Vertical and adiabatic excitations in anthracene from quantum Monte Carlo: Constrained energy minimization for structural and electronic excited-state properties in the JAGP ansatz

Nicolas Dupuy' , Samira Bouaouli' , Francesco Mauri' , Sandro Sorella' , and Michele Casula'

Citation: The Journal of Chemical Physics **142**, 214109 (2015); doi: 10.1063/1.4922048
View online: http://dx.doi.org/10.1063/1.4922048
View Table of Contents: http://aip.scitation.org/toc/jcp/142/21
Published by the American Institute of Physics

Vertical and adiabatic excitations in anthracene from quantum Monte Carlo: Constrained energy minimization for structural and electronic excited-state properties in the JAGP ansatz

Nicolas Dupuy,[1,a)] Samira Bouaouli,[2,b)] Francesco Mauri,[3,c)] Sandro Sorella,[4,d)] and Michele Casula[3,e)]

[1]*Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, Université Pierre et Marie Curie, case 115, 4 place Jussieu, 75252 Paris Cedex 05, France*
[2]*Laboratoire de Chimie Théorique, Université Pierre et Marie Curie, case 115, 4 place Jussieu, 75252 Paris Cedex 05, France*
[3]*CNRS and Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, Université Pierre et Marie Curie, case 115, 4 place Jussieu, 75252 Paris Cedex 05, France*
[4]*International School for Advanced Studies (SISSA), Via Beirut 2-4, 34014 Trieste, Italy and INFM Democritos National Simulation Center, Trieste, Italy*

(Received 18 December 2014; accepted 21 May 2015; published online 5 June 2015)

We study the ionization energy, electron affinity, and the $\pi \rightarrow \pi^*$ (^1L_a) excitation energy of the anthracene molecule, by means of variational quantum Monte Carlo (QMC) methods based on a Jastrow correlated antisymmetrized geminal power (JAGP) wave function, developed on molecular orbitals (MOs). The MO-based JAGP ansatz allows one to rigorously treat electron transitions, such as the HOMO → LUMO one, which underlies the ^1L_a excited state. We present a QMC optimization scheme able to preserve the rank of the antisymmetrized geminal power matrix, thanks to a constrained minimization with projectors built upon symmetry selected MOs. We show that this approach leads to stable energy minimization and geometry relaxation of both ground and excited states, performed consistently within the correlated QMC framework. Geometry optimization of excited states is needed to make a reliable and direct comparison with experimental adiabatic excitation energies. This is particularly important in π-conjugated and polycyclic aromatic hydrocarbons, where there is a strong interplay between low-lying energy excitations and structural modifications, playing a functional role in many photochemical processes. Anthracene is an ideal benchmark to test these effects. Its geometry relaxation energies upon electron excitation are of up to 0.3 eV in the neutral ^1L_a excited state, while they are of the order of 0.1 eV in electron addition and removal processes. Significant modifications of the ground state bond length alternation are revealed in the QMC excited state geometry optimizations. Our QMC study yields benchmark results for both geometries and energies, with values below chemical accuracy if compared to experiments, once zero point energy effects are taken into account. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4922048]

I. INTRODUCTION

The peculiar low-lying excited-states properties of the π-conjugated hydrocarbons (πCH) are at the heart of fundamental biochemical processes, which govern phenomena such as photo-synthesis and vision. The πCHs spectrum is commonly peaked in the energy range of visible light, which makes the πCH-based compounds photo-active. They are thus natural knobs for complex processes driven by the interaction of photons with matter. For instance, πCHs are major actors in light harvesting and energy transfer mechanisms and are functional building blocks of carotenoids,[1] green fluorescent chromophores,[2] and retinal proteins such as rhodopsin.[3] Moreover, in technological applications, πCH-based dyes have been effectively used to improve the power-conversion efficiency of dye-sensitized solar cells.[4,5]

An interesting and fundamental aspect behind the functional role played by the πCHs is the interplay between the electronic states and their geometries. A π-conjugated electronic structure typically leads to the so-called bond length alternation (BLA) in the carbon-carbon distance of the linear or cyclic molecular geometries. However, important structural changes may happen upon photo-excitation, with a significant modification of the BLA.[6] An illuminating example is provided by the rhodopsin photochromic interconversion responsible of the twilight vision of vertebrates. BLA changes between the S_0 and photo-excited S_1 states, and highly modulates the π-bond destruction and reconstruction, allowing an efficient and ultra-fast (~100 fs) photoisomerization between the rhodopsin and bathorhodopsin, its *trans* analog,[7] via a conical intersection.

In order to understand and model these fundamental processes, a very robust theoretical framework is necessary, able

a)nicolas.dupuy@impmc.upmc.fr
b)samira.bouaouli@lct.jussieu.fr
c)francesco.mauri@impmc.upmc.fr
d)sorella@sissa.it
e)michele.casula@impmc.upmc.fr

to resolve not only energy differences with an accuracy below 0.1 eV but also to describe structural modifications across excited states. Photo-driven excitations in πCHs, such as the $\pi \rightarrow \pi^\star$ transitions, are particularly challenging, as a highly correlated theory is usually needed. Indeed, post-Hartree Fock (HF) methods, such as multi-reference perturbation theory (CASPT2), symmetry-adapted cluster-configuration interaction (SAC-CI),[8] and quantum Monte Carlo (QMC) methods,[2,9] have been used to compute $\pi \rightarrow \pi^\star$ vertical excitations and compare their outcome against cheaper but less accurate time dependent density functional theory (TDDFT), including a large variety of functionals. The disagreement between theories can be as large as 0.5 eV[8] and strongly dependent on the reference geometry.[1] If compared to experiment, TDDFT is affected by substantial errors, showing up in the $\pi \rightarrow \pi^\star$ excitation energies, particularly for long π-conjugated chains. Indeed, a well known TDDFT failure is the ^1L_a (HOMO-LUMO) excitation in linear oligoacenes, polycyclic aromatic hydrocarbons (PAH) characterized by a zig-zag edge. The TDDFT discrepancy with respect to the experiment gets larger as the number of rings in the chain increases,[10] owing to a poor description of charge-transfer energies. Moreover, a direct comparison with experiments is also limited by structural relaxation effects between the ground and excited states. These lead to sizable differences between vertical and adiabatic excitations of up to a few tenths of eV. Relaxation energies are usually computed at the density functional theory (DFT) level, as in post-HF methods ionic gradients are time consuming and not readily available for the highest levels of theory. A quite strong dependence of the quality of the BLA on the DFT functional choice has been pointed out in Ref. 1.

In this paper, we present a QMC approach which is able to provide not only accurate energies but also precise relaxed geometries in both ground and excited states, at the variational level. The possibility to obtain optimal geometries of excited states in a highly correlated framework is an appealing and promising feature of our method, which allows us to take into account non-trivial structural modifications across the electron excitations and to compare our results to experimental adiabatic excitation energies. Very recently, a similar route has been taken by Guareschi and Filippi,[9] who assessed the quality of QMC optimized geometry in the excited states of small organic molecules. Their method, however, differs significantly from ours. Indeed, their work is based on a many-body wave function written as a linear combination of symmetry-adapted configuration state functions (CSFs)—which is the most general form—while the determinantal part of our variational wave function is constrained to take a symmetry-adapted antisymmetrized geminal power (AGP) form. On one hand, the AGP allows one to exploit its efficient determinant evaluation and some peculiar features of the geminal expansion in atomic orbitals, such as its locality properties. On the other hand, optimizing the AGP at chosen molecular orbital (MO) configurations requires a fixed-rank projection of the geminal function in order to have a stable and accurate energy minimization of ground and excited states.

We apply our QMC approach to study one of the simplest molecules in the oligoacene family, i.e., the three-ring anthracene molecule, which retains all the difficulties of PAHs, and more generally of πCHs, while being a strict test case for any new theory, as accurate benchmark data are available both experimentally[11–15] and theoretically.[10,16–19] Indeed, from the theoretical point of view, the small size of the molecule allows for an accurate extrapolation on both theory (up to the coupled cluster single double (triple) (CCSD(T)) level) and basis set (up to aug-cc-pVQZ), at a given frozen geometry.[16,17] Moreover, high-quality X-ray diffraction data are available for the ground state geometry[20]—although in the crystal condensed phase—while information on structural modifications across the ^1L_a transition is provided by ultrahigh-resolution spectroscopy,[14] probing the rotovibrational spectrum in the gas phase. We show that the quality of the electronic QMC wave function is of crucial importance to get accurate ionic equilibrium positions. The possibility to relax the geometry in both ground and excited states is in turn necessary to reach the chemical accuracy for adiabatic excitation energies.

Beside its importance as benchmark molecule, anthracene is interesting also because it belongs to the acene PAH group, which has quite strong correlation effects, and shows non-conventional features as a function of the chain length, such as a strong multiradical character.[21] The presence of many competing states in the long length limit could trigger the formation of an open-shell singlet ground state[22] and shrink the spin-triplet gap.[23,24] Moreover, in this context, a quite interesting link can be made with the physics of graphene nanoribbons, where a spin polarized edge should appear in the zig-zag configuration.[25]

The paper is organized as follows. In Sec. II A, we describe the Jastrow correlated antisymmetrized geminal power (JAGP) wave function used throughout our work, and in Sec. II B, we report the computational details of our calculations. In Sec. II C, we present the method employed to minimize the energy and optimize the geometry within a molecular orbital framework, useful to treat both ground and excited states by QMC. In Sec. III, we show our results for the anthracene ground state (Sec. III A), its ionization energy (IE) (Sec. III B), electron affinity (EA) (Sec. III C), and the low-lying ^1L_a neutral excited state (Sec. III D). Finally, we close the section by a thermodynamic comparison of the zig-zag anthracene molecule with phenanthrene, its armchair-like analog (Sec. III E). The conclusions and perspectives of this work are drawn in Sec. IV.

II. METHODS

A. Pseudopotentials, wave function, basis set

All the calculations are carried out for a first-principles Hamiltonian with core electrons replaced by pseudopotentials. The carbon atom is described by a HF energy consistent pseudopotential with scalar relativistic corrections by Burkatzki *et al.*,[26] while the Coulomb singularity of the hydrogen electron-ion potential has been replaced by a short-range non-diverging pseudopotential obtained within the same HF energy-consistent framework.[27] The HF energy-consistent pseudopotentials are particularly suited for correlated quantum chemistry calculations.

The quantum Monte Carlo variational ansatz which minimizes the energy of the first-principles Hamiltonian is the JAGP wave function ($\Psi_{\mathrm{JAGP}} = J\Psi_{\mathrm{D}}$), applied so far to a large variety of systems, and described in length elsewhere.[28–31] It yields very accurate variational energies and structural properties, by providing a correct description of dynamical and static correlations.

