

Energy-Based Molecular Orbital Localization in a Specific Spatial Region

Tommaso Giovannini* and Henrik