechanical gating effect is observed already in DFT-LDA [top curves in Figs. 2(b) and 2(c) and open circles in Fig. 2(d)], there is a large discrepancy with the experimental data: the DFT-LDA LUMO reaches the Fermi level far too late (at 12 Å), and it does not sharpen as much as in experiment. Because of this sharpening, the experimental resonance at the Fermi level has been discussed in terms of dynamic correlations (Kondo effect). \( (2, 9) \)

A systematic approach to the Kondo effect is achieved by taking the single-particle mean-field spectrum of the LUMO (at each \( z \)) as a starting point for a NRG simulation (see Sec. II A for details). The DFT-LDA spectrum should not be interpreted as a realistic mean-field spectrum since it lacks physical significance, being essentially an auxiliary quantity within a ground-state total-energy approach. However, realistic mean-field spectra can be addressed within the standard \( GW \) approximation of MBPT. \( (7) \) In particular, MBPT takes nonlocal correlation into account. In the present case, however, a full \( GW \) calculation is too demanding. We therefore employ a simplified, perturbative \( LDA + \Delta GW \) approach, \( (20-22) \) which yields reliable QP energies \( E^{GW}_{n} = E^{LDA}_{n} + \Delta^{GW}_{n} \) on top of DFT-LDA (see Sec. II A and Ref. 22 for details).
We note which yields the many-body corrections for the LUMO from the various methods. Note that the metallic polarizability of the molecule becomes combined with the nonmetallic one, depending on \( z \). When PTCDA is adsorbed on the Ag(111) surface, the nonmetallic corrections for PTCDA) is thus weakened by the metal, depending on \( z \) for the state being empty, being partially occupied (at \( n_{\text{LDA+GdW}} \)), and being filled. (b) Intrastate interaction \( U \) [from Eq. (4)]. (c) Occupancy of the LUMO from the various methods. Note that \( n \) refers to one spin channel and the total number of electrons in the state is \( 2n \). (d) Resulting Kondo temperature.

QP corrections from LDA + GdW differ from standard GW data by less than 20%.\(^{20-23}\) For gas-phase PTCDA, LDA + GdW can be checked against a standard GW calculation and against experiment. Figure 2(a) shows that the GdW shifts (−1.8 eV for the HOMO and +1.7 eV for the LUMO) agree well with standard GW data (both with our own and with those of Ref. 27), in particular for the LUMO, and with experiment,\(^{27,28}\) establishing LDA + GdW as a suitable method.

When PTCDA is adsorbed on the Ag(111) surface, the nonmetallic polarizability of the molecule becomes combined with the metallic response of the surface in a nonlocal way (due to Coulomb-interaction convolution or local-field effects, as expressed in the inversion of the dielectric matrix leading to \( W \)). The nonmetallic \( W \) of the molecule (causing the nonzero QP corrections for PTCDA) is thus weakened by the metal,\(^{22,29,30}\) depending on \( z \). This leads to three features in \( GdW \) and \( E_{\text{GdW}} \): (i) due to the metallic nature of the substrate, which screens the interaction inside the molecule nonlocally,\(^{22,29,30}\) the \( GdW \) corrections are substantially reduced as compared to gas-phase PTCDA. As shown in Fig. 3(a), the QP corrections are weaker than ±0.5 eV for \( z = 5 \) Å instead of almost ±2 eV for the free molecule. Note, however, that as the molecule is removed from the metallic substrate, the influence of the substrate on the screening inside the molecule decreases, and both \( W \) and \( A_{\text{GdW}} \) grow again with increasing \( z \). The increase (decrease) of the self energy of an occupied (empty) molecular level upon approach to a metallic surface, as shown in Fig. 3(a), can also be interpreted as an image-potential effect. Note that image-potential effects are an intrinsic feature of the \( GW \) (and \( GdW \)) self-energy operator.\(^{31,32}\) (ii) The QP correction of the LUMO changes sign as a function of its occupation \( n \) (referring to one spin channel, i.e., \( 0 < n < 1 \)). If the state were empty (occupied) it would observe a GdW correction of \( \Delta_{n=0} \) (\( \Delta_{n=1} \)) > 0 (\( \Delta_{n=1} \) < 0), as shown in Fig. 3(a). At partial occupation \( 0 < n < 1 \) (through charge transfer from the metal) we obtain an \( n \)-dependent QP shift \( \Delta_{n=0} = (1 - n)\Delta_{n=0} + n\Delta_{n=1} \) (iii) Since the \( \Delta_{n=0} \) shifts the whole LUMO spectrum \( f(E) \) rigidly with respect to the Fermi level, thereby changing its occupation by \( \Delta n \), the classical (screened) Coulomb energy changes by 2(\( n \))U (the factor of two results from spin degeneracy), which must be added to \( \Delta_{n=0} \), yielding an effective GdW correction of \( \Delta_{\text{eff}} = (1 - n)\Delta_{n=0} + n\Delta_{n=1} + 2(n - n_{\text{LDA}})U \). Since the QP shift in turn determines \( n = -1 E_{\text{GdW}}(E) \) via \( E_{\text{GdW}}(E) = f^{-1}(E - \Delta_{\text{GdW}}(n)) \), \( n \) and \( \Delta_{\text{GdW}} \) must be determined self-consistently for each \( z \) [see Figs. 3(a) and 3(c)]. At small \( z \), we obtain a negative QP shift and a slight increase of \( n \) (as compared to \( n_{\text{LDA}} \)) while at large \( z \) the QP shift approaches zero. Although the QP correction turns out to be small, there is a clear trend from \( \Delta_{\text{GdW}} = -0.17 \) eV for \( z = 5 \) Å to \( \Delta_{\text{GdW}} \approx 0 \) eV for \( z = 12 \) Å. This leads to an increased slope of the LDA + GdW QP peak position for \( z \) between 7 and 10 Å [see Fig. 2(d)].