The AGP part has the determinantal form $\Psi_{\mathrm{D}} = \det(A_{ij})$. Given a system with $N^{\uparrow}$ up-spin electrons, and $N^{\downarrow}$ down-spin electrons (with the convention of $N^{\uparrow} \geq N^{\downarrow}$), A is a matrix with symmetric entries $A_{ij} = \Phi(\mathbf{r}_i^{\uparrow}, \mathbf{r}_j^{\downarrow})$ for $1 \leq \mathrm{i,j} \leq N^{\downarrow}$, and $A_{ij} = \Phi_j(\mathbf{r}_i^{\uparrow})$ for the unpaired electrons with $N^{\downarrow}+1 \leq \mathrm{i} \leq N^{\uparrow}$, where $\{\mathbf{r}_i^{\sigma}\}$ are the electron spatial coordinates. Φ is called geminal, or pairing function, while Φ_j are the unpaired orbitals. In our formulation, Φ can be either expanded on a localized basis set $\mathcal{G}_i$, taking a valence bond (VB) representation, or developed on MOs ϕ_k^{MO}, within the framework of a MO picture,

$$\Phi(\mathbf{r},\mathbf{r}') = \sum_{i,j=1}^{N_{\mathrm{basis}}} \lambda_{ij}\mathcal{G}_i(\mathbf{r})\mathcal{G}_j(\mathbf{r}') = \sum_{k=1}^{N_{\mathrm{MO}}} \lambda_k^{\mathrm{MO}}\phi_k^{\mathrm{MO}}(\mathbf{r})\phi_k^{\mathrm{MO}}(\mathbf{r}'), \tag{1}$$

where N_{basis} is the full basis set size, comprising the local orbitals centered on all atoms, and N_{MO} ($= N_{\mathrm{basis}}$) is the total number of MOs. One can easily pass from the VB to the MO picture by diagonalizing the VB λ_{ij} representation, yielding λ_k^{MO} as eigenvalues and ϕ_k^{MO} as eigenvectors, and from the MO back to the VB picture, by expanding the MOs in their local basis set components, $\phi_k^{\mathrm{MO}} = \sum_{i=1}^{N_{\mathrm{basis}}} \alpha_i^k \mathcal{G}_i$. Analogously, the unpaired orbitals Φ_j are expressed in the same local basis as $\Phi_j = \sum_{i=1}^{N_{\mathrm{basis}}} \beta_i^j \mathcal{G}_i$. It is worth noting that the AGP ansatz is not restricted to symmetric matrices λ_{ij}, but can be extended to generic real matrices λ. In the latter case, the diagonalization is performed via a singular eigenvalue decomposition having the spin-down (spin-up) MOs as left (right) eigenvectors (see Sec. II C). In this work, however, we are dealing with restricted spin-symmetric states; therefore, the geminal is a singlet, and the λ's are always symmetric.

In the case $N_{\mathrm{MO}} < N_{\mathrm{basis}}$, the relation in Eq. (1) is only approximated. However, taking the MOs with the largest $|\lambda_k^{\mathrm{MO}}|$ is usually enough to have very accurate wave functions, with a significant reduction in the total number of orbitals.

The minimum number of MOs is $N_{\mathrm{MO}} = N^{\downarrow}$, namely, the rank of λ is exactly equal to the number of opposite-spin pairs in the system, and the AGP ansatz reduces to a single Slater determinant. On the other hand, $N_{\mathrm{MO}} > N^{\downarrow}$ gives larger ranks, where excited state orbitals can contribute to lower the ground state energy, particularly in the situation of near-degenerate electronic shells, or perfect orbital degeneracy due to symmetry constraints. The system is then "resonating" between different configurations, and the ansatz, correlated by the Jastrow factor, is the *ab initio* representation of the resonating valence bond (RVB) picture proposed by Pauling in the context of aromatic molecules,[32] and by Anderson in the framework of spin models and superconductivity.[33]

In this paper, we show that by playing with the orbital weights, related to their occupations, given by the λ_k^{MO} in Eq. (1), the AGP wave function allows not only for an accurate description of the ground state but also of excited states, as one can promote electrons from occupied MOs (for $k \leq N_{\mathrm{HOMO}}$) to empty ones (for $k \geq N_{\mathrm{LUMO}}$) in the geminal expansion. This flexibility will be used, for instance, to compute the ^1L_a ($\pi \rightarrow \pi^{\star}$) excitation in the anthracene (see Sec. III D).

The localized basis set $\mathcal{G}_i$ is made of Gaussian type orbitals (GTOs). The GTO primitive basis is $(10s8p4d)$ for carbon and $(8s4p)$ for hydrogen. The initial α_i^k, which define the starting MOs, are obtained by DFT calculations in the local density approximation (LDA),[34] performed within the same GTO basis set. The GTO exponents, MOs, and unpaired orbitals are then optimized in the QMC framework, after applying the Jastrow correlating factor to the AGP wave function.

The Jastrow factor is decomposed in one-, two-, and three-four body parts, $J = J_1 J_2 J_{3-4}$.

The two-body Jastrow term reads

$$J_2(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \exp\left(\sum_{i<j}^{N} u(r_{ij})\right), \tag{2}$$

with $N = N^{\uparrow} + N^{\downarrow}$, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ the distance between the i-th and j-th electrons. The function u in Eq. (2) is $u(r) = 0.5r/(1+\gamma r)$, with γ a variational parameter. It is a homogeneous function (translationally invariant), correlating pairwise the electrons. This allows one to fulfill the antiparallel-spin cusp conditions. We neglect the parallel-spin cusp conditions, which are of less importance because of Pauli repulsion. This avoids spin contamination induced by the use of spin dependent Jastrow factors, and the wave function is an eigenstate of both S^2 and S_z operators.[35]

The three-four body Jastrow term is defined as

$$J_{3-4}(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \exp\left(\sum_{i<j}^{N} \Phi^J(\mathbf{r}_i,\mathbf{r}_j)\right), \tag{3}$$

where $\Phi^J(\mathbf{r}_i,\mathbf{r}_j) = \sum_{l,m=1}^{N_{\mathrm{basis}}^J} g_{lm}\mathcal{G}_l^J(\mathbf{r}_i)\mathcal{G}_m^J(\mathbf{r}_j)$. The development of the function Ψ^J on the GTO basis set $\mathcal{G}_l^J$ of size N_{basis}^J is formally equivalent to the one in Eq. (1) for the determinantal part. At variance with Eq. (2), the J_{3-4} factor is inhomogeneous, as it depends on the nuclear positions through the Jastrow GTO basis $\mathcal{G}_l^J$. We distinguish between a purely 3-body contribution, when both $\mathcal{G}_l^J$ and $\mathcal{G}_m^J$ are centered on the same atom, from terms in the sum involving different atoms and different ions, thus called 4-body contributions.

Finally, the one-body term is

$$J_1(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \exp\left(\sum_{i=1}^{N}\sum_{l=1}^{N_{\mathrm{basis}}^J} h_l\mathcal{G}_l^J(\mathbf{r}_i)\right), \tag{4}$$

which involves one electron and one ion at the time. The $\mathcal{G}_l^J$ basis set is the same as the one of the 3-4 body term in Eq. (3). Note that Eq. (4) does not have any cusp, as there is no singularity in the electron-ion potential of the *ab initio* Hamiltonian, thanks to the use of non-diverging pseudopotentials.[26,27] The one-body term takes care of keeping the electron density optimal, because the density needs to be readjusted when affected by the many-body correlations introduced by Eqs. (2) and (3).

In the present calculations, we used a GTO basis set $\mathcal{G}_l^J$ made of $(2s2p1d)/[1s1p1d]$ contractions for carbon and $(2s\,2p)/[1s1p]$ contractions for hydrogen. We combine two different radial functions for each contracted Gaussian, namely, $\alpha e^{Z_1|\mathbf{r}-\mathbf{R}|^2} + \beta|\mathbf{r}-\mathbf{R}|e^{Z_2|\mathbf{r}-\mathbf{R}|^2}$ times the angular dependence, where $\mathbf{R}$ is the Gaussian ionic center, and α, β, Z_1, and Z_2 are variational parameters per contraction. We found that combining these two radial dependencies in the contracted basis set (CBS) helps to get a better Jastrow factor, with lower energy and variance. The exception is represented by the s Gaussian of carbon, based on the standard contraction of two GTOs with the same radial dependence, namely, $\alpha e^{Z_1|\mathbf{r}-\mathbf{R}|^2} + \beta e^{Z_2|\mathbf{r}-\mathbf{R}|^2}$.

B. Computational details

The initial ground state geometry of both the anthracene and phenanthrene molecules has been relaxed in plane-waves DFT calculations with the Perdew, Burke, and Ernzerhof (PBE) functional,[36] using the QUANTUM ESPRESSO package[37] and ultrasoft pseudopotentials for carbon and hydrogen atoms. We used a Gaussian smearing of 10^{-4} Ry and an energy cutoff of 60 Ry in a periodic box large enough to get converged results.

The PBE-DFT optimal geometries are the starting structural guess for further QMC calculations, performed with the TurboRVB package.[38] The initial ϕ_k^{MO} MOs are obtained by LDA-DFT calculations carried out by the TurboRVB DFT driver with the same GTO basis set and HF energy-consistent pseudopotentials as the ones in the subsequent QMC analysis (see Sec. II A for details on basis set and pseudopotentials).

Once the determinantal part Ψ_{D} (Eq. (1)) is initialized, the Jastrow factor is first optimized by changing the linear coefficients g_{lm} and h_l, with the contracted basis set $\mathcal{G}_l^J$ frozen and the 2-body parameter γ set to 0.5. This QMC energy minimization step is done with the stochastic reconfiguration (SR) method,[28,39,40] which is very cheap and efficient, as it requires only the calculation of the wave function derivatives with respect to the variational parameters. The energy minimization is very stable as the linear Jastrow coefficients are the easiest to optimize.

In the second QMC optimization step, all Jastrow parameters (linear coefficients, γ, and basis set exponents) are relaxed by means of the so-called "linear method" (LM),[41] which requires the calculations of some Hessian matrix elements to improve the convergence rate and the stability of the minimization.[41–43]

In the third step, also the determinantal orbitals are optimized by LM: α_i^k linear coefficients, β_i^k for the unpaired orbitals (if any), and Gaussian exponents, together with the fully relaxed Jastrow factor. The MOs evolve according to the algorithm devised in Ref. 30 and recalled in Sec. II C, driven by the wave function derivatives with respect to λ_{ij}, under the constraint of fixed-rank AGP.

Finally, in the fourth step, we perform the geometry relaxation, where the structural parameters are displaced by the ionic forces according to the steepest descent algorithm. The QMC ionic forces are computed with finite variance estimators and analytic derivatives, as described in Ref. 44. The electronic structure is changed on the fly in such a way that the ions follow the Born-Oppenheimer energy surface.

For the anthracene, we compute not only the ground state but also some selected excited states: the electron addition and removal excitations and the difficult case of the ^1L_a excited state, i.e., the $\pi \to \pi^\star$ (HOMO $\to$ LUMO) transition. For the ionic wave functions, we set $N^\uparrow = N^\downarrow + 1$. For the ground state, anion, and cation, we set $N^{\mathrm{MO}} = N^\uparrow$, namely, we use a single Slater determinant in the antisymmetric part of the wave function. For the ^1L_a excited state instead, $N^{\mathrm{MO}} = N^\uparrow + 1$, and the wave function is obtained by mixing the HOMO and LUMO obtained from DFT in a procedure described in Sec. III D. The starting geometry for all excited states is the optimal QMC one for the ground state. This allows us to compute both vertical and adiabatic excitation energies, once the excited state geometry is also relaxed.

