Bandgaps of Atomically Precise Graphene Nanoribbons and Occam's Razor

Aristides D. Zdetsis*

Molecular Engineering Laboratory, Department of Physics, University of Patras, Patras 26500 GR, Greece, and

Institute of Electronic Structure and Laser, Foundation for Research &Technology Hellas, Vassilika Vouton, P.O. Box 1385, Heraklion, Crete GR-71110, Greece

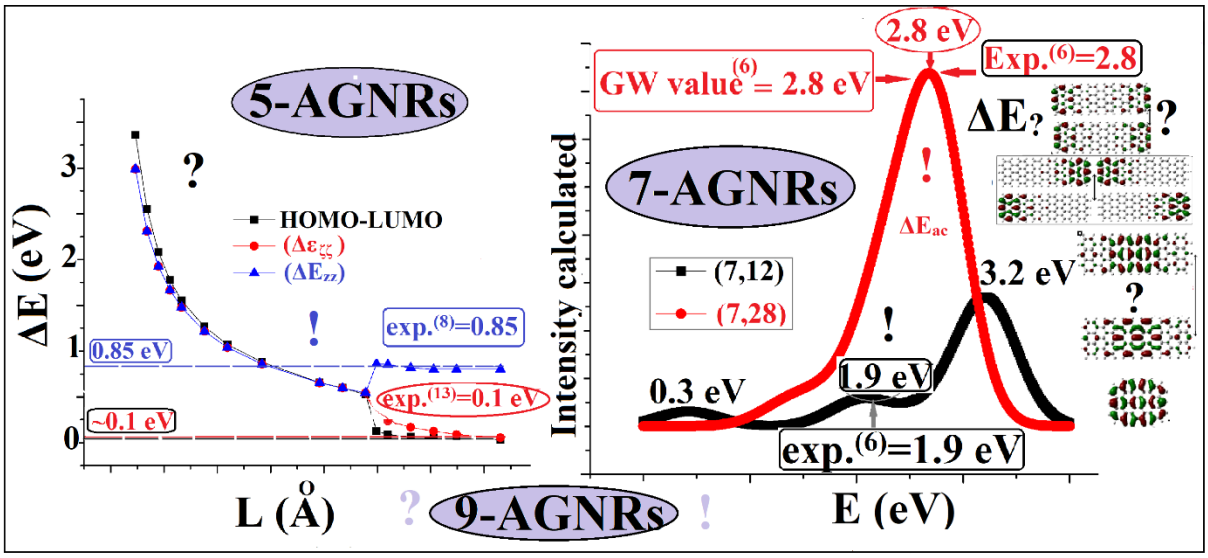
ABSTRACT

Rationalization of the "bulk" (ΔE_{ac}) or "zigzag-end" (ΔE_{zz}) energy gaps of atomically precise AGNRs, which are directly related to fundamental applications in nanoelectronics, could be challenging and largely controversial with respect to their magnitude, origin, substrate influence (ΔE_{sb}), and spin-polarization, among others. Hereby a simple self-consistent, "economical" and highly successful interpretation is presented based on "appropriate" DFT (TDDFT) calculations, general symmetry principles, and plausibility arguments, which is fully consistent with current experimental measurements for 5-, 7-, and 9-AGNRs within less than 1%, although at variance with some prevailing views or interpretations for ΔE_{ac}, ΔE_{zz}, and ΔE_{sb}. The excellent agreement with experiment and the new insight gained is achieved by invoking the approximate equivalence of Coulomb correlation energy with the staggered sublattice potential. Breaking established stereotypes, we suggest that the measured STS gaps are virtually independent of the substrate, essentially equal to their free-standing values, and that the "true" lowest energy state is closed singlet with no conventional magnetism. The primary source of discrepancies is the finite length of AGNRs together with inversion/reflexion symmetry conflict and the resulting topological end/edge states. Such states invariably mix with other "bulk" states making their unambiguous detection/distinction difficult. This can be further tested by eliminating end-states (and ΔE_{zz}), by eliminating "empty" zigzag rings.

*zdetsis@upatras.gr

Keywords: Armchair Graphene Nanoribbons (AGNRs), Atomically Precise AGNRs, "bulk"/"surface" Energy Gaps, Topological end-states, Aromaticity, Inversion symmetry, sublattice structure, scanning tunnelling spectroscopy (STS)

Graphical Abstract



Conflicting results and open questions (?) on various "bulk" (ΔE_{ac}) and "surface" (ΔE_{zz}, $\Delta\varepsilon_{\zeta\zeta}$) energy gaps of AGNRs have very simple answers in accord with Occam's razor. Based on "inexpensive" DFT calculations and inversion-sublattice symmetry arguments we achieve very high accuracy (!)

1. Introduction Edge or end states in graphene nanoribbons (GNRs), and in particular armchair GNRs (AGNRs) have attracted very much interest lately,[1-8] due to their anticipated magnetic properties,[9-10] although their presence in finite nanographenes (NGRs) has been predicted long time ago.[11] However, the significance and importance of end states for AGNRs was recognized only recently[1-8], after the pioneering bottom-up synthesis of atomically precise AGNRs of finite lengths L with short zigzag ends, and their characterization by scanning tunnelling microscopy (STM) and spectroscopy (STS). [3-16] Clearly no end states appear in the common infinite AGNRs fabricated by the usual top-bottom techniques, which are theoretically described by periodic boundary conditions at their two ends[1]. The new developments have brought to the forefront new concepts and properties such as the "bulk band gaps" ΔE_{ac} (or Δ_{ac}[1,6]) i.e., the energy gaps between delocalized states, and the energy separation of the zigzag-end-localized "end-states", denoted here by (ΔE_{zz}) (or Δ_{zz} [6-8]), thus increasing both quantity and quality of key properties to be rationalized, understood, or interrelated at the atomic scale. At the same time, despite the increased complexity, such advances have also allowed the study of the L-dependence of key-quantities such as the bandgaps[4-7] (both, ΔE_{ac} and ΔE_{zz}), conductivity, aromaticity[1, 3-4], and even Raman spectra.[17] The L-dependence studies[4-5] revealed that the changes in such properties versus length are not always gradual (or smooth). The presence of a metal-insulator-like phase transition at a critical length L_c was advocated by two different recent works, Lawrence *et al.*[8] and Zdetsis *et al.*[4], almost simultaneously. However, the two works have offered different assessments and interpretations for the nature of the transition and magnetism, as well as the value of L_c.[4, 8] This is not something new or unusual in a rapidly grown pioneering field like this[1], and this is not the only existing "discrepancy". Other conflicting (or conflicting-looking) results (experimental and theoretical) include the magnitude and nature of the bandgaps[1, 6-7, 12-13], the existence and nature of magnetism in the edge states[1, 3-7], as well as the magnitude of the substrate influence on these properties.[1, 5-7] For example, the magnitude of the bandgap for the 5-AGNRs has been measured by (at least) three

different groups[8, 12-13] to be 0.85eV[8], 2.8 eV[12], and 0.1 eV[13] respectively, while the theoretical values vary from 0.1 eV[1] to 1.7 eV.[14] For the 7-AGNRs the measured values of ΔE_{zz} vary between 1.9 eV[6] and 2.5 eV[7], whereas the measured ΔE_{ac} values range from 2.3 eV to 3.2 eV[6-7,9,15], overlapping significantly with the range of ΔE_{zz}. Thus, the unambiguous distinction between ΔE_{ac} and ΔE_{zz} is another subtle point together with the bridging of the measured and calculated ΔE_{ac} values, which also vary widely from 2.3 eV to 3.7 eV.[1, 6, 7, 14] Some of the (different) measured or calculated values correspond to AGNRs of different length, but in the literature the quoted values are usually given without reference to the actual length which is, thus, treated as a hidden variable. However, the biggest problem seems to be the large difference between the measured values of the gap(s) in relation to the "official" theoretical values obtained by the GW method[14], which are widely recognized as an almost universal point of reference. Such large differences (almost ~1.5 eV for the 7-AGNRs) between experimental and theoretical GW gaps (ΔE_{ac}) are usually attributed to the screening from the metallic (Au) substrate ΔE_{sb}, even though identical values of gap (within the experimental uncertainties) have been obtained for AGNRs grown on non-metallic substrates, such as NaCl[6] and MgO[7]. This is clearly (at least) problematic. As a result, it appears that there are several conflicting results and interpretations or high-braw "solutions" about the STS gaps, although the real solution could be much simpler (but not always obvious), as could be possibly argued on the basis of Occam's principle. Along these lines the present work aims at deciphering all these subtle points, also including the confusion in distinguishing between ΔE_{ac} and ΔE_{zz} gaps. Thus, the present work can be considered as a positive synthesis of various conflicting views. Based on previous experience,[1, 3, 19] it is expected that such synthesis should be proven successful and constructive, facilitating the successful and accurate functionalization of AGNRs for realistic applications. As is demonstrated below, we can fully rationalize all known experimental data for the 5-, 7-, and 9-AGNRs within less than 1% accuracy, predict non-measured gaps, and pinpoint at the same time the sources of discrepancies.