After inclusion of the GdW QP corrections, the peak positions still converge toward the Fermi level much slower than found in experiment [see Fig. 2(d)]: the LDA + GdW LUMO reaches the Fermi level far too late (at 12 Å), and its spectrum does not sharpen as much as in experiment [see Fig. 2(c)]. As expected, QP corrections alone (from nonlocal correlation effects), although being mandatory for the following procedure, do not provide a complete physical picture since dynamical correlations (leading to the Kondo effect) are still missing.\(^{25}\) This will be the topic of the next section.

IV. CORRELATED SPECTRUM AND KONDO EFFECT FROM NRG THEORY

Having calculated the mean-field spectrum \( f^{\text{GdW}}(E,z) \) of the molecule-metal contact from first principles, we use this spectrum as input to NRG theory to implement dynamical correlation. Figure 2 indicates that the physics of the PTCDA/Ag(111) contact near the Fermi level is dominated by a single nondegenerate orbital, so the single impurity Anderson Hamiltonian is appropriate. Employing the calculated values of \( U(z) \) (see next paragraph) we extract the bare LUMO level positions \( \epsilon_0(z) \) and coupling functions \( \Gamma(E,z) \) from the LDA + GdW spectra at each \( z \) (see Sec. II A for more details). Together with \( U(z) \), the so obtained \( \epsilon_0(z) \) and \( \Gamma(E,z) \) enter the NRG calculation,\(^{3,33}\) which yields the many-body LUMO spectra displayed in Figs. 2(b) and 2(c).\(^{34}\) We note that while previous work\(^{35}\) extracts only some characteristic parameters from the LDA, we have used here the full spectral information \( \Gamma(E,z) \) from LDA−GdW and the calculated screened Coulomb repulsion \( U \) as input to our NRG.

One key element of our calculations is the intrastate interaction energy \( U \). It is given (for each \( z \)) by

\[
U = \int |\psi_{\text{LUMO}}(r)|^2 W(r,r') |\psi_{\text{LUMO}}(r')|^2 d^3r d^3r'.
\]

Figure 3(b) reveals that as a consequence of efficient metallic screening the \( U \) parameter is rather small (below 1 eV) for \( z < 8 \) Å, but increases for larger \( z \). Similar to the QP corrections, this results from the diminishing influence of the metal on the
screening inside the molecule as \( z \) increases. For gas-phase PTCDA, \( U \) amounts to 3.0 eV.