Additional PBE, B3LYP, and CCSD(T) calculations (energy and geometry optimization) have been performed in a triple zeta Gaussian basis set (pVTZ) for anthracene and phenanthrene ground states with the carbon pseudopotential by Burkatzki *et al.*,[26] by using the Molpro *ab initio* quantum chemistry package.[45]

C. Excited states and their geometry relaxation by QMC methods

In Ref. 30, some of us introduced for the first time the JAGP ansatz developed on MOs, together with an efficient way to optimize them, constraining the rank of the geminal matrix λ during the QMC total energy minimization. The computational scheme devised there allows one to optimize both the Jastrow factor and the AGP in the MO representation, with the aim at reaching the closest variational wave function to the true ground state.

It is notorious that the QMC excited states are harder to compute and optimize than the ground state. Indeed, one needs to enforce some orthogonality conditions not to fall to the lowest variational state during the optimization. For wave functions belonging to different spin or particle number sectors, this is trivial, as the orthogonality is automatically imposed by construction, and the task reduces to find the lowest state within that particular sector. Thus, the same minimization algorithm used for the ground state can be applied also there. Things get more challenging when one looks for excited states with fixed spin and particle number, but with electronic part belonging to different irreducible representations of the same molecular point group. Indeed, a variational Monte Carlo (VMC) algorithm which optimizes also the determinantal orbitals by energy minimization is potentially capable to flip from one irreducible representation to another at lower energy. It is worth noting that it is precisely the determinantal part of the QMC wave function which sets the spatial symmetry of the many-body state, as the Jastrow term is a fully symmetric bosonic factor, and thus A_{1g} symmetry invariant. Therefore, optimizing the determinantal part by preserving its symmetry is of paramount importance if one would like to compute accurate excited state properties. Usually, the determinantal part is developed in terms of spin-adapted CSFs, which are built according to the selected

symmetry of the excited states and then multiplied by a common Jastrow factor.[2,46,47]

In recent years, an important progress has been done in this direction.[2,9,46,48] Energy minimization algorithms have been devised to optimize CSFs linear coefficients, single particle orbitals, and Jastrow factor all together. The LM has been effectively extended to fully optimize excited states with the lowest energy of a given symmetry.[2,47] This approach shows the same stability and robustness as the ground state minimization. Instead, getting the partial or full spectrum of a given symmetry is a more difficult and challenging case. However, even in this harder situation, significant methodological advancements have been made. The main route taken so far follows the idea that the orthogonality between different solutions can be enforced by a simultaneous QMC optimization of all excited states spanning a given manifold, an approach which keeps track of the overlap between them. The orbital optimization is then performed in a state-average framework, where one minimizes a global energy estimator summed over the exited states taken into account with appropriate weights.[2] In this case, the MC sampling is carried over a global fictitious distribution which is devised to have a finite overlap with all excited states in the manifold. Thus, reweighting is needed to compute single excited state properties, but this can become less efficient than a state-specific optimization. Moreover, the parameters are then stationary with respect to the weighted state-average and not to each single excited state. This breaks the zero-variance principle for each individual state. Despite these drawbacks, this approach has led to important results for the accurate description of chromophores and photoluminescence in biochemistry.

In this paper, we take a different route to compute excited state properties. Instead of dealing with a linear combination of symmetry-adapted CSFs, the determinantal part of our wave function is a symmetry-adapted AGP. Its form is of course less general than the former, and not all exited states could be reproduced within this ansatz. However, its interest is twofold. First, the multideterminant expansion implied by the AGP is evaluated at a computational cost of a single Slater determinant. Second and more importantly, the AGP expansion in terms of atomic orbitals allows one to exploit its locality properties by dramatically reducing the number of independent λ_{ij} parameters. Indeed, the λ_{ij}s depend implicitly on the ion-ion distance, through the orbital index and its corresponding nucleus, such as $\lambda_{ij} = \lambda_{ij}(\mathbf{R}_i - \mathbf{R}_j)$. On one side, the point group of the molecular geometry is used to reduce the independent λ_{ij}s, as commonly done in any quantum chemistry code. On the other side, the λ_{ij}s are explicitly optimized only below a given cutoff radius ($\forall i, j$ such that $|\mathbf{R}_i - \mathbf{R}_j| < R_{\text{cutoff}}$), as introduced in Ref. 49, while the "long-range" λ_{ij}s are initialized according to the starting DFT orbitals. In the case of anthracene and phenanthrene molecules presented in this work, $R_{\text{cutoff}} = 2.86$ Å, i.e., the explicitly optimized λ_{ij}s live in the nearest and next-nearest neighbor shells. This locality approach is supposed to be more and more effective as the system size increases. Three-ring PAHs have already a size where locality can play a critical role.

The flexible AGP atomic expansion is symmetry-adapted in its MO representation, and its energy minimization is constrained, based on the AGP-MO optimization scheme detailed in Ref. 30, which turns out to be useful not only to minimize the ground state energy but also to optimize excited states which belong to the same spin sector and different spatial irreducible representations. Under certain conditions, the method is supposed to provide the lowest variational energy within a given spin and spatial symmetry. In the parameter space, the excited state can be either a local minimum or a saddle point. While the former case is the easiest, the latter is not trivial, as the QMC noise can make the minimization to drift from the targeted symmetry. Our method, which is based on a constrained fixed-rank optimization of the AGP wave function, has proven to be stable with absent or insignificant symmetry contamination of the starting excited state symmetry.

Let us assume to work with the most general AGP, expressed in the VB representation (Eq. (1)) by means of an arbitrary non-symmetric λ matrix, which can be written as

$$\lambda_{ij} = \sum_{k=1}^{N_{\text{MO}}} \lambda_k^{\text{MO}} \alpha_i^k \bar{\alpha}_j^k, \tag{5}$$

where we have defined the spin-down MOs as $\bar{\phi}_k^{\text{MO}} = \sum_{i=1}^{N_{\text{basis}}} \bar{\alpha}_i^k \mathcal{G}_i$, while for the spin-up, we use the same definition as the one introduced in Sec. II A. Both spin-up and spin-down MOs are orthonormal, i.e., $\langle \phi_k^{\text{MO}} | \phi_{k'}^{\text{MO}} \rangle = \langle \bar{\phi}_k^{\text{MO}} | \bar{\phi}_{k'}^{\text{MO}} \rangle = \delta_{kk'}$. The derivation carried on here holds for orthonormal and non-orthonormal basis sets $\mathcal{G}_i$. In the case of non-orthonormal basis set, as the GTO one in our calculations, the overlap matrix

$$\sigma_{ij} = \langle \mathcal{G}_i | \mathcal{G}_j \rangle \tag{6}$$

must be computed and used whenever necessary, whereas for orthonormal basis sets, σ reduces to the identity. By introducing the extended $2N_{\text{basis}} \times 2N_{\text{basis}}$ symmetric matrices

$$M = \begin{pmatrix} 0 & \lambda \\ \lambda^\dagger & 0 \end{pmatrix} \text{ and } S = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}, \tag{7}$$

the spin-up and spin-down MOs and their related coefficients λ_k^{MO} are obtained from the generalized eigenvalue equation,

$$SMS\psi_\pm^k = \pm\lambda_k^{\text{MO}} S\psi_\pm^k, \tag{8}$$

where the pair of eigenvectors $\psi_\pm^k$ read in the vector notation,

$$\psi_\pm^k = \begin{pmatrix} \alpha^k \\ \pm\bar{\alpha}^k \end{pmatrix}. \tag{9}$$

Therefore, as already mentioned in Sec. II A, one can always go from the VB to the MO representation, diagonalizing Eq. (8) and keeping the first N_{MO} MOs with the largest eigenvalues $|\lambda_k^{\text{MO}}|$. In that case, the λ matrix, as written in Eq. (5), has clearly rank N_{MO}.

In order to minimize the variational energy of the JAGP wave function containing the λ_{ij} entries as variational parameters, the energy derivatives $\partial E/\partial\lambda_{ij}$ are computed in the QMC framework, where $E = \langle \Psi_{\text{JAGP}} | \hat{H} | \Psi_{\text{JAGP}} \rangle / \langle \Psi_{\text{JAGP}} | \Psi_{\text{JAGP}} \rangle$ is the expectation value of the Hamiltonian $\hat{H}$ on the variational wave function Ψ_{JAGP}. According to the SR or LM algorithm, and based on the derivatives information, the λ matrix evolves into $\lambda' = \lambda + \delta\lambda$ after an optimization step. One would like λ'

to have the same rank as λ to preserve the correlation level and symmetry of the wave function, while reducing the variational energy. As proven in Ref. 30, a necessary and sufficient condition for λ' to have rank N_{MO} up to the first order in the parameters' variations is

$$(\hat{I} - \hat{P}^{\uparrow})\delta\hat{\lambda}(\hat{I} - \hat{P}^{\downarrow}) = 0, \tag{10}$$

where $\delta\hat{\lambda} = \sum_{i,j=1}^{N_{\mathrm{basis}}} |g_i\rangle\delta\lambda_{ij}\langle g_j|$ is the variation of the geminal operator, $\hat{I} = \sum_{i,j=1}^{N_{\mathrm{basis}}} |g_i\rangle\sigma_{ij}^{-1}\langle g_j|$ is the identity in the non-orthogonal basis set, and $\hat{P}^{\uparrow} = \sum_{k=1}^{N_{\mathrm{MO}}} |\phi_k^{\mathrm{MO}}\rangle\langle\phi_k^{\mathrm{MO}}|$ and $\hat{P}^{\downarrow} = \sum_{k=1}^{N_{\mathrm{MO}}} |\bar{\phi}_k^{\mathrm{MO}}\rangle\langle\bar{\phi}_k^{\mathrm{MO}}|$ are the projectors on the manifold spanned by the spin-up and spin-down MOs, respectively. In order to satisfy the condition in Eq. (10), λ' must be accordingly projected as

$$\hat{\lambda}' - (\hat{I} - \hat{P}^{\uparrow})\hat{\lambda}'(\hat{I} - \hat{P}^{\downarrow}) = \hat{P}^{\uparrow}\hat{\lambda}' + \hat{\lambda}'\hat{P}^{\downarrow} - \hat{P}^{\uparrow}\hat{\lambda}'\hat{P}^{\downarrow}, \tag{11}$$

namely, those λ' components which leak outside the starting manifold on *both* spin-up and spin-down channels have to be filtered out, in order to be consistent with a rank N_{MO} optimization.

Imagine now that one is working with an initial $\hat{\lambda}$ belonging to a given spatial irreducible representation of the molecular point group. This is set by initializing the single particle MOs occupations via a proper combination of λ_k^{MO} with starting MOs which fulfill the spatial symmetries of the molecule. For instance, the starting MOs can be taken from DFT, such as in our case, or from whatever other theory which provides MOs without breaking the symmetries of the appropriate point group. After a step of energy minimization and fixed rank projection (Eq. (11)), the new geminal operator $\hat{\lambda}'$ is such that the variation with respect to the starting $\hat{\lambda}$ is

$$\delta\hat{\lambda} = \sum_k \left(|\phi_k\rangle\lambda_k\langle\delta\bar{\phi}_k| + |\delta\phi_k\rangle\lambda_k\langle\bar{\phi}_k| + |\phi_k\rangle\delta\lambda_k\langle\bar{\phi}_k|\right). \tag{12}$$

By diagonalizing Eq. (12) via the generalized eigenvalue problem in Eq. (8), one obtains new N^{MO} MOs with eigenvalues λ_k^{MO} of order 1 and further states with eigenvalues $\lambda_k^{\mathrm{MO}} \propto O(\epsilon^2)$ for $k > N^{\mathrm{MO}}$, where the order ϵ is set by the parameters variation $\delta\lambda$, tunable in the SR and LM minimization algorithms. We found that for ϵ (i.e., the parameter variation) properly tuned and for a large enough statistics, the evolved wave function never leaves the given starting manifold, and within the present scheme, one can successfully carry out constrained energy minimizations.