2. Theoretical framework. For a consistent and transparent understanding and interpretation of the origin and magnitude of ΔE_{ac}, ΔE_{zz} as well as the factors that influence their size, it is important to realize that practically all these quantities are dominated by the influence of the ("many-body") Coulomb correlation energy combined with sublattice frustration, which gives rise to the staggered sublattice potential[20] across the zigzag ends of finite length AGNRs. In fact, the sublattice frustration, which is the driving force for the generation of the end/edge states, as we have illustrated earlier,[2-4] constitutes the largest (or even the full) contribution on the Coulomb correlation energy. The understanding that most (or all) of the Coulomb correlation energy is devoted to counterbalance the topological frustration between sublattice and molecular symmetry-groups is the starting (and key) point of the present investigation. This principle together with the established[2-5] (hidden) strong contributions of aromaticity and shell structure[2-5] constitute the basis for the deeper understanding of all these quantities (ΔE_{ac}, ΔE_{zz}, $\Delta\varepsilon_{\zeta\zeta}$, and ΔE_s). Thus, if we can properly alleviate the sublattice-molecular group symmetry frustration (which is equivalent with inversion/reflection symmetry conflict), under the natural constrains of shell structure and aromaticity, we could effectively account for the (largest part of) Coulomb correlation energy. At the same time this would explain why the open shell states (singlet or triplet) are not the real lower energy states, but rather "pseudo-states".[21]

2.1 Calculation of ΔE_{zz} and ΔE_{ac} from the one-body DFT calculations. Within the 1-electron approximation underlining the DFT and Hartree-Fock (HF) self-consistent fields, the symmetry frustration between molecular (D_{2h}) and sublattice (C_{2V}) symmetry groups can be alleviated by effectively breaking (or redefining) the symmetry of the additional degrees of freedom (besides spatial coordinates) i.e., the spin and/or pseudospin (for real-space calculations). In the first case we can introduce non-zero spin values preserving the molecular symmetry,[3] whereas in the second case we are forced to break molecular symmetry, by introducing open-shell singlet states, which when optimized geometrically converge normally to C_{2v} symmetric geometries compatible with sublattice symmetry, thus breaking the molecular symmetry as well. This occurs because the HOMO (and

LUMO) orbitals of the open singlet are obtained, by construction, by mixing the HOMO and LUMO orbitals of the closed singlet. These orbitals have opposite parities, u, or g, (and opposite behaviour under σ_y or σ_{xz} reflection plane). Thus, by losing the σ_y (or σ_{xz}) reflection plane of the D_{2h} symmetry group, as is illustrated in Fig. 1(a), we can get sublattice distribution with opposite sublattice points at the two ends (or antisymmetric with respect to y-axis, but symmetric with respect to the x-axis, which is the axis of the AGNR). This facilitates frontier molecular orbitals (HOMO, LUMO) localized only at one end (left or right) of the AGNR, producing an antisymmetric (pseudo)spin density, as is shown in Fig. 1(b), reflecting the sublattice symmetry and structure. Obviously, the reverse picture with the pseudospins interchanged is equally valid. On the other hand, the molecular D2h symmetry demands same type (same sublattice) atoms at the two ends, as shown in the lower part of Fig. 1(a), thus producing a symmetric, with respect to the y-axis, (pseudo)spin distribution (bottom of Fig. 1(b)). In both cases the (pseudo)spin distribution is almost zero at the middle part of the AGNR. This is reproduced in the corresponding "spin" densities (b). Such spin densities, in Figs. 1(b), 1(c) and 1(d), which are generated self consistently (through the DFT convergence process) are in fact pseudospin densities. As mentioned above, the open singlet reflects the sublattice symmetry (with different type atoms at the two ends) with equal number of up and down (pseudo)spins, or A and B sublattice points, resulting in a balanced (nearly zero spin distribution in the middle). The triplet state on the other hand has also a region of zero spin in the middle, exactly where the sublattice imbalance occurs. In the ordinary "atomistic" calculations the sublattice degree of freedom does not enter in the spatial Hamiltonian and can only be introduced as (pseudo) spin. Then, due to the better account of Coulomb interaction, open shell states (triplet or singlet) appear energetically lower than the closed singlet state. This is because the additional degree of freedom of "pseudospin", introduced to take care of the sublattice topology (and the staggered potential), facilitates the optimization of Coulomb interaction by keeping away of each other electrons of different spin (for which Pauli repulsion is not operative), but of identical pseudospins. Moreover, based on the shell model[2-3], the

unoccupied states of the "previous" (shell number smaller by 1) AGNR are the occupied of the current AGNR. This is responsible for the interplay between odd and even parity HOMOs as the width of AGNRs is growing (3n AGNRs have odd HOMO and even LUMO, whereas 3n+1 AGNRs are characterized by even HOMO and odd LUMO).[2-3]