The \( z \)-dependent renormalized peak position of the LDA + GdW + NRG data (full squares) can now be compared to the experimental data as shown in Fig. 2(d). As in experiment, the calculation shows a fast-rising QP peak (slope 0.15 eV/Å for \( z \) between 8 and 10 Å), which also becomes pinned at \( E_F \) as soon as it reaches the Fermi energy (at \( z \approx 10 \) Å). In both experiment and LDA + GdW + NRG, the pinned resonance is substantially sharper than the QP peak that moves up to \( E_F \) [see Fig. 2(c)]. Since this sharpening to 14 meV (experiment at 9.7 Å) or 15 meV (theory at 11 Å) is neither observed in LDA nor in LDA + GdW, we can ascribe it to a dynamical correlation effect.

Note that for \( z \lesssim 8 \) Å, the LDA + GdW and LDA + GdW + NRG spectra are very similar [see Fig. 2(b)]. This is to be expected, since a nearly fully occupied state below \( E_F \) shows very weak dynamical correlation. In contrast, for \( z \gtrsim 9 \) Å the difference between LDA + GdW and LDA + GdW + NRG becomes dramatic. While LDA + GdW still yields a single LUMO peak [see Fig. 2(c)], the LDA + GdW + NRG spectrum shows a three-peak structure, with two single-particle side bands (at \(-0.9 \) eV and \(+0.3 \) eV for \( z = 11 \) Å, separated by \( -U \)) and a Kondo resonance at \( E_F \) [see Fig. 2(c)] and sharp horizontal line in Fig. 2(d).

In Fig. 3(d), the evolution of the Kondo temperature \( T_K(z) \) is plotted. \( T_K \) has been determined from the NRG calculation at the temperature at which 40% of the local moment in the LUMO is screened (Wilson criterion). For \( z < 9 \) Å, \( T_K \) is a few thousand K because in this \( z \) range \( U \) and \( \Gamma \) are of similar size, and hence, there is no clear separation between spin and charge fluctuation energy scales. In other words, in this regime, \( T_K \) is equivalent to the charge fluctuation scale \( \Gamma \). Physically, this means that screening of the small residual moment that is left by the total occupancy of 1.8 electrons in the LUMO is accomplished by uncorrelated charge fluctuations in and out of the level. At \( z = 12 \) Å, \( T_K \) has dropped to \( T_K < 1 \) K, caused by the increase of the ratio \( U/\Gamma \), which leads to the clear separation of spin and charge energy scales. Accordingly, \( T_K \) is now a Kondo temperature in the proper sense, related to the energy scale of correlated screening of a fixed moment in the LUMO by virtual spin fluctuations. In experiment, the Kondo resonance can only be followed up to \( z \approx 10 \) Å because of decreasing signal strength, but at 9.7 Å the Kondo resonance has a FWHM of 14 meV, in good agreement with a predicted Kondo temperature of the order of 100 K [see Fig. 3(d)].

One remaining difference between our LDA + GdW + NRG data and experiment is the \( z \) position at which the QP peak reaches \( E_F \) and turns into the resonance. This could be due to a slightly too low level position of all LDA + GdW data compared to full-GW data and experiment [see Fig. 2(a)], resulting from the approximations within the method. To account for this, all QP corrections might have to be increased by \(+0.2 \) eV. When this is included in our LDA + GdW + NRG data [see open squares in Fig. 2(d)], the LUMO peak is in fact shifted to higher energies (matching the experimental data point for \( z = 5 \) Å) and reaches the Fermi level at about \( z = 9 \) Å, in much better agreement with the measured data.

V. SUMMARY

In conclusion, we have presented an approach to calculating electronic spectra that systematically includes nonlocal and dynamic correlation effects. It correctly predicts the mechanical gating of the tip/PTCDA/Ag(111) junction. Because of its computational efficiency, this state-of-the-art approach to electronic spectra should be widely applicable.

ACKNOWLEDGMENTS

We thank the Deutsche Forschungsgemeinschaft for financial support (Grants RO 1318/6-1, AN 275/6-2, and TA 244/5-2). Computational resources have been provided by the John von Neumann Institut für Computing (NIC) at the Forschungszentrum Jülich.

---

1. Corresponding author: michael.rohlfing@uos.de

34 Since the LDA calculation indicates a weak influence of the tip onto the spectra, we also neglect the voltage dependence of the spectral function which would be only significant for a symmetric junction.2,36