A caveat is in order here. In the case of degenerate states, one has to ensure that all states in the given degenerate manifold are included in the $\hat{P}^{\uparrow}$ and $\hat{P}^{\downarrow}$ projection operators, otherwise the diagonalization of the M' matrix could bring the new solution out of the starting manifold. As the AGP has a multi-reference character, all the degenerate manifold can be effectively and consistently included in the AGP expansion.

Here, we detailed the optimization algorithm that we have implemented, which is made of the following five iterative steps:

0. by DFT (or other lower-level method), generate the starting $\{\phi_k^{\mathrm{MO}}\}_{k=1,N_{\mathrm{MO}}}$ and $\{\bar{\phi}_k^{\mathrm{MO}}\}_{k=1,N_{\mathrm{MO}}}$ MOs for the spin-up and spin-down sectors, respectively, which belong to the targeted symmetry representation;
1. compute the derivatives $\partial E/\partial\lambda_{ij}$ for the current Ψ_{JAGP} variational wave function, with the AGP part developed on $\{\phi_k^{\mathrm{MO}}\}_{k=1,N_{\mathrm{MO}}}$ and $\{\bar{\phi}_k^{\mathrm{MO}}\}_{k=1,N_{\mathrm{MO}}}$ MOs which give the current λ matrix in Eq. (5);
2. constrain the derivatives of local energy and wave function by applying the condition in Eq. (10), namely,
$$\widetilde{\frac{\partial E}{\partial\lambda_{ij}}} = \frac{\partial E}{\partial\lambda_{ij}} - (I - P^{\uparrow})\frac{\partial E}{\partial\lambda_{ij}}(I - P^{\downarrow}), \tag{13}$$
after constructing the projectors $\hat{P}^{\uparrow} = \sum_{k=1}^{N_{\mathrm{MO}}} |\phi_k^{\mathrm{MO}}\rangle\langle\phi_k^{\mathrm{MO}}|$ and $\hat{P}^{\downarrow} = \sum_{k=1}^{N_{\mathrm{MO}}} |\bar{\phi}_k^{\mathrm{MO}}\rangle\langle\bar{\phi}_k^{\mathrm{MO}}|$;
3. apply the SR or LM algorithm to find the corresponding variation $\delta\lambda$, and update the AGP matrix as $\lambda' = \lambda + \delta\lambda$;
4. diagonalize the λ' matrix as in Eq. (8) to find the new $\{\phi_k^{\mathrm{MO}}\}_{k=1,N_{\mathrm{MO}}}$ and $\{\bar{\phi}_k^{\mathrm{MO}}\}_{k=1,N_{\mathrm{MO}}}$ MOs, by keeping only the first largest N_{MO} eigenvalues, and reconstruct the new λ matrix according to Eq. (5);
5. go back to point 1. and iterate till convergence in the variational energy.

The steps devised above for the MOs optimization (involving λ^{MO} and α) of ground and excited states can be combined with the optimization of other wave function parameters that do not change the rank N_{MO} of the matrix λ, such as all the ones corresponding to the Jastrow factor, but also the basis set exponents and the contraction coefficients. Indeed, in the SR or LM method, the variation of those parameters enters in the SR or Hessian matrix as additional degrees of freedom, which are treated on the same footing, via SR or Hessian inversion based on their related energy derivatives $\partial E/\partial a$ and local operators $\partial \log\Psi/\partial a$, where a is a generic wave function parameter. Therefore, in this generalized optimization method, in step (1), more derivatives are computed without any projection, and in step (3), the new set of parameters comprises more components than the ones related to the MOs.

This flexibility allows one to include also the ionic coordinates as additional parameters, implying that a structural relaxation is possible even for excited states. The difficulty of dealing with force estimators with finite variance and being efficiently computed has been recently overcome, thanks to a constant development in the field.[29,44,50–55] Therefore, we have been able to perform a complete optimization of the excited states, in both electronic structure and geometry, which yields vertical and adiabatic excitation energies within the JAGP level of theory.

III. RESULTS

A. Ground state energies and geometry

As explained in Sec. II B, the ground state QMC wave function has been initialized starting from the PBE geometry and LDA MOs. We take $N^{\mathrm{MO}} = N^{\uparrow}$ MOs, i.e., we deal with a single Slater determinant. In the first QMC step, the Jastrow factor is optimized with geometry and DFT Slater determinant frozen. At the second stage, a larger fraction of correlation energy is recovered by optimizing the Jastrow and MOs together. The MOs are evolved by following the algorithm presented in Sec. II C, where the energy gradient $\partial E/\partial\lambda_{ij}$ is calculated

TABLE I. Ground state energy and variance (σ^2) for the neutral anthracene molecule obtained by various levels of QMC optimization, with contracted and uncontracted basis sets (CBS and UBS, respectively). In the last row, both orbitals and geometry are optimized ("opt.") at the QMC level, by using an uncontracted basis set (UBS) for the determinant. The last but one row shows the intermediary step where only the determinant is optimized with the UBS with the geometry frozen at the level where it was optimized using the contracted basis step (CBS).

Orbitals/geometry	VMC (Ha)	σ^2 (Ha^2)	LRDMC (Ha)
LDA/PBE	−85.4417(5)	1.446(5)	...
Opt. CBS/PBE	−85.4580(5)	1.341(5)	...
Opt. CBS/opt. CBS	−85.4620(5)	1.344(4)	−85.6252(5)
Opt. UBS/opt. CBS	−85.4651(5)	1.320(2)	−85.6260(8)
Opt. UBS/opt. UBS	−85.4677(5)	1.310(4)	−85.6275(8)

on a contracted atomic basis set (one contraction per angular momentum channel), which dramatically reduces the size of the derivative matrix, and makes the convergence in energy faster. Unless otherwise specified, the results presented in this section are always obtained in the CBS.

In practice, the variation of the AGP $\delta\lambda$ is performed in a reduced Hilbert space (defined by the contractions), while the starting DFT MOs are defined in the full GTO primitive basis set ($(10s8p4d)$ for carbon and $(8s4p)$ for hydrogen). The new MOs are then obtained by diagonalizing the matrix λ' updated in the primitive basis set, solving the generalized eigenvalue problem in Eq. (8). The gain in energy yielded by the MOs optimization is significant (160 mHa), as reported in Table I. Also the variance is significantly reduced, as the variational wave function gets closer to the true ground state.

In the final optimization step (at the fixed contracted basis set), we relax the geometry, following the ionic forces according to the steepest descent algorithm. During the geometry relaxation, the electronic part of the wave function stays nearly optimal at each ionic configuration, as the electronic parameters evolve according to the linear method. Once the equilibrium geometry is reached, the energy is slightly lower than the VMC one at the LDA structure (by 40 mHa). This small energy change has, however, a large impact in the final QMC geometry, which gets closer to the experimental structure.

To check whether our VMC geometry is converged with respect to the basis set, we performed an additional optimization within an uncontracted basis set (UBS) for the electronic parameters. In this case, at variance with the former calculations, $\partial E/\partial\lambda_{ij}$ has been evaluated directly in the primitive basis set, which was chosen to be $(9s8p4d)$ for carbon and $(8s4p)$ for hydrogen, without use of contractions, and thus having a much larger flexibility. The corresponding VMC and lattice regularized diffusion Monte Carlo (LRDMC)[56] energies are reported at the last row of Table I. When we relax both the electronic and ionic parts in UBS, the energy lowers by 57 and 23 mHa in the VMC and LRDMC, respectively. Also the variance is further reduced, with respect to the CBS reference. After these calculations in the primitive basis set, we can safely conclude that our best VMC geometries are unbiased with respect to sizable basis set errors.

TABLE II. Carbon-carbon bond lengths for the anthracene molecule. For comparison, we report also the B3LYP/pVTZ and PBE/pVTZ results obtained for the same carbon pseudopotential. The bond label conventions are drawn in Fig. 1.

Bond label	Expt.[20] (Å)	QMC (Å)	B3LYP/pVTZ (Å)	PBE/pVTZ (Å)
1-2	1.356(9)	1.3570(3)	1.362	1.371
2-3	1.430(6)	1.4243(4)	1.424	1.425
3-4	1.400(9)	1.3899(4)	1.393	1.399
7-8	1.410(10)	1.4218(6)	1.420	1.421
3-12	1.435(7)	1.4328(4)	1.438	1.445

As one can see in Table II and Fig. 2, the QMC equilibrium geometry reproduces very well the experimental bond minimum ("1-2") and maximum ("3-12") elongations. In between the shortest and the largest bond lengths, their values fluctuate on the molecular zig-zag edge, as shown in Fig. 2. This so-called BLA is a characteristic feature of aromatic compounds and π-conjugated molecules. It depends on the π weight in the carbon-carbon double bond, whose strength is set by the superposition of allowed resonating structures, as predicted by Clar's theory.[57] It turns out that the JAGP wave function reproduces very accurately this feature, which plays a key role in many important biochemical processes, such as the conformers photo-activation in rhodopsin.[7]

For comparison, we performed also B3LYP and PBE DFT structural relaxation calculations in the pVTZ basis set with the Molpro package.[45] This improves upon previously published B3LYP geometries, obtained in smaller basis sets (cc-pVDZ[58] and 6-311G**[59]). Our calculations have been carried out for

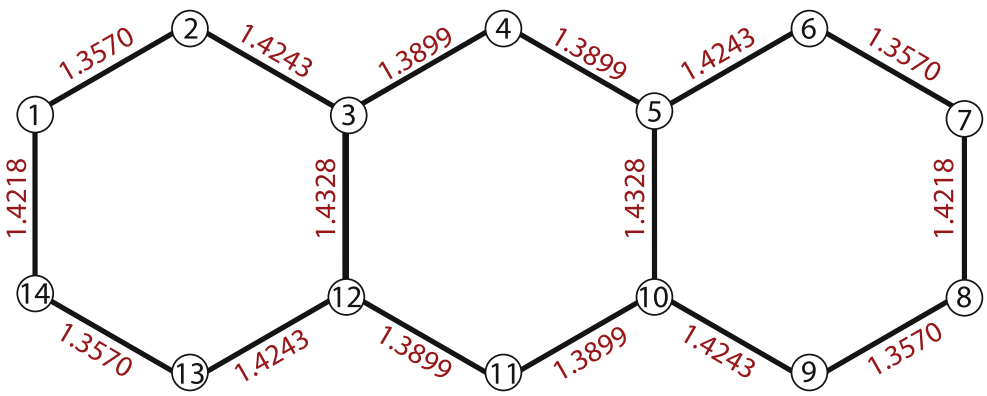

FIG. 1. Carbon site labels and QMC bond length in Å for the anthracene ground state.