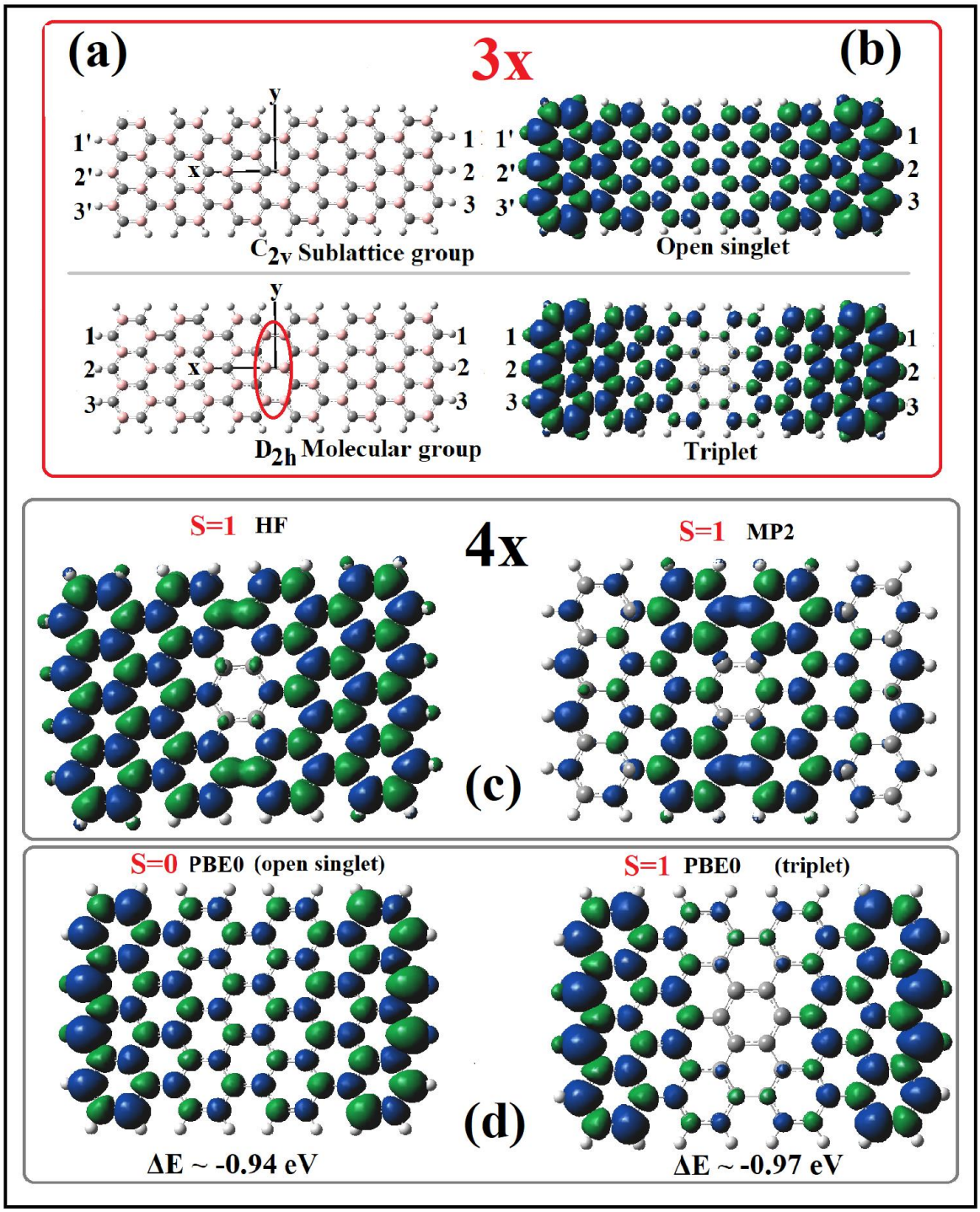


FIGURE 1. Molecular and sublattice symmetry of the 3x6 (7,12) AGNR (a), reflected in the spin densities (b). Isovalue=0.004. The vertical elliptic curve indicates the region of sublattice imbalance (and frustration). Comparison of the corresponding spin densities for the 4x4 (9, 8) AGNR at the HF, MP2 and DFT/PBE0 level is given in (c) and (d).

This is also responsible for the well-known 3n, 3n±1 width rule for AGNRs.[3] Note that the (pseudo)spin densities invariably reflect the sublattice (pseudospin) structure within the frustrated molecular (D_{2h}) symmetry[3] in the first case, or the sublattice symmetry (C_{2v}) in the latter (see also Fig. S1), where opposite end sites have opposite spins. It should be emphasized at this point that for wider AGNRs (where n>1 in the above rule for width[3]), higher spin states are required[3] to lower the total energy (within the molecular D_{2h} symmetry group). Such larger (pseudo) spin-polarized states optimize better the sublattice distribution (within the D_{2h} molecular group),[3] whereas the open-shell singlets lye higher in energy and revert to the closed singlet state. This illustrates emphatically that the open-singlet state is not the true ground (lowest energy) state of AGNRs (and, consequently, no conventional magnetism is truly present). Nevertheless, the open singlet state is still a very useful and efficient concept for the description of end-states, as is illustrated below. It should be emphasized that in both cases of Fig. 1, when correlation is introduced even at the MP2 level, the energetical ordering is reversed and the lowest energy structure is a closed singlet.[4] In addition, the MP2 correlated "spin" density of the triplet, as we can see in Fig. 1 (c), is rather correcting the HF failure (having the opposite sign) than reflecting the full sublattice structure. Note also in Fig. 1 (c, d) that the triplet state is slightly lower than the open singlet, and that the energy difference of the open shell singlet and triplet states (which are practically isoenergetic) from the closed singlet is about 0.95 eV. This should be a good estimate of the "missing" Coulomb energy in this case, and, based on the (approximate) electron-hole symmetry, the expected (HOMO-LUMO) separation of the open singlet (or the triplet) should be about twice as large (~2 eV). Indeed, the calculated open-singlet HOMO-LUMO gap for the 9-AGNRS (or 4x) is 2.2 eV, and so is ΔE_{zz} (*vide infra*). Even more important is

the fact that the corresponding value for the 3x6 or (7,12) AGNR is also about 0.93 eV, suggesting an open-singlet gap of about 1.9 which is in excellent agreement with both the measured value[6] of ΔE_{zz} (1.90 eV), and the calculated open singlet gap. It is important to observe also that the open-singlet value "ΔE_{zz}" = 1.2 eV for the 5-AGNRs and the 1.9 eV open singlet gap for the 7-AGNRs are practically equal to the correlation improved GW-LDA bandgap differences,[14] which is highly suggestive for the essential correctness of our claim. Thus, within the one-electron approximation we have established the correct basis for discussion and analysis of both ΔE_{zz} and ΔE_{ac}. ΔE_{zz} is identified as the open-singlet HOMO-LUMO gap, whereas ΔE_{ac} can be identified as the difference |(HOMO-1)-(LUMO+1)|, with the understanding that both HOMO and LUMO are end-states. Nevertheless, as will be illustrated below, similar values of gaps and analogous estimates can be found in the triplet state as well. Furthermore, it should be emphasized that the central meaning of ΔE_{zz} is only valid for lengths L longer than the critical length ($L \geq L_c$), although the open-singlet HOMO-LUMO is defined for almost all lengths and is practically constant, as is verified also by Wang et al.[6]

2.2 Introducing many-body corrections to the gaps. For both gaps (ΔE_{ac} and ΔE_{zz}) we can further correct if we wish their (one-body) values by considering additional many body contributions through time-dependent DFT (TDDFT), which has been shown[1] to provide very good ("many-body") estimates of the gaps, so that the STS spectrum overall looks very much alike the (luminous) optical spectrum, because both are dominated by molecular overlaps between transition states. This is further illustrated and "verified" from the results below. Furthermore, the use of TDDFT allows the clear and unambiguous identification of the energy separation of the end/edge states, which according to the present investigation is not given by ΔE_{zz}, as Wang *et al.*[6] have suggested, but by another type of gap which here is denoted as $\Delta\varepsilon_{\zeta\zeta}$. In the usual one-body approximation $\Delta\varepsilon_{\zeta\zeta}$ corresponds to the HOMO-LUMO separation of the closed singlet true ground state for $L \geq L_c$, which is always only a few 0.1 eV (~0.1 eV, for $L \rightarrow \infty$) in accord with the association of the end states with the Dirac points[3-4] (and charge neutrality points[4]) located "very close" to the

fermi level. TDDFT indeed verifies that in contrast to $\Delta\varepsilon_{\zeta\zeta}$ which involves transition from one purely end-localized HOMO state to an opposite-parity end-localized LUMO state, the ΔE_{zz} gap always involves transitions from a mixture (~ 60% - ~ 40%) of "surface"-"bulk" states to another state of about equal amount of mixing. Thus, although ΔE_{zz} involves a large amount of localized end-states, it should not be associated with the energy separation of the end-states. Another way, besides TDDFT, to distinguish between "bulk" and "surface" energy gaps is by comparing to the corresponding "edge-modified" AGNRs,[5] obtained by eliminating "empty" (i.e., non-aromatic) end-rings, which also eliminates topological end-states (and, therefore, ΔE_{zz} and $\Delta\varepsilon_{\zeta\zeta}$).

2.3 The Substrate influence on the measured STS gaps. As we have mentioned above, the main crucial property under possible dispute is the magnitude of the substrate influence (screening) ΔE_{sb} on the measured STS gap. According to our earlier estimates[1] ΔE_{sb} should be of the order of a few 0.1 eV. However, almost in all cases ΔE_{sb} larger than 1 eV is needed to bridge the experimental STS measurements for AGNRs deposited on metal surfaces (usually Au) and the theoretical values for free standing AGNRs. The theoretical values widely recognized as an almost universal point of reference are the GW results of Yang *et al*.,[14] which among the theoretical values reported earlier are clearly the largest, and many times by far. As a result, ΔE_{sb} which is defined as the difference of the STS measurements and the theoretical reference values are unrealistically large. For example, for the 7-AGNRs the theoretical GW gap[14] is 3.7 eV, whereas the experimental STS gap value obtained by various groups[6-7, 9,15] is 2.5±0.2 eV. Thus, ΔE_{sb} should be at least 1.2 eV. However, the STS value of 2.5 eV was also obtained for 7-AGNRs deposited on non-metallic substrates, such as NaCl [6] and MgO[7], for which such large ΔE_{sb} value is clearly unrealistic. On the basis of their STS measurements on samples grown on MgO, Kolmer *et al*.[7] concluded that ΔE_{sb} should be marginal, which is in full agreement with our present results. However, the general consensus, with few exceptions[1,7,18] is largely different (up to now). In this work we are led to conclude that the GW results[14] overestimate the bandgaps mainly due to the size effect, since the