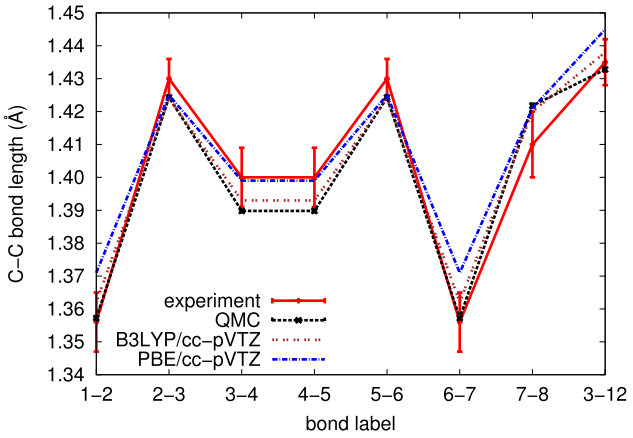

FIG. 2. Anthracene ground state experimental (Ref. 20), QMC, B3LYP /pVTZ, and PBE/pVTZ bond lengths. All theoretical values are obtained in the present work. The bond label conventions are reported in Fig. 1. We plot the error bars for both the experimental results and the QMC geometries. In the latter case, the error bars magnitude is comparable with the line thickness.

the same carbon pseudopotential as the one in QMC calculations. The B3LYP/pVTZ geometry is the closest to the QMC one, and this confirms the quality of the B3LYP functional in providing accurate structures. The PBE/pVTZ is systematically poorer than B3LYP/pVTZ if we take the QMC geometry as reference (see Fig. 2). The B3LYP itself has a slight tendency to overestimate the bond length, as already pointed out elsewhere,[60] but it is almost on top of the QMC results.

We conclude this section by saying that the experimental Ref. 20, so precious for validating the reliability of different theoretical methods to predict the ground state geometry, has a quite large error bar and is not completely bias-free either. Indeed, it comes from X-ray diffraction results obtained from various samples of anthracene crystal. The effect of the crystal environment could slightly affect the measured molecular equilibrium geometry with respect to the gas phase. On the other hand, we believe that Ref. 20 provides one of the best experimental estimates of the anthracene geometry, as it is an average over a quality selected ensemble of as many as 68 different samples, compiled from the Cambridge Structural Database.

B. Ionization energy

The ionization (or electron removal) potential is a spectroscopic property of paramount importance to benchmark new theoretical methods and characterize electro-optically active materials. In the anthracene, it involves an electron mainly taken from the π HOMO orbital, with a concomitant charge redistribution due to the partial change of bond character between carbons.

We compute the vertical IE by removing an electron from the ground state and optimizing the electronic part in the contracted basis set at frozen VMC ground state geometry. The VMC and LRDMC results are shown in Table III.

By relaxing the geometry, the gain in energy is quite large, about 0.1 eV. The equilibrium structure is reported in Table IV and plotted in Fig. 2. The impact of the electron removal on the geometry is significant, with the shortest bond ("1-2") which gets looser in the ionic state, and with a milder BLA. Also these features can be qualitatively explained within Clar's theory, as the number of resonance configurations is smaller, due to one electron less, and so there is less room to modulate the π weights of the carbon-carbon bonds.

TABLE III. Vertical and adiabatic ionization energies (IE) for the anthracene molecule by VMC and LRDMC calculations. We report also the results of Ref. 18, obtained at the B3LYP/6-311++ G(d,p) level, and the benchmark (nominally "exact") calculations based on the focal point analysis (FPA) of Ref. 16. The vertical and adiabatic "exact" FPA results do not include the zero point energy (ZPE).

IE	VMC (eV)	LRDMC (eV)
Vertical QMC	7.440(17)	7.476(22)
Adiabatic QMC	7.338(14)	7.385(20)
Expt.	7.415 eV[11]	
Vertical B3LYP	7.16 eV[18]	
Adiabatic B3LYP	7.09 eV[18]	
Vertical "exact" (FPA)	7.47 eV[16]	
Adiabatic "exact" (FPA)	7.40 eV[16]	

The projective LRDMC has a tendency of yielding larger IE than VMC, for both vertical and adiabatic excitations. Both LRDMC IE values are in better agreement with the focal point analysis (FPA) performed in Ref. 16, which provides benchmark quantum chemistry results for vertical and adiabatic ionization energies by extrapolating to both complete basis set (up to 1070 basis set elements for anthracene) and theory (up to the CCSD(T) level). The slight improvement of LRDMC with respect to the VMC results, if compared to the FPA benchmark, could be due to a weak residual basis set bias in VMC. This effect is gone in LRDMC, because it is basis set insensitive.

Our best vertical IEs are in a reasonable agreement with the experimental value, although slightly underestimated, by 75 and 30 meV at VMC and LRDMC levels, respectively. These differences are almost of the order of the error bar in LRDMC calculations. LRDMC confirms the tendency of a 0.1 eV reduction of IE after geometry relaxation. According to Ref. 18, B3LYP calculations significantly underestimate the IE, by more than 0.3 eV. The B3LYP relaxation energy of 0.07 eV is however in quite good agreement with respect to the QMC estimates. The residual energy difference between the QMC and experimental values can be related to the additional zero point energy (ZPE) effect, which leads to a weak correction of 16 meV, according to B3LYP/pVTZ estimates.[16]

TABLE IV. Carbon-carbon bond lengths for the anthracene molecule in the ground and exited states, considered in this paper.

Bond label	Ground state (Å)	Cation (Å)	Anion (Å)	^1L_a state (Å)
1-2	1.3570(3)	1.3750(5)	1.3816(7)	1.3929(7)
2-3	1.4243(4)	1.4104(8)	1.4148(5)	1.4028(5)
3-4	1.3899(4)	1.3960(5)	1.3994(6)	1.4055(17)
7-8	1.4218(6)	1.4008(5)	1.3988(6)	1.3853(7)
3-12	1.4328(4)	1.4328(14)	1.4418(21)	1.4338(14)

TABLE V. Vertical and adiabatic electronic affinities (EAs) for the anthracene molecule by VMC and LRDMC. In the last two row, we report benchmark calculations carried out in Ref. 17 with a focal point analysis (FPA). FPA are "exact" only for the electronic part, as they do not include zero point energy effects.

EA	VMC (eV)	LRDMC (eV)
Vertical QMC	0.139(18)	0.260(22)
Adiabatic QMC	0.260(14)	0.352(22)
Expt.	0.53 eV[12]	
Vertical "exact" (FPA)	0.279 eV[17]	
Adiabatic "exact" (FPA)	0.382 eV[17]	

C. Electron affinity

The electron affinity in the anthracene molecule is extremely sensitive to correlation. Indeed, the EA computed at the HF level is negative, namely, the electron attachment is not stable. On the other hand, the experimental EA is positive (0.53 eV[12]); therefore, the stabilization energy mainly comes from electron correlation. As a consequence, the *ab initio* EA estimate strongly depends on the level of theory and on the ability of the method to treat correlation effects. This has been pointed out in Ref. 17, which provides one of the most accurate vertical EAs (reported in Table V), through a FPA-based extrapolation in both basis set and theory. Anthracene is also the smallest PAH to show a stable electron attachment, being the measured EA of benzene (single ring) and naphthalene (two rings) both negative. For longer PAH, the EA is always positive and increases with the PAH length. Thus, the quite weak EA in anthracene results from a subtle combination of correlation and electron delocalization along the three ring chains, which places the molecule close to a critical behavior for electron addition processes.

Our QMC results are reported in Table V. The VMC underestimates both vertical and adiabatic EAs by 0.1 eV. The LRDMC improves upon the VMC values and brings them in agreement with the vertical and adiabatic EA benchmarks. The remaining difference between the adiabatic LRDMC results and the experiment can be accounted for by a ZPE effect. Indeed, B3LYP calculations of the zero point motion in the anion and ground state give a ZPE difference of 0.144 eV, which is one order of magnitude larger than the calculated ZPE difference between the cation and the ground state. This could be related to a significant softening of the carbon-carbon bond in the anion. Fig. 3 shows that the BLA is milder in the anion than in the ground state, with the same behavior as in the cation. However, in the anion, the carbon-carbon bond length is always larger than that in the cation, which signals that the bond is globally softened by the electron attachment. The QMC bond lengths are reported in Table IV.

By adding the zero point energy contribution to the LRDMC best estimate, one gets a total adiabatic EA of 0.496(22) eV, in agreement with the experimental value of 0.530 eV.

D. ^1L_a excited state

Linear PAHs present spectral properties with similar features across the whole family. They have been intensively studied since the pioneering work by Clar, Klevens, Platt, Jones, and others, in the late 1940s.[61–64] The lowest lying singlet excitations are characterized by two transitions denoted by Platt as ^1L_a and ^1L_b. The ^1L_a band is intense and polarized transversally with respect to the molecular longest symmetry axis, whereas the ^1L_b band is very weak and polarized longitudinally. ^1L_b is lower than ^1L_a in benzene and naphthalene, while they switch their order with increasing ring number. The ^1L_a becomes the singlet lowest energy excitation starting from anthracene and is one of the most challenging excitations in organic chemistry and photo-chemistry. It is a singlet HOMO-LUMO transition, which excites the ^1A_g ground state to a $^1B^+_{2u}$ (^1L_a) state in the D_{2h} point group (Platt's) notations.[64] It is a low-lying $\pi \to \pi^\star$ excitation, which involves a promotion of one electron from a π bonding to a $\pi^\star$ antibonding orbital. This singlet-singlet transition can be described by our JAGP, replacing one HOMO b_{3g} orbital (plotted in Fig. 4(a)) by a LUMO b_{1u} single-particle state (see Fig. 4(b)) in the AGP ^1A_g expansion, in such a way that the resulting $^1B^+_{2u}$ AGP is a symmetric function under the $\mathbf{r}^\uparrow \leftrightarrow \mathbf{r}^\downarrow$ swap. The latter condition implies a singlet state, as the antisymmetry under particle exchange must then come from the spin sector of the geminal. Therefore, the AGP for the $^1B^+_{2u}$ state reads

$$\begin{aligned}\Phi^{^1B_{2u}}(\mathbf{r}^\uparrow,\mathbf{r}^\downarrow) = &\sum_{k=1}^{\mathrm{HOMO}-1} \lambda_k \phi_k(\mathbf{r}^\uparrow)\phi_k(\mathbf{r}^\downarrow)\\ &+ \epsilon\phi_{\mathrm{HOMO}}(\mathbf{r}^\uparrow)\phi_{\mathrm{LUMO}}(\mathbf{r}^\downarrow)\\ &+ \epsilon\phi_{\mathrm{LUMO}}(\mathbf{r}^\uparrow)\phi_{\mathrm{HOMO}}(\mathbf{r}^\downarrow),\end{aligned} \tag{14}$$

where $\{\lambda_k = 1\}_{k=1,\ldots,\mathrm{HOMO}-1}$, and $\epsilon \ll 1$ is an arbitrary small prefactor, which makes the weights of double excited states with $(\mathrm{LUMO})^2$ occupations vanishing, i.e., of order ϵ^2, negligible with respect to the dominant $O(\epsilon)$ contributions. This construction is necessary, as $\det(\Phi(\mathbf{r}_i^\uparrow,\mathbf{r}_j^\downarrow))$ takes into account all possible N-particle states made of $N/2$ pairs included in the expansion Φ, with relative weights given by the product of λ_k, where k runs over the $N/2$ pairs of each allowed N-particle state. Double excitations are therefore weighted as ϵ^2 in $\det(\Phi)$, while at first (leading) order in ϵ, $\det(\Phi)$ is represented by two Slater determinants each one made by the HOMO → LUMO

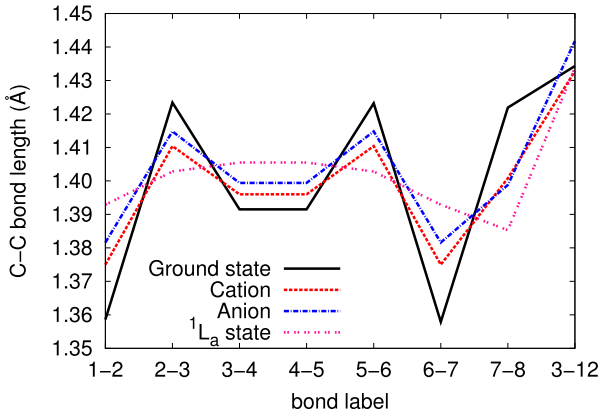

FIG. 3. QMC geometries for the positive ion, the negative one, and the ^1L_a state compared to the QMC neutral ground state.

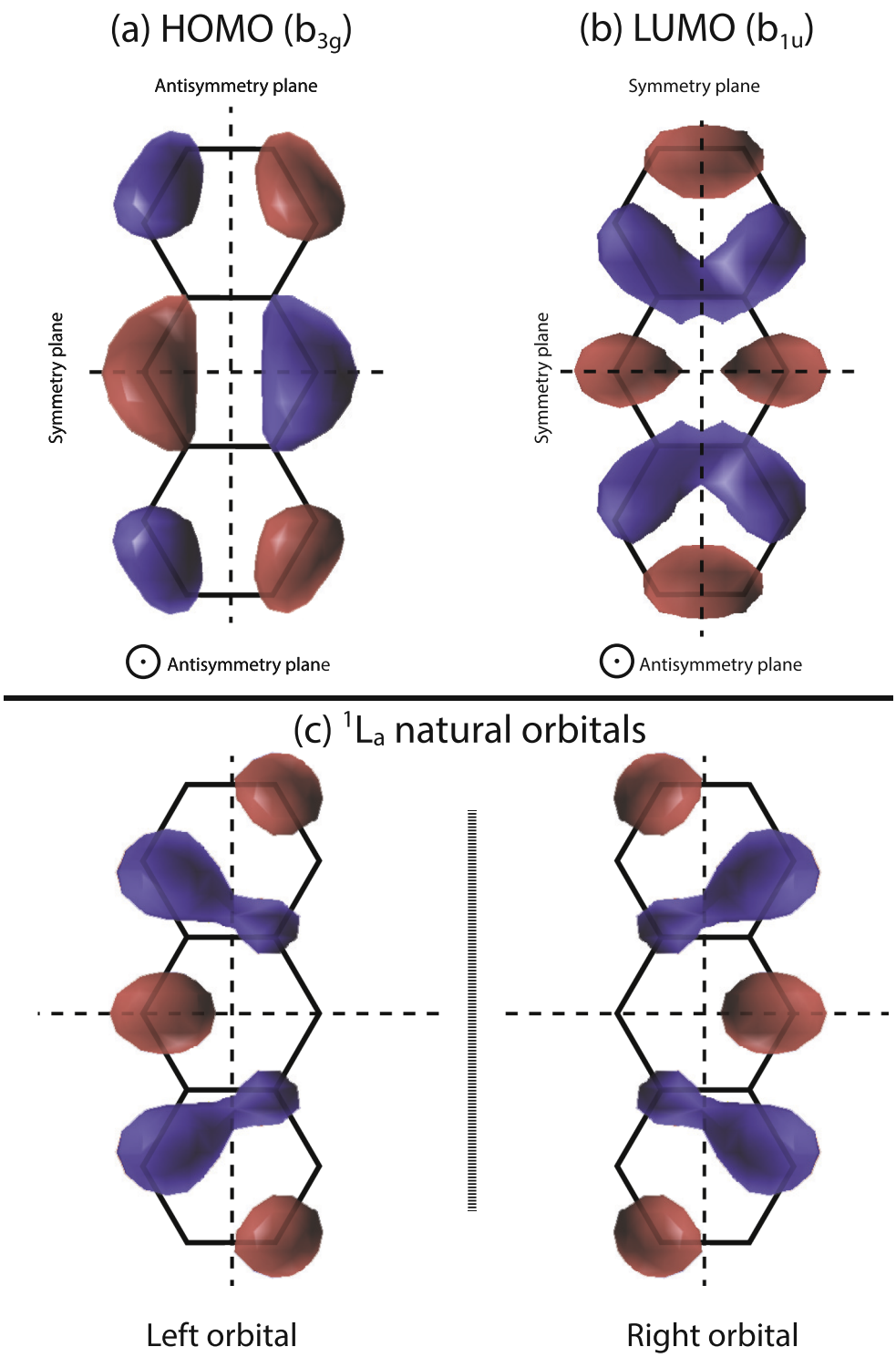

FIG. 4. (a) HOMO orbital for the anthracene molecule calculated in DFT LDA for the ground state; (b) LUMO DFT orbital for the anthracene molecule calculated in DFT LDA for the ground state; (c) molecular natural orbitals obtained from the $^1B_{2u}$ AGP diagonalization process encoding the HOMO-LUMO transition, as described in the text.

transition in the spin-down and spin-up sectors, respectively. ϵ has been further relaxed by energy minimization as additional variational parameter, and its optimal value is about 10^{-3}. From this analysis, it is thus clear that the $^1B^+_{2u}$ state as described by the AGP ansatz is multideterminant, with (mainly) two determinants equally weighted, and all single determinant methods fail in computing this excitation.

However, this is not the only difficulty in dealing with the $^1B^+_{2u}$ state. Indeed, associated to its multideterminant character, there is a strong charge transfer along the short symmetry axis (the y axis in our geometry notations), which leads to a large polarization along the same direction. To understand this, let us diagonalize the AGP in Eq. (14). We get the equivalent AGP representation, expressed in terms of "natural orbitals,"

$$\Phi^{^1B_{2u}}(\mathbf{r}^\uparrow,\mathbf{r}^\downarrow) = \sum_{k=1}^{\mathrm{HOMO}-1} \lambda_k \phi_k(\mathbf{r}^\uparrow)\phi_k(\mathbf{r}^\downarrow) + \epsilon\phi_{\mathrm{left}}(\mathbf{r}^\uparrow)\phi_{\mathrm{left}}(\mathbf{r}^\downarrow) - \epsilon\phi_{\mathrm{right}}(\mathbf{r}^\uparrow)\phi_{\mathrm{right}}(\mathbf{r}^\downarrow), \quad (15)$$

where $\phi_{\mathrm{left}} = (\phi_{\mathrm{HOMO}} + \phi_{\mathrm{LUMO}})/\sqrt{2}$ and $\phi_{\mathrm{right}} = (\phi_{\mathrm{HOMO}} - \phi_{\mathrm{LUMO}})/\sqrt{2}$ are the "left" and "right" MOs depicted in Fig. 4(c). The "left" ("right") MO breaks the reflection symmetry with respect to the x axis, and it has lobes more localized to the left (right) side of the molecule. Those are prototypical edge states. Equation (15) makes the physical interpretation transparent. The $^1B_{2u}^+$ AGP, in its natural orbital representation, is a resonance between two states, where a singlet electron pair is localized on one of the two edges of the molecule. The pair localization is related to a strong charge transfer from the ground state arrangement. The point group symmetry is restored by the linear combination of the "left" and "right" states, with a net large polarization of the resulting multi-reference wave function. Early analysis in the VB framework soon highlighted the ionic character of the $^1B_{2u}^+$ state. However, the underlying $N-2$ electrons are not simple spectators, but they react to the charge transfer taking place at the HOMO-LUMO level of energy. Indeed, the remaining electrons contribute to screen this charge fluctuation, a process which is dubbed as "σ-π correlation" in organic chemistry. Accurately describing this kind of correlation is a challenge for every quantum chemistry method.

The results of our JAGP QMC calculations for the ^1L_a energy excitations are reported in Table VI. The geometry optimization performed at the VMC level gives relaxation energies of 300(24) meV and 553(24) meV, in CBS and UBS, respectively. These values are systematically reduced by further LRDMC calculations—performed at frozen optimal VMC geometry—which give 216(38) meV and 340(50) meV, in CBS and UBS. Despite the reduction, these values are still slightly larger then previous TDDFT-B3LYP estimates, which give a relaxation energy of 0.22 eV.[10]

UV spectral data are available for this transition. The early value of 3.38 eV,[13] used as benchmark in many subsequent theoretical calculations, was obtained in solution after applying a solvent correction. A more recent experiment[14] studied the ^1L_a energy spectrum directly in the gas phase, by means of ultrahigh-resolution spectroscopy. This gives a value of 3.43 eV, which is thus the best experimental reference up to now. However, as shown in Table VI, our adiabatic excitation energies of 3.778(16) eV and 3.754(27) eV, at VMC-CBS and LRDMC-CBS levels, respectively, are not quite consistent with experiment.

A possible bias can come from the CBS, as polarization effects usually require a very flexible basis. To check this hypothesis, we performed further QMC calculations in a Gaussian primitive basis set (UBS) of $(9s8p4d)$ for carbon and $(8s4p)$ for hydrogen. These results are reported in Table VI as well. Optimizing both the electronic and ionic parts in the UBS lowers the energy of the excited state, and yields a VMC value of 3.640(16) eV, which is closer to experiment. LRDMC calculations carried out at the frozen optimal VMC geometry in the UBS confirm this excitation energy, but with a twice larger statistical error bar.

The remaining discrepancy between VMC and experiment can be due to vibronic effects, i.e., the ZPE. Indeed, phonon calculations performed in Ref. 14 at the CASSCF(4,4) level of theory give a zero-point energy of the ground state larger than the $^1B_{2u}^+$ one by 0.151 eV. This difference would push the 0_0^0 spectral band to smaller values than the wall-to-wall estimates, and would explain the remaining mismatch found in VMC. In fact, the VMC-UBS value corrected by the ZPE yields a 0_0^0 spectral band located at 3.489(17) eV, 3.5 sigma further from the experimental value. Therefore, our VMC value of 3.640(16) eV can be taken as the newest most accurate theoretical benchmark for the wall-to-wall excitation.

Interestingly enough, if one adds the vibronic (0.151 eV) and geometry relaxation (≈0.3 eV) contributions to the most recently measured 0_0^0 excitation (3.43 eV), one finds a vertical excitation energy of 3.88 eV. This is significantly larger than the commonly used reference (3.60 eV) for the vertical ^1L_a transition. The correct vertical excitation reference falls in between previous CIS(D)[18] and CC2[10] calculations and partially resumes the reliability of the high level CIS(D) results[65] for this particular case. In this perspective, the TDDFT results obtained with the B3LYP functional[10,18] look even worse than before. As already pointed out, the TDDFT failure for this excitation is severe, probably due to the adiabatic approximation in the time dependent functional which is particularly rough in case of strong charge fluctuation and polarization. The use of constricted variational DFT seems to improve over the TDDFT results already at the LDA level.[19]

The recent ultrahigh-resolution spectroscopy allows us to compare with experiment not only the transition energies but also the structural changes across the excitation. Indeed, from the analysis of the spectral fine structure, one can get precious information on the rotovibrational spectrum in the initial and final states, and from that one can measure the rotational constants and moments of inertia along the main rotational axes of the molecule. Those are global probes, as they depend on the sum over all ionic positions in the molecular structure. Nevertheless, they reveal the major structural changes, when an electron is excited, added, or removed from the molecular ground state.