GW results of Yang *et al.*[14] were obtained for infinite AGNRs, whereas the atomically precise AGNRs have finite length (and topological end-states). This could be sufficient to explain the resulting unrealistically large ΔE_{sb} values. Yet, besides the infinite size (and the corresponding periodic boundary conditions) the lack of exact exchange in the LGA wavefunctions building the Green's function could be also important since exchange interaction is very sensitive to inversion symmetry frustration. Nevertheless, judging from our TDDFT results, it is more reasonable to attribute the gap difference between the infinite and the finite size AGNRs (size effect) to the mixing of edge/end states with the infinite "bulk" states (and the scattering at the zigzag edges) which can drastically reduce the gap. This is corroborated by the GW results[6] of Wang *et al.*[6] for the finite (7, 24) AGNR (and slightly longer), who obtained a gap of 2.8 eV clearly closer to the measured (by several groups) gap, and substantially smaller compared to the 3.7 eV (G_0W_0) value[14] for the infinite 7-AGNR. Parenthetically, it should be mentioned at this point that even in the worst-case scenario where the substrate interaction is strong (especially when the distance of STS tip from the surface is small), this leads to mixed substrate-AGNR states[1] which can be easily recognized (and excluded) from the measurements by comparing the two separate STS spectra. Moreover, such states would be expected to have low overlap with the pure AGNR- excited-states, and consequently the corresponding transition(s) would have very low intensity and would be difficult to detect. Thus, we assert here that ΔE_{sb} should indeed be marginal, in full agreement with the experimental results (for 7-AGNRs grown directly on MgO substrate) and conclusions of Kolmer *et al.*[7].

2.4 Computational details. The theoretical and computational details of the present investigation have been described in references 1 through 5. The computations, as before, have been performed with the Gaussian[22] program package using the DFT and TDDFT employing the PBE0 functional[23] (which includes "exact" exchange, in contrast to PBE) and the 6-G31(d) basis set. The same computational package was used for the Møller-Plesset many-body perturbation theory of 2nd order

(MP2) calculations, with the same basis set. The visualization of the results was accomplished using the GaussView software.[24]

2.5 Synopsis of the theoretical approach. For the atomically precise AGNRs examined here it has been illustrated that although the closed singlet is the correct lowest energy state, due to (inversion) symmetry frustration (which is also a sublattice frustration), open shell states, such as open shell singlet and triplet appear energetically lower due to better account of the Coulomb correlation energy (within the 1-electron approximation).[3] Thus, such open shell states should be considered as "pseudo-states", and their resulting spin distribution within the one-electron DFT framework should be characterized as "pseudospin" distribution. The full amount (or most of it) of the Coulomb correlation energy can be estimated from the total energy difference of the closed singlet and the open singlet or triplet states. Thus, using standard DFT calculations (with functionals including exact exchange, such as the PBE0[23]) for the closed and open singlet, and/or triplet states we can have a very good estimate of all ("surface" and "bulk") energy gaps, measured by STS as follows:

1st) The HOMO-LUMO gap of the closed singlet corresponds to the real energy separation of the end-states localized at the zigzag ends, $\Delta\varepsilon_{\zeta\zeta}$. This is true for long enough AGNRs for which both HOMO and LUMO are zigzag-end-localized.

2nd) In this case the (HOMO-1)-(LUMO+1) difference corresponds to the "bulk" bandgap ΔE_{ac} between states delocalized over the entire AGNR.

3rd) ΔE_{ac} can be further verified, if desired, by comparing to the HOMO-LUMO gap of the edge modified AGNRs[4-5], which contain no zigzag end bonds, and no end-states.

4th) On the other hand, the zigzag-end-localized HOMO-LUMO gaps of the open-shell singlet or the triplet states provide a very good estimate of the "mixed" ΔE_{zz} gap.

5th) Further refinement of all these three gaps ("bulk", "surface") can be obtained, if needed, by the analysis of the orbital composition of the main peaks of the excitation spectrum obtained by TDDFT. This can also help the correct identification (and nature) of the gaps.

3. Results and discussion

3.1 5-AGNRs. Figure 2 summarizes the present results for the 5-AGNRs (or 2x AGNRs) which, as mentioned earlier, have been also studied by several groups.[1,4, 8, 12-13, 16]

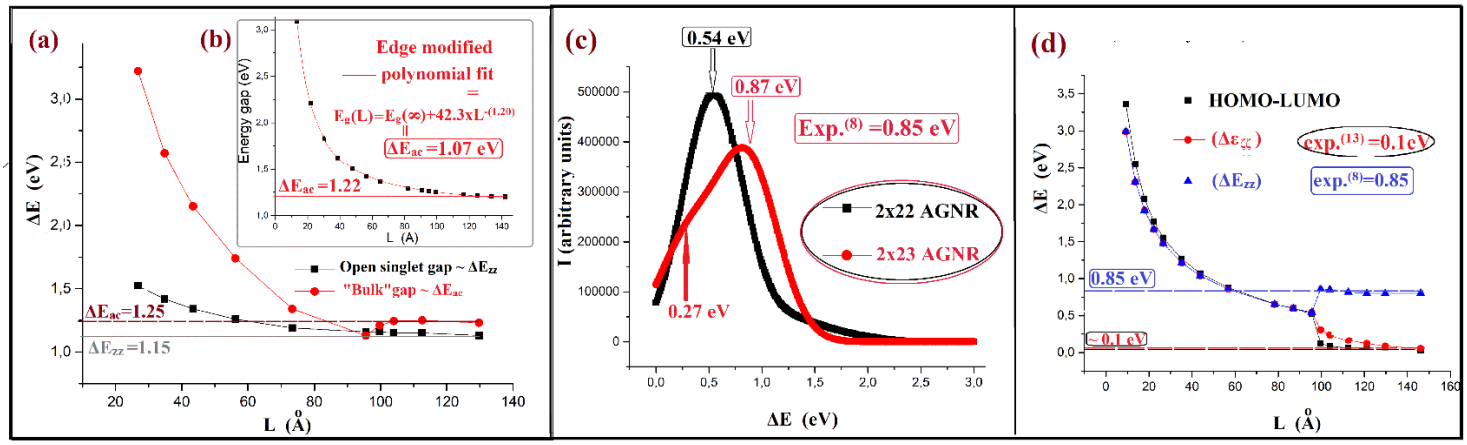


FIGURE 2. (a): Variation of the open singlet, ΔE_{zz}, and the "bulk" |(HOMO-1)-(LUMO+1)|= ΔE_{ac} gaps (in eV) in terms of length L (in Å) for the 5-AGNRs (2x). **(b):** Variation of ΔE_{ac} gap (in eV) as a function of length for the edge modified 5-AGNRs together with the usual polynomial fit (see text). **(c):** excitation spectrum of the 2x22 and 2x23 AGNRs. Intensity (I) is in arbitrary units and excitation energy (ΔE) in eV. **(d):** Variation with length of the HOMO-LUMO, and ΔE_{zz} , $\Delta\varepsilon_{\zeta\zeta}$ gaps, calculated by TDDFT as "first" and "second" optical gaps respectively, including the corresponding experimental values from refs. 8, and 13 (see text).