TABLE VI. Vertical and adiabatic ^1L_a excitation energies for the anthracene molecule by VMC and LRDMC, with both contracted (CBS) and uncontracted basis set (UBS) optimization. The best QMC results are obtained for the adiabatic calculations in the UBS. We report also recent calculations and experimental results extrapolated to the gas phase,[13] or directly performed in the gas phase.[14]

^1L_a excitation energy	VMC (eV)	LRDMC (eV)
Vertical (CBS) QMC	4.078(17)	3.97(3)
Vertical (UBS) QMC	4.193(17)	4.00(4)
Adiabatic (CBS) QMC	3.778(17)	3.754(27)
Adiabatic (UBS) QMC	3.640(17)	3.66(3)
Expt.	3.38 eV,[13] 3.43 eV[14]	
Vertical TDDFT(B3LYP)	3.21 eV[10,18]	
Vertical CVDFT	3.68 eV[19]	
Vertical CC2	3.69 eV[10]	
Vertical CIS(D)	4.05 eV[18]	

TABLE VII. Relative difference in rotational constants across the ^1L_a excitation of the anthracene molecule as measured by ultrahigh-resolution spectroscopy and computed from the VMC relaxed geometries. A, B, and C are the GS rotational constants related to moments of inertia axes oriented along the x, y, and z directions, respectively, where the Cartesian frame is taken according to the conventions adopted to draw the (x, y) coordinates in Fig. 1, with the z axis pointing outwards. We present results for both contracted (CBS) and uncontracted (UBS) basis set to underline the importance of having an accurate wave function to obtain a good description of structural changes.

Rotational constants variation $\Delta\alpha/\alpha = \frac{(\alpha(^1L_a)-\alpha(\mathrm{GS}))}{\alpha(\mathrm{GS})}$	Expt.[14]	VMC (CBS)	VMC (UBS)
$\Delta A/A$	0.0108	0.0064(4)	0.0112(5)
$\Delta B/B$	−0.0136	−0.0116(10)	−0.0132(2)
$\Delta C/C$	−0.0099	−0.0085(9)	−0.0091(2)

Table VII shows the rotational constants relative difference between the ^1L_a and ^1A_g states. Both experimental and VMC values are reported. The VMC rotational constants are computed as $\alpha = \hbar/4\pi c I_\alpha$, where I_α is the moment of inertia along one of the three main rotational axes of the molecule. I_α is evaluated according to the optimal VMC structure, obtained in UBS, and with H and C masses based on their natural isotope partitioning. The relative variation of the rotational constants is well reproduced by QMC, signaling that the global structural deformation the molecule undergoes through the electronic excitation is correctly described by the VMC geometry relaxation.

It is interesting to see that the ^1L_a structural change is indeed significant. In Fig. 3, the bond alternation clearly present in the ^1A_g GS is gone in the ^1L_a excited state. This is another consequence of the strong ionic character of the excitation, which destroys the VB resonance structure, mainly covalent in nature, of the GS. As we have discussed, adiabatic and vibronic effects are important to quantitatively reproduce the experimental 0_0^0 band of the UV spectrum. The reason appears now clear as both the equilibrium positions and their vibrations are related to the very different nature of the two electronic states.

E. Zig-zag versus armchair edge: A comparison with phenanthrene

Beside the oligoacenes, which are characterized by a zig-zag edge, another important class of PAH is the one of phenacenes, based on the fusion found in phenanthrene (see Fig. 5), namely, leading to an armchair border. For each oligoacene member, there is a corresponding one in the phenacenes, having the same number of rings, but different arrangements. The edge shape has a dramatic impact on the electronic properties of the PAH. The oligoacenes have a tendency toward metallicity in the infinite chain limit, with low-lying excitations,[23,24] such as the singlet-triplet, whose gap shrinks as the PAH length increases,[21] and the possible formation of an open-shell singlet ground state.[22] On the other hand, the phenacenes spectra show a gap which is weakly dependent on the molecular size. Similar features have been found in graphene nanoribbons. They share the same structures as oligoacenes and phenacenes except for their width, which is usually much larger than one ring unit. The armchair graphene nanoribbons are semiconductors, while theoretical *ab initio* predictions for the zig-zag nanoribbons yield a small gap insulator, spin polarized at the edges, which becomes a half-metal in case a strong in-plane electric field is applied, leading to very peculiar spintronic properties.[25]

The isomeric analog of anthracene is the phenanthrene molecule. Studying the differences between anthracene and phenanthrene, and how they are captured by *ab initio* methods, is fundamental to assessing the accuracy of those methods in these simplest chemical systems with the aim at validating their use in larger PAHs and in graphene nanoribbons. In this section, we report the ground state study of phenanthrene carried out by VMC and LRDMC with the same procedure (basis set, PBE starting geometry, and wave function) as the one used for anthracene. Both the wave function and geometry have been fully relaxed in the VMC framework, giving the ground state equilibrium geometry shown in Fig. 6 and Table VIII. From the figure, it is apparent that the BLA is no longer present in phenanthrene, as the molecule is less symmetric than anthracene. Nevertheless, there is still a strong carbon-carbon bond length variation, ranging from 1.3510 Å to 1.4458 Å, roughly the same bond stretching found in anthracene. The comparison with B3LYP/pVTZ and PBE/pVTZ geometries reveals that PBE is poorer than B3LYP, and the latter is almost on the top of the QMC geometries, a systematic behavior already found for anthracene in Sec. III A.

The reference experimental geometries[66] are obtained from X-ray crystallographic data. We would like to stress once

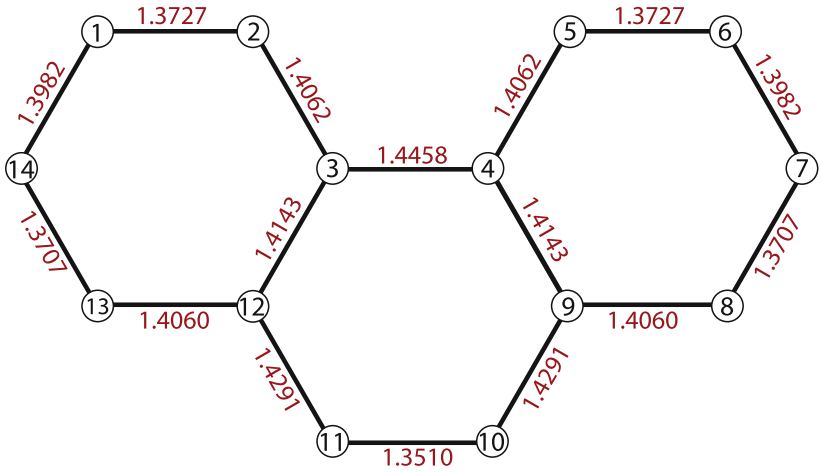

FIG. 5. Carbon site labels and QMC bond lengths in Å for the phenanthrene ground state.

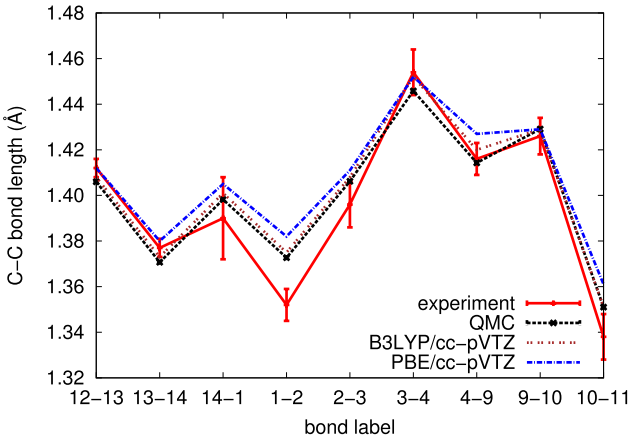

FIG. 6. Phenanthrene ground state experimental (Ref. 66), QMC, B3LYP/pVTZ, and PBE/pVTZ bond lengths. All theoretical values are obtained in the present work. The bond label conventions are reported in Fig. 5. The error bars are plotted for both experimental results and QMC geometries.

again that the equilibrium geometry measured in the crystal could differ from the gas phase one, due to a possible bias provided by the crystalline environment.

A key property which depends on the edge shape is the relative thermochemical stability of the PAH species. It is well known that the oligoacenes are very reactive due to their multi-radical character.[21] This is reflected by the chemical instability of zig-zag mono-hydrogenated and unreconstructed graphene nanoribbons, which prevents them from being synthesized and measured. The armchair nanoribbons are instead stable and widely analyzed experimentally. Recent PBE-DFT based *ab initio* calculations showed that indeed zig-zag graphene nanoribbons are thermally unstable over a very large range of experimental conditions.[67] A related situation happens for the phenanthrene and anthracene molecules. The former shows a lower enthalpy of formation than the latter, according to experimental Ref. 15. Therefore, the armchair edge has a larger thermochemical stability than the zig-zag one. Those data can be directly compared to the total energy differences between anthracene and phenanthrene computed by QMC at their corresponding VMC equilibrium geometries. In this case, the total energy difference does not need to be corrected by the zero-point energies, which are basically the same in the isomeric compounds.[68] The VMC and LRDMC energy differences are reported in Table IX. They both agree with experiment within chemical accuracy, as they lay within one error bar (<1 kcal/mol) from the experimental value. CCSD(T) calculations in the pVTZ basis set performed at the QMC geometries with the Molpro implementation are in very good agreement with the QMC energy differences. B3LYP and PBE energy differences, also reported in Table IX and computed at their corresponding relaxed geometries (see Tables II and VIII), slightly underestimate the stability of phenanthrene with respect to anthracene, a behavior which is more apparent in PBE.

TABLE VIII. Carbon-carbon bond lengths for the phenanthrene molecule. For comparison, we report also the B3LYP/pVTZ and PBE/pVTZ results for the same carbon pseudopotential. The bond label conventions are drawn in Fig. 5.

Bond label	Expt.[66] (Å)	QMC (Å)	B3LYP/pVTZ (Å)	PBE/pVTZ (Å)
12-13	1.412(4)	1.4060(5)	1.408	1.412
13-14	1.377(4)	1.3707(5)	1.373	1.380
14-1	1.390(18)	1.3982(5)	1.401	1.405
1-2	1.352(7)	1.3727(5)	1.375	1.382
2-3	1.396(10)	1.4062(5)	1.408	1.411
3-4	1.454(10)	1.4458(5)	1.452	1.452
4-9	1.416(7)	1.4143(5)	1.420	1.427
9-10	1.426(8)	1.4291(5)	1.429	1.429
10-11	1.338(10)	1.3510(5)	1.351	1.361

TABLE IX. Energy differences between the anthracene and phenanthrene molecules, obtained in this work with various methods, and compared to the experiment. This shows the stability of the armchair edge with respect to the zig-zag one in the three-ring PAHs.

	$E_{\text{anthra}} - E_{\text{phena}}$ (kJ/mol)
PBE/pVTZ	18.6
B3LYP/pVTZ	20.8
CCSD(T)/pVTZ	25.0
VMC	26.8(1.8)
LRDMC	24.4(3.0)
Expt.	24.7(2.9)[15]

IV. CONCLUSIONS

In this paper, we have shown that the variational QMC framework based on the JAGP is capable to accurately reproduce not only the ground state properties but also the excited-state energies and geometries.

We described an algorithm, already used in the ground state molecular orbital energy minimization, which has proven to be stable also in the excited state electronic optimization, and in the full geometry relaxation at a given excited state configuration. It is based on a constrained (fixed-rank) minimization of the geminal expansion, followed by an eigenvalue selection of the determinant density matrix spectrum. In practice, after the constrained variation of the geminal matrix λ, one keeps the target number of natural orbitals by selecting those with largest eigenvalues. This method gives new molecular orbitals which largely lay in the same symmetry sector of the wave function as in the former iteration. The statistics and parameters acceleration which guarantee this stable behavior do not differ from those of a regular ground state optimization.

We have applied this scheme to study the anthracene addition and removal spectrum, and the difficult case of the $\pi \rightarrow \pi^\star$ neutral excitation. Anthracene is an ideal test case, as it is one of the simplest PAH, which retains all the complexity of larger or more general aromatic hydrocarbons. Moreover, the PAHs represent a very interesting family to study, as their spectral properties are appealing and nowadays exploited in many photo-chemistry applications and devices.

The anthracene ground state energy and geometry turn out to be well described by our QMC framework. It correctly captures the resonance between singlets responsible of the chemical bond stabilization in the aromatic rings, because the bond alternation is well reproduced. Moreover, the energy

difference with phenanthrene, which shares the same number of rings but in an armchair arrangement, is in agreement with thermochemistry data.

In anthracene, we proved the importance of performing geometry relaxation to get excitation energies in better agreement with the experiment. Indeed, the difference between the vertical and adiabatic excitations can be sizable, and strongly dependent on the type of transition, leading to corrections of up to 0.2-0.3 eV. Relaxing the geometry in the excited state electronic configuration is, therefore, very important. We showed that the present QMC approach is able to carry the double task of dealing with non-trivial excited states and optimizing their geometry, both performed within the same JAGP framework. Good agreement with experiment is finally obtained by the QMC adiabatic excitation energies. The remaining gap between QMC and experimental values can be related to the zero point energy variation, that we have not addressed in this paper. However, the possibility to efficiently evaluate ionic forces also in the excited states by QMC allows one to reliably compute phonon properties as well, as already reported elsewhere for the electronic ground state of small molecules.[69]

The computational cost of our JAGP-QMC calculations for the three-ring PAH analyzed here is moderate. The VMC calculations took 4 h (wall time) on 64 cores (Intel Xeon-Curie supercomputer thin nodes) to get an error of 12 meV in the total energy, the geometry relaxations took about 20 wall time hours on 256 cores to converge, while the LRDMC cost (with a lattice space of 0.25 a.u.) is about 12 wall time hours on 512 cores to yield a 24 meV error in the fixed-node total energy. Given the QMC favorable scaling (N^3-N^4) with the system size, it is therefore possible to perform large scale calculations on larger PAH, at the same accuracy as the one obtained here.

To conclude, in this paper, we reported benchmark calculations of the ground state properties, ionization energy, electron affinity, and neutral singlet HOMO-LUMO excitations in anthracene. The interplay between electronic and structural modifications in the excited states is deeply explored within the same QMC framework. This opens the way to new applications in a large variety of chemical systems. Our scheme is promising also for further studies on larger PAH, where the tight competition between low-lying energy states can give rise to new peculiar structural and chemical properties of the hydrocarbon chain, such as the occurrence of an open-shell singlet ground state, and strong diradical or multiradical character.

ACKNOWLEDGMENTS

One of us (M.C.) acknowledges computational support provided by the GENCI Grant Number 096493. This research was also supported by COFIN-PRIN2010 MIUR.

[1]E. Coccia, D. Varsano, and L. Guidoni, J. Chem. Theory Comput. **10**, 501 (2014).
[2]C. Filippi, M. Zaccheddu, and F. Buda, J. Chem. Theory Comput. **5**, 2074 (2009).
[3]O. Valsson, P. Campomanes, I. Tavernelli, U. Rothlisberger, and C. Filippi, J. Chem. Theory Comput. **9**, 2441 (2013).
[4]B. E. Hardin, H. J. Snaith, and M. D. McGehee, Nat. Photonics **6**, 162 (2012).
[5]C.-L. Wang, Y.-C. Chang, C.-M. Lan, C.-F. Lo, E. Wei-Guang Diau, and C.-Y. Lin, Energy Environ. Sci. **4**, 1788 (2011).
[6]O. Valsson and C. Filippi, J. Chem. Theory Comput. **6**, 1275 (2010).
[7]I. Schapiro, M. N. Ryazantsev, L. M. Frutos, N. Ferré, R. Lindh, and M. Olivucci, J. Am. Chem. Soc. **133**, 3354 (2011).
[8]D. Bousquet, R. Fukuda, P. Maitarad, D. Jacquemin, I. Ciofini, C. Adamo, and M. Ehara, J. Chem. Theory Comput. **9**, 2368 (2013).
[9]R. Guareschi and C. Filippi, J. Chem. Theory Comput. **9**, 5513 (2013).
[10]S. Grimme and M. Parac, ChemPhysChem **4**, 292 (2003).
[11]P. M. Mayer, V. Blanchet, and C. Joblin, J. Chem. Phys. **134**, 244312 (2011).
[12]J. Schiedt and R. Weinkauf, Chem. Phys. Lett. **266**, 201 (1997).
[13]D. Biermann and W. Schmidt, J. Am. Chem. Soc. **102**, 3163 (1980).
[14]M. Baba, M. Saitoh, K. Taguma, K. Shinohara, K. Yoshida, Y. Semba, S. Kasahara, N. Nakayama, H. Goto, T. Ishimoto *et al.*, J. Chem. Phys. **130**, 134315 (2009).
[15]J. Pedley, *Thermochemical Data and Structures of Organic Compounds*, Trc Data Series (CRC Press, Taylor & Francis Group, 1994).
[16]M. S. Deleuze, L. Claes, E. S. Kryachko, and J.-P. Franois, J. Chem. Phys. **119**, 3106 (2003).
[17]B. Hajgat, M. S. Deleuze, D. J. Tozer, and F. De Proft, J. Chem. Phys. **129**, 084308 (2008).
[18]E. S. Kadantsev, M. Stott, and A. Rubio, J. Chem. Phys. **124**, 134901 (2006).
[19]M. Krykunov, S. Grimme, and T. Ziegler, J. Chem. Theory Comput. **8**, 4434 (2012).
[20]F. H. Allen, O. Kennard, D. G. Watson, L. Brammer, A. G. Orpen, and R. Taylor, J. Chem. Soc., Perkin Trans. 2 (12), S1 (1987).
[21]J. Hachmann, J. J. Dorando, M. Avils, and G. K.-L. Chan, J. Chem. Phys. **127**, 134309 (2007).
[22]M. Bendikov, H. M. Duong, K. Starkey, K. Houk, E. A. Carter, and F. Wudl, J. Am. Chem. Soc. **126**, 7416 (2004).
[23]H. Chakraborty and A. Shukla, J. Phys. Chem. A **117**, 14220 (2013).
[24]B. Hajgat, D. Szieberth, P. Geerlings, F. De Proft, and M. S. Deleuze, J. Chem. Phys. **131**, 224321 (2009).
[25]Y.-W. Son, M. L. Cohen, and S. G. Louie, Nature **444**, 347 (2006).
[26]M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007).
[27]M. Dolg and C. Filippi, private communication (2013).
[28]M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).
[29]M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).
[30]M. Marchi, S. Azadi, M. Casula, and S. Sorella, J. Chem. Phys. **131**, 154116 (2009).
[31]M. Dagrada, M. Casula, A. M. Saitta, S. Sorella, and F. Mauri, J. Chem. Theory Comput. **10**, 1980 (2014).
[32]L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry* (Cornell University Press, Ithaca, New York, 1939).
[33]P. W. Anderson, Science **235**, 1196 (1987).
[34]W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).
[35]C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996).
[36]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).
[37]P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, J. Phys.: Condens. Matter **21**, 395502 (2009).
[38]S. Sorella, people.sissa.it/sorella/web/index.html; accessed 21 May 2012.
[39]S. Sorella and L. Capriotti, Phys. Rev. B **61**, 2599 (2000).
[40]S. Sorella, Phys. Rev. B **64**, 024512 (2001).
[41]C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).
[42]S. Sorella, Phys. Rev. B **71**, 241103 (2005).
[43]S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).
[44]A. Zen, Y. Luo, S. Sorella, and L. Guidoni, J. Chem. Theory Comput. **9**, 4332 (2013).
[45]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2012.1, a package of *ab initio* programs, 2012, see http://www.molpro.net.
[46]F. Schautz and C. Filippi, J. Chem. Phys. **120**, 10931 (2004).
[47]P. M. Zimmerman, J. Toulouse, Z. Zhang, C. B. Musgrave, and C. J. Umrigar, J. Chem. Phys. **131**, 124103 (2009).
[48]F. Schautz, F. Buda, and C. Filippi, J. Chem. Phys. **121**, 5836 (2004).
[49]A. Zen, Y. Luo, G. Mazzola, L. Guidoni, and S. Sorella, J. Chem. Phys. **142**, 144111 (2015).
[50]C. Filippi and C. J. Umrigar, Phys. Rev. B **61**, R16291 (2000).
[51]R. Assaraf and M. Caffarel, Phys. Rev. Lett. **83**, 4682 (1999).
[52]R. Assaraf and M. Caffarel, J. Chem. Phys. **119**, 10536 (2003).
[53]C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501 (2008).
[54]S. Sorella and L. Capriotti, J. Chem. Phys. **133**, 234111 (2010).
[55]M. Barborini, S. Sorella, and L. Guidoni, J. Chem. Theory Comput. **8**, 1260 (2012).

[56]M. Casula, C. Filippi, and S. Sorella, Phys. Rev. Lett. **95**, 100201 (2005).
[57]T. Wassmann, A. P. Seitsonen, A. M. Saitta, M. Lazzeri, and F. Mauri, J. Am. Chem. Soc. **132**, 3440 (2010).
[58]J. M. Martin, J. El-Yazal, and J.-P. François, J. Phys. Chem. **100**, 15358 (1996).
[59]K. B. Wiberg, J. Org. Chem. **62**, 5720 (1997).
[60]J. M. L. Martin, J. El-Yazal, and J.-P. François, Mol. Phys. **86**, 1437 (1995).
[61]R. N. Jones, Chem. Rev. **32**, 1 (1943).
[62]E. Clar, Ber. Dtsch. Chem. Ges. (A and B Ser.) **69**, 607 (1936).
[63]H. Klevens and J. Platt, J. Chem. Phys. **17**, 470 (1949).
[64]J. R. Platt, J. Chem. Phys. **17**, 484 (1949).
[65]M. Head-Gordon, D. Maurice, and M. Oumi, Chem. Phys. Lett. **246**, 114 (1995).
[66]V. Petříček, I. Císařová, L. Hummel, J. Kroupa, and B. Březina, Acta Crystallogr., Sect. B: Struct. Sci. **46**, 830 (1990).
[67]T. Wassmann, A. Seitsonen, A. Saitta, M. Lazzeri, and F. Mauri, Phys. Rev. Lett. **101**, 096402 (2008).
[68]W. Herndon, Chem. Phys. Lett. **234**, 82 (1995).
[69]Y. Luo, A. Zen, and S. Sorella, J. Chem. Phys. **141**, 194112 (2014).