In Fig.2(a) the open singlet HOMO-LUMO gap, ΔE_{zz}, and the "bulk" gap ΔE_{ac} are plotted versus length L. Here, following the discussion above for the open singlet gap and its relation to ΔE_{zz}, we have defined ΔE_{zz} as the HOMO-LUMO gap of the open singlet state, contrary to the original definition of Wang *et al.*[6] as the energy separation of the end states. Obviously, for an open singlet

ground state both definitions are equivalent, but this is not the case. As we can see in Fig. 2(a), the "one-body" ΔE_{ac}=|(HOMO-1)-(LUMO+1)| gap after the discontinuity (or transition) at L≈100Å, which we have discussed in detail in a previous work,[4] starts opening up at L_c, contrary to the "one-body" ΔE_{zz} (i.e., the open singlet HOMO-LUMO gap) which varies slowly and smoothly over the entire range of lengths. This is very strange indeed, if ΔE_{zz} is going to represent the real separation of the edge states, since ΔE_{zz} first appears at and after the transition at L_c. Such behaviour (smooth variation) should be better suited for ΔE_{ac}. This is indeed verified in Fig. 2(b), which shows the HOMO-LUMO gap of the "edge modified AGNRs", which seems to saturate to the value of 1.22 eV, very close to the value of 1.25 eV, suggested from the behavior of the "normal" AGNRs in Fig. 2(a). The edge modified AGNRs by construction have no edge states and their HOMOs and LUMOs are delocalized over their entire length,[4] and therefore their fundamental gap corresponds to ΔE_{ac}. Such edge-modified AGNRs are obtained by eliminating the empty (non-aromatic) end-rings[5] of the standard AGNRs, which also eliminates end-states and zigzag end-bonds.[5] This is a clear manifestation of the importance of aromaticity for AGNRs (and graphene itself).[2-3] Comparing the behavior of the "bulk gap" in Figs. 2(a) and 2(b), we can see that due to quantum confinement (both lateral and longitudinal) the (HOMO-1) and (LUMO+1) states defining the "one-body" ΔE_{ac} are also affected by the abrupt appearance of the edge states, in sharp contrast to the ("one-body") open singlet gap which seems to be practically insensitive to the appearance of the end-states, contrary to what is expected from its original definition. This in fact emphasizes the "many-body" nature of the end states through their connection with inversion symmetry conflict, which is further supported from Figs. 2(c) and S1. The "correct" behavior (with length variation) of the "one-body" ΔE_{ac} is given by the (delocalized) HOMO-LUMO gap of the edge-modified AGNRs in Fig. 2(b). As we can see in Fig. 1(b) the value of ΔE_{ac} (HOMO-LUMO gap of the edge-modified AGNRs) as a function of length, as $L \rightarrow \infty$, seems to saturate to the value of 1.22 eV. This could be misleading since only lengths up to about 140 Å have been considered. To remedy this

problem we have recently suggested[25] to fit the calculated ΔE_{ac} as a function of L efficiently and transparently[1, 25] to a polynomial of the form $\Delta E_{ac}(L)=A+B\times L^{-C}$, where the value A corresponds to the gap at infinity, $\Delta E_{ac}(\infty)=A$, and the constant C to some short of effective ("fractal") dimensionality (here equal to 1.20).[1, 25] As we can see in the inset in Fig. 2(b), the projected ΔE_{ac} value is 1.07 eV, which is also verified by the TDDFT result ΔE_{ac} =1.01 eV (see Fig. S1). The TDDFT value (1.01 eV) is clearly closer to the value of 0.85 eV measured by Lawrence *et al.*[8], assuming a very reasonable substrate screening (of about 0.15 eV), as we have suggested recently.[4] However, further correct information is given in Fig. 2(c), showing the spectra of the 2x22 and 2x23 AGNRs immediately before and after transition, respectively (see Fig.2(d) too). As is illustrated in Fig. 2(c), in the 2x23 AGNR (immediately after the transition) there is a strong peak value at 0.87 eV, very close to the recently measured[8] STS gap of 0.85 eV. Detailed analysis of the TDDFT results shows that this peak includes transitions involving end-states to a large percentage (about 60%). Thus, the calculated value of 0.87 eV and the measured[8] gap should be assigned to ΔE_{zz}. This, contrary to the "one-body" gap, restores the expected correct behavior of ΔE_{zz} at (and after) L_c. Even more interesting is the fact that extrapolating to longer AGNRs gives a gap of 0.85 eV (exactly), which is an unexpected full agreement with experiment, as is shown in Fig. 2(d). Fig. 2(d) also shows that, contrary to the "one-body" (open-singlet) ΔE_{zz} gap of Fig. 2(a), both "many-body" gaps, ΔE_{zz}, and $\Delta\varepsilon_{\zeta\zeta}$ (the latter corresponding to the "real energetical separation of the end-states), and the one-body HOMO-LUMO gap, which involve end-states, change discontinuously at the critical length (~100 Å), where $\Delta\varepsilon_{\zeta\zeta}$ and HOMO-LUMO gaps drop, while ΔE_{zz} increases. Thus, the observed[8] gap opening (of about 0.30 eV) is due to the increased aromaticity at the critical length, and the mixing of bulk and end-states at an almost equal amount. Lawrence *et al.*[8] have attributed such gap opening to the different electrostatic potential felt by valence electrons at different regions of the ribbon due to the positive partial charge on the hydrogen atoms along the sides of the AGNR. However, the paradigm of edge-modified AGNRs contradict such

interpretation.[4] Our present work reveals that the gap opening is a many-body effect related with the aromatic transition and the change from bulk-like (ΔE_{ac}) to coupled "surface-bulk" end-states (ΔE_{zz}). On the other hand, the calculated $\Delta\varepsilon_{\zeta\zeta}$ gap of 0.1 eV in Fig. 2(d) is in full agreement with the results of Kimouche *et al.*[13] Thus, Kimouche *et al.*[13], and Lawrence *et al.*[8], have apparently ("correctly") measured different kinds of gaps. Moreover, the same could be true for the value of 2.8 eV measured by Zhang *et al.*[12], which could be assigned as a tentative ΔE_{ac} value, either for very short AGNRs (without end-states), or for longer AGNRs with a strong "bulk" transition from deep occupied states (well below HOMO-1 orbital) to higher unoccupied states (well above the LUMO+1), and thus much larger than the real ΔE_{ac} (which is technically determined by the HOMO-1, LUMO+1 difference). We can also observe in the 5-AGNRs that differences between the "one-body" and "many-body" (TDDFT) methods for assigning ΔE_{ac}, ΔE_{zz}, and $\Delta\varepsilon_{\zeta\zeta}$ are relatively large (or even unusual) compared to the 7- and 9-AGNRS, discussed below, where the corresponding differences are of the order of 0.1-0.2 eV. This could be related to the fact that the 5-AGNRs (contrary to 7- and 9-AGNRs) are topological and aromatic mixtures.[3] Thus, the three seemingly conflicting measurements[8, 12-13] for the 5-AGNRs could be attributed to different length samples (and/or different positions of the STS tip). Yet, alternatively, one could claim, based on the GW results[14], that there is a substrate interaction of equal magnitude (0.85 eV) and the "real gap" is 1.7 eV. Such conclusion is clearly considered here as highly improbable, in view of equally good (in fact better) agreement for the 7- and 9-AGNRs, not to mention Occam's principle. Moreover, if this is indeed a general trend, it clearly illustrates that elaborate correlation calculations (e.g., GW) could be avoided (see also ref. 23) if topological frustration can be taken into account appropriately by simple DFT (one particle) calculations, provided that the DFT functionals include "exact" exchange which is sensitive to inversion symmetry conflict.[4]

3.2 7-AGNRs. Figure 3 summarizes the results for the three spin states (closed singlet, open singlet, and triplet) for the 7- AGNRs, and in particular the (7, 12) or 3x6 AGNR. First of all, we

can comment on the significance of the exact exchange in the DFT functional, which was discussed above. The calculated DFT/PBE0 open singlet ΔE_{zz} gap is 1.9 eV in full agreement with the measured[6] ΔE_{zz} gap for the 3x6 (7, 12) AGNR. In contrast the ΔE_{zz} gap calculated with the PBE functional, which does not include "exact exchange", is less than half this value (~0.5 eV, in agreement with the PBE calculations of Wang et al.[6]). As we can see in Figs. 3(a), 3(b), and 3(c), which show the one-body DFT picture for the triplet, closed singlet, and open singlet, respectively, there are gaps in all of them between HOMO (or HOMO-1) and LUMO, which are equal or very nearly equal to the measured ΔE_{zz} value of 1.9 eV. We must rremember also that this value is practically equal to the correlation energy obtained from the difference between the GW and LDA values[14] for these AGNRs. Let us first focus on the open singlet, which is commonly accepted as the "ground state". Figure 3(d) is practically identical with figure(s) 2(b) and 2(c) of Wang et al.[6] where the definitions of ΔE_{zz} and ΔE_{ac} (which are designated as Δ_{zz} and Δ_{ac} respectively) are illustrated. Moreover, the calculated DFT/PBE0 ΔE_{zz} and ΔE_{ac} values (contrary to those of DFT/PBE, with no "exact exchange", used by Wang et al.[6]) are practically identical to the measured values for the 3x6 (7,12) AGNR deposited on non-metallic NaCl substrate (in full analogy to similar results for the 5-AGNRs, described above). Based on the closed singlet ground state, it becomes clear that the "real" energy separation of the end states is the HOMO-LUMO gap of the closed singlet state which is (almost always) about 0.1-0.3 eV (depending on the length). This is corroborated by the TDDFT results, giving rise to the $\Delta\varepsilon_{\zeta\zeta}$ gap, which discussed earlier for the 5-AGNRs, and is consistent with the appearance of Dirac points close and around the Fermi level, whereas ΔE_{zz} is due to mixed transitions involving both "end" and "bulk" states. This is verified by Fig. 3(d) which shows the excitation spectrum of the closed singlet state for the normal 3x6 (7,12) AGNR, in which there are two characteristic maxima at 1.9 eV and 3.2 eV, which practically coincide with the measured ΔE_{zz}, and ΔE_{ac} values respectively for this AGNR.[6] As we can see in the left part of Fig. 3(d), ΔE_{ac} involves transitions between (mixtures of) "bulk states" (from

HOMO-1 and HOMO-2, to LUMO+1 and LUMO+2), whereas ΔE_{zz} corresponds to transitions from mixt , "bulk" + "surface" (HOMO-3 and HOMO) to LUMO+3 and LUMO. Thus, ΔE_{zz}, although not equal to energy separation of the end states, is clearly associated with the first (lowest energy) transition involving end and bulk states, corresponding to the measured ΔE_{zz} value of 1.9 eV and the magnitude of the open singlet gap.

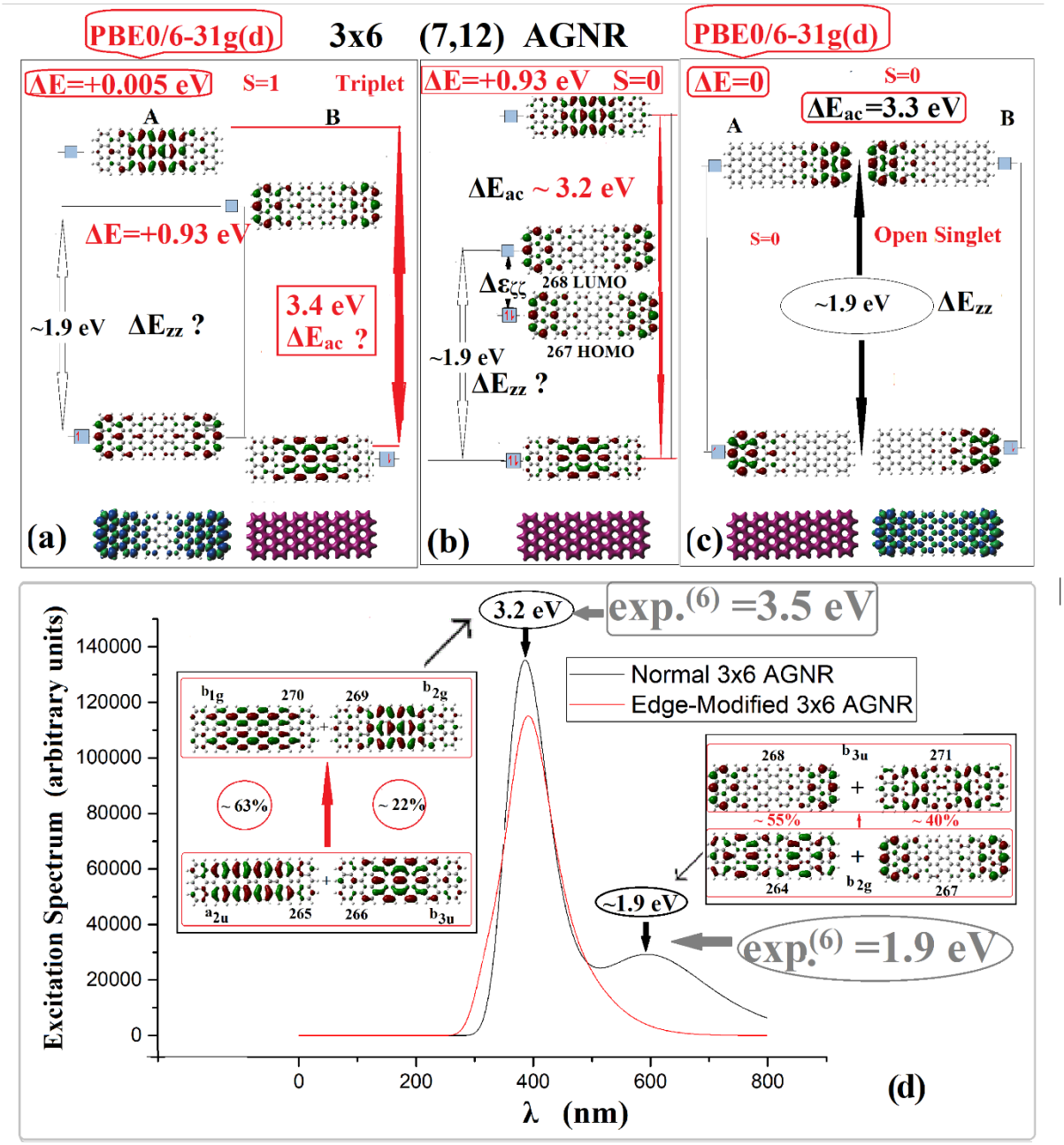


FIGURE 3. Spin states of the 3x6 (7, 12) AGNR: (a) Triplet, (b) Closed Singlet, and (c) Open Singlet states, showing frontier MOs, and gaps, together with charge density and spin density (see

text). (d): Excitation spectrum for the standard (black line) and edge-modified (red line on line) AGNRs. Intensity is given in arbitrary units, and excitation energy in eV. Boxes (rectangular and elliptical) emphasize the calculated values in agreement with experimental measurements. Question marks indicate alternative possibilities and/or identifications of the corresponding gaps (see text).

This is also supported by the TDDFT results in Fig. 3(d) showing the spectrum of the edge-modified closed singlet in which the peak of 1.9 eV is totally absent, whereas the peak of the "bulk" gap ΔE_{ac} is present and identical to the 3.2 eV peak of the normal (7,12) AGNR. The position of the ΔE_{ac} peak, contrary to ΔE_{zz}, changes (decreases) as the length increases. Thus, for the 3x14 (7,28) AGNR we found a ΔE_{ac} value of 2.8 eV, as is shown in Fig. S2(a). This value of 2.8 eV, as could be expected, is in perfect agreement with the calculated GW value[6] and the experimental measurements for the (7, 24-28) AGNR(s) on insulating NaCl substrate.[6]

3.3 9-AGNRs. We can observe in Fig. S2(c) that the overall spectrum of the 4x6 AGNR which has the same length with the 3x6 AGNR, except for a suppression of the $\Delta\varepsilon_{\zeta\zeta}$ peak, looks at a first sight very much alike the one for the 3x6 AGNR. Clearly a (deep) "bulk" gap could be expected not to vary very much or be so sensitive to the exact AGNR's width; but for the peak around 2.0 eV, which up to now was associated with the ΔE_{zz} gap of the 3x- AGNRs, further investigation is needed, which is described in Fig. 4. Figures 4(a), 4(b), and 4(c) are the corresponding analogues of figures 3(a), 3(b), and 3(c) respectively. However, contrary to the 7-AGNRs, the experimental data for the 9-AGNRs are very limited.[16] Therefore most of the results shown in Fig. 4 should be considered as predictions of the present work. As we can see in Fig. 4(c) for the open singlet the two fundamental gaps ΔE_{zz} and ΔE_{ac} are very close together (2.2 eV and 2.4 eV respectively) and not exactly equal to the corresponding 3x6 gaps This is also true for the almost

equal values of ΔE_{zz} and ΔE_{ac} obtained from triplet (and closed singlet). Thus, the peak around 2.0 eV in Fig. 4(d) is the result of the overlap of the ΔE_{zz} and ΔE_{ac} gaps, whereas the peak around 3.2 eV in the same figure, Fig. 4(d), although of "bulk" type (similarly to the 3x6 AGNR) is not the smallest "bulk" gap, and the real ΔE_{ac} for the 4x6 (9,12) AGNR should be around 2.0 eV.

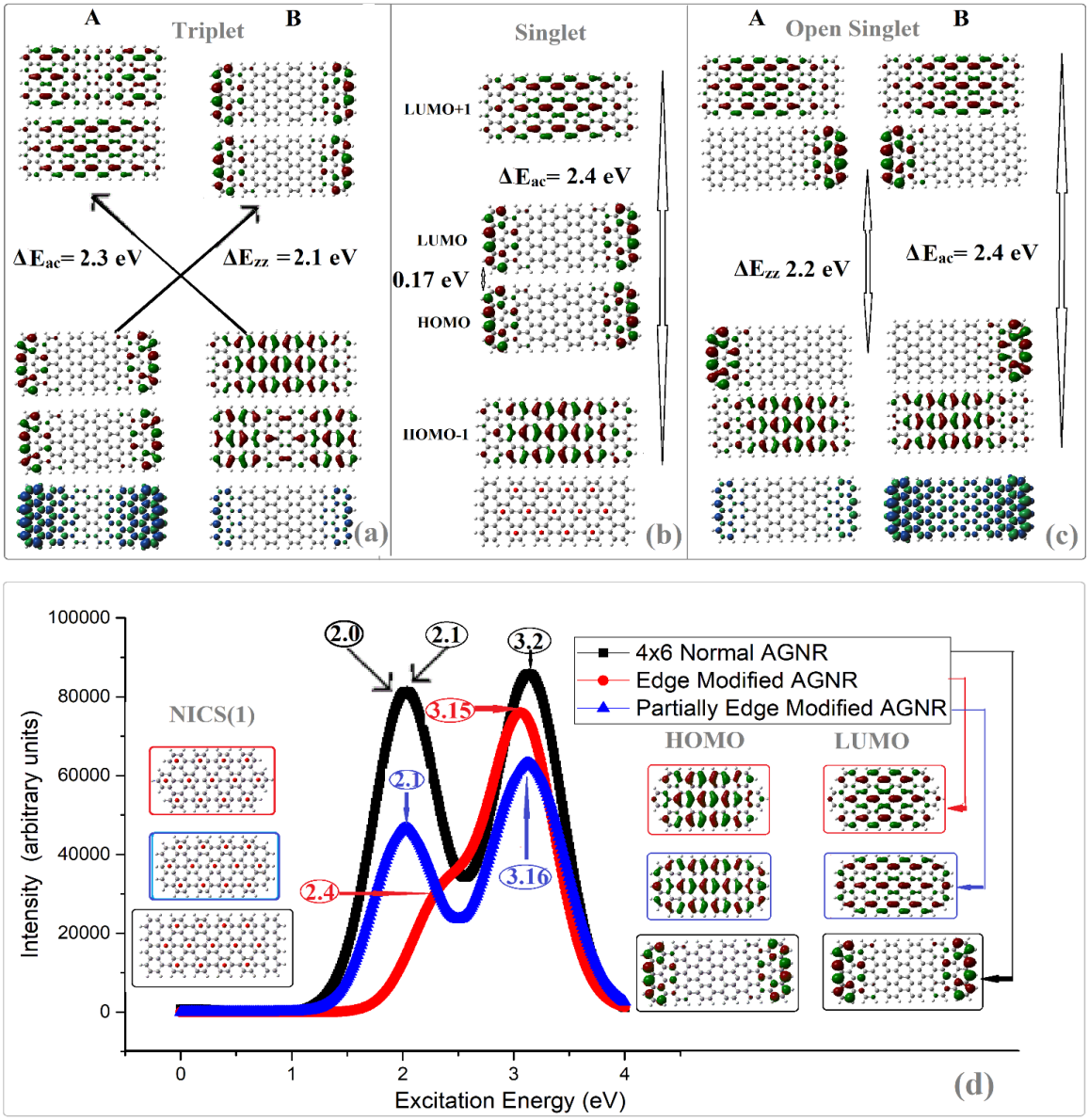


FIGURE 4. Spin states of the 4x6 (9, 12) AGNR: (a) Triplet, (b) Closed Singlet, and (c) Open Singlet, showing frontier MOs, gaps, and spin densities; (d)Excitation spectrum for the standard (black line), partially edge-modified (blue line on line), and fully edge modified AGNRs (red line on line). Intensity is given in arbitrary units, and excitation energy in eV. The frontier orbitals and

aromaticity patterns of partially and fully edge-modified AGNRs are shown in the right and left portions of the figure.

This is verified in Fig. 4(d), which shows that the 2.1 eV "bulk" peak (together with the "deeper" 3.2 eV "bulk" peak) survives the elimination (total and partial) of the empty (non-aromatic) end-rings which generates the edge modified AGNRs (without end states, and ΔE_{zz}). As is well known, this "bulk" peak value, decreases as the length of the AGNR increases. For the 4x13 AGNR we find ΔE_{ac} =1.6 eV, but for the longer 4x18 (9, 36), and 4x24 (9, 48) AGNRs of lengths L ≈ 78Å, and L≈104 Å respectively, we obtain (by TDDFT) for both of them ΔE_{ac} = 1.45 eV. This value is in very good agreement with the recently measured gap of 1.4 eV by Talirz *et al.*[16], as is illustrated in Fig. S3. The peak at 1.45 eV is further verified by Fig. S4 which shows the spectrum of the edge-modified 4x24 AGNR. We can also clearly see in Figs. S3(b), S3(c) the "surface" ΔE_{zz} gap at about 2.1-2.2 eV. Thus, for the 9-AGNRs the predicted values for the gaps are ΔE_{ac} = 1.45±0.1 eV, and ΔE_{zz} = 2.1±0.1 eV.

4.Conclusions. We have achieved an excellent agreement (within 1% or less) with the measured STS gaps ("bulk" and "surface") for the known 5-, 7- and 9-AGNRs, although the "surface" gaps, as is illustrated in Table 1. Namely:

a) For the 5-AGNRs the measured[8] gap value is 0.85 eV. The calculated here gap with DFT/PBE0 is 1.07, whereas the TDDFT/PBE0 value is exactly 0.85 eV, indicating also that this is a ΔE_{zz} gap.
b) Moreover, the measured[13] 0.1 eV gap is recognized to fully coincide with the calculated here (by both DFT-TDDFT/PBE0) $\Delta\varepsilon_{\zeta\zeta}$ gap.
c) For the 7-AGNRs the measured[6,7,9,10,15] ΔE_{ac} gap of 2.3±0.2 eV coincides with the calculated here ΔE_{ac} gap (with both DFT-TDDFT/PBE).

d) Furthermore, for the (7, 28) AGNR the measured[6] and GW-calculated[6] 2.8 eV gap fully coincides with the calculated here ΔE_{ac} gap (with both DFT-TDDFT/PBE0), whereas for the (7,12) AGNR the measured[6] and calculated[6] ΔE_{ac} gap is ~3.2 eV.

e) The measured[6] ΔE_{zz} gap of 1.9 eV for the 7-AGNRs, (7,12), and longer, is clearly identical to the calculated here ΔE_{zz} gap of 1.9 eV (with both DFT-TDDFT/PBE0).

f) For the 9-AGNRs the only known (to the present author) measurement[16] for the gap is 1.4 eV. The present calculations (TDDFT/PBE0) yield a ΔE_{ac} value of 1.45 eV, and also predict ΔE_{zz}=2.1±01 eV, quite close to the corresponding gap for the 7-AGNRs.

Note that for the 5- and 9-AGNRs the ΔE_{zz} gaps are theoretical predictions of the present work, due to the lack of analogous experimental data.

TABLE 1. Calculated and measured gaps for the 5-, 7-, and 9-AGNRs (in eV). Numbers with asterisk denote the present values, while numbers in parenthesis indicate the reference numbers of the original works. Numbers in bold emphasize the agreement between theoretical and experimental results, whereas underlined numbers in italics indicated the results of GW calculations.[14]

AGNR	$\Delta\varepsilon_{zz}$ Calculated	$\Delta\varepsilon_{zz}$ Measured	ΔE_{zz} Calculated	ΔE_{zz} Measured	ΔE_{ac} Calculated	ΔE_{ac} Measured
5-	**0.1*** **(1)**	**0.1**(13)	**0.85** *	**0.85**(8)	1.1*, *1.7*(14)	-
7-	0.1* (1)	-	**1.9***	**1.9**(6), 2.5(7)	**2.5***, **2.8***, 2.8(6), *3.7*(14)	**2.8**(6), **2.5**±0.2(6,7,9,15)
9-	0.1* (1)	-	2.2*	-	**1.45***, 1.6 (1), *2.0*(14)	**1.4** (16)

*Values obtained in the present work.

At the same time the present work has provided a simple physical understanding/rationalization of the origin and properties of these gaps. We have shown that such excellent agreement can be obtained by a transparent approach, using a minimum of computational resources, avoiding high level many-body methods, such as the advanced GW approach.[14] This is accomplished by

recognizing (and suitably exploiting) the fact that most (or all) of the Coulomb correlation energy is devoted to offset the (inversion) symmetry conflict. This is an added insight. Thus, simple DFT calculations with or without "fictitious" open shell states (such as open shell singlets or triplets) can give accurate results, especially when augmented by TDDFT calculation which can further refine the results, provided that the chosen DFT functional includes the "exact" exchange (such as the PBE0 functional[23] used here, proven to provide excellent results[1-5, 24]), and the finite length of the AGNRs is taken into account (recall the synopsis of the theoretical approach in section 2.5). Under the same provisions (finite size of AGNRs, and "exact exchange") the GW approach would also give the correct results, as is illustrated in ref. 6, where taking account the finite size of the 7-AGNRs has lowered the GW gap by about 1 eV, in very good agreement with the measured STS value. As a result, a similarly large overestimation of the expected substrate screening would be avoided, since the GW results[14] are widely used as reference values for the free standing AGNRs. This is corroborated by STS measurements of AGNRs on non-metallic substrates.[6-7] Thus, the measured STS gaps are practically independent of the substrate and virtually equal to the free-standing values, obtained by any of the three computational methods: DFT, TDDFT, and GW (from the simplest to the more complex), provided the finite size and the "exact" exchange are taken into account. Obviously, the simplest (and computationally most economical) approach should be normally preferred, in accord also with Occam's principle. A combination of DFT and TDDFT, as is used here, should be considered ideal.

Additional supplementary material with more details and comparisons for the spectra of 5-, 7-, and 9-AGNRs is given in the ***Supplementary Information.***

References

1. Zdetsis, A. D.; and Economou, E.N. Rationalizing and Reconciling Energy Gaps and Quantum Confinement in Narrow Atomically Precise Armchair Graphene Nanoribbons. *Carbon* **2017**, 116, 422-434.
2. Zdetsis, A. D. Bridging the Physics and Chemistry of Graphene(s): From Hückel's Aromaticity to Dirac's Cones and Topological Insulators. *J. Phys. Chem. A* **2020**, 124, 976-986
3. Zdetsis, A. D. Do We Really Understand Graphene Nanoribbons? A New Understanding of the 3n, 3n±1 Rule, Edge "Magnetism" and Much More. *J. Phys. Chem. C* **2020**, 124, 7578−7584
4. Zdetsis A. D.; Economou, E. N. Topological Metal-Insulator Transition in Narrow Graphene Nanoribbons? *Carbon* **2021**, 176, 548-557
5. Zdetsis, A. D. 4n+2=6n? A Geometrical Approach to Aromaticity? *J. Phys. Chem. A* **2021**, 125, 6064-6074
6. Wang, S. *et al*. Giant edge state splitting at atomically precise graphene zigzag edges. *Nat. Commun.* **2016** 7 : 1150. DOI: 10.1038/ncomms115077
7. Kolmer, M. et al. Rational synthesis of atomically precise graphene nanoribbons directly on metal oxide surfaces. *Science* **2020**, 369, 571-575DOI: 10.1126/science.abb8880
8. Lawrence, J. *et al*. Probing the Magnetism of Topological End States in 5-Armchair Graphene Nanoribbons. *ACS Nano* **2020**, 14, 4499–4508
9. Chen, Y-C; Oteyza, D. G.; Pedramrazi, Z; Chen, C.; Fischer, F. R.; Crommie, M. F. Tuning the Band Gap of Graphene Nanoribbons Synthesized from Molecular Precursors. *ACS Nano*. **2013**, 6123–6128.
10. Song, S. *et al*. On-surface synthesis of graphene nanostructures with π-magnetism *Chem. Soc. Rev.* **2021**, 50, 3238-3262
11. Nakada K.; Fujita M.; Dresselhaus G.; Dresselhaus M. S. Edge state in graphene ribbons: Nanometer size effect and edge shape dependence. *Phys. Rev. B* **1996** 54, 17954

12. Zhang, H.; Lin, H; Sun, K. *et al*. On-surface synthesis of rylene-type graphene nanoribbons. *J. Am. Chem. Soc.* **2015** 137, 4022–4025.
13. Kimouche, A; Ervasti, M. M.; Drost, R.; Halonen, S.; Harju, A.; Joensuu, P. M.; Sainio, J.; and Liljeroth, P. Ultra-narrow metallic armchair graphene nanoribbons. *Nature Comm.* **2015**, 6:10177, 1-6.
14. Yang, L.; Park, C.-H.; Son, Y.-W.; Cohen, M. L.; Louie, S. G. Quasiparticle energies and band gaps in graphene nanoribbons. *Phys. Rev. Lett.* **2007**, 99, 186801
15. Ruffieux, P.; Cai, J.; Plumb, N. C.; Patthey, L.; Prezzi, D.; Ferretti, A.; Molinari, E.; Feng, X.; Müllen, K.; Pignedoli, C. A.; et al. Electronic Structure of Atomically Precise Graphene Nanoribbons. *ACS Nano* **2012**, 6, 6930–6935.
16. Talirz, L.; Söde, H; Dumslaff, T. *et al*. On-Surface Synthesis and Characterization of 9-Atom Wide Armchair Graphene Nanoribbons. *ACS Nano*. **2017**, 11, 1380-1388.
17. Overbeck, J. *et al*. A Universal Length-Dependent Vibrational Mode in Graphene Nanoribbons. ACS Nano **2019,** 13, 13083-13091
18. Varykhalov, A.; Sánchez-Barriga, J.; Shikin, A. M.; Biswas, C.; Vescovo, E.; Rybkin, A. ; D. Marchenko, D.; Rader, O. Electronic and Magnetic Properties of Quasifreestanding Graphene on Ni. *Phys. Rev. Lett.* **2008**, 101, 157601.
19. Zdetsis. A. D. The real structure of the Si_6 cluster, *Phys. Rev. A* **2001,** 64, 23202
20. Son, Y.-W. Cohen, M. L.; Louie, S. G. Energy Gaps in Graphene Nanoribbons. *Phys. Rev. Lett.* **2006**, 97, 216803
21. Huzak, M.; Deleuze, M.S.; Hajgató B. Half-metallicity and spin-contamination of the electronic ground state of graphene nanoribbons and related systems: An impossible compromise?. *J. Chem. Phys*. 2011, 135, 104704
22. Frisch, M. J. et al. Gaussian 09, revision C.01, Gaussian, Inc.: Wallingford, CT, 2009.

23. Adamo, C.; and Barone,V. Toward Reliable Density Functional Methods Without Adjustable Parameters: The PBE0 model. *J. Chem. Phys.*, 1999, 110 6158-69.
24. GaussView, version 5.09, Dennington, R.; Keith, T.; Millam, J. *Semichem Inc.:* Shawnee Mission, KS, **2016**.
25. Zdetsis, A. D.; Niaz, S. From Zero to Infinity: Customized Atomistic Calculations for Crystalline Solids —Applications to Graphene and Diamond. *Adv. Mater. Lett.* **2021**, 12(9), 21091659. DOI:10.5185/amlett.2021.091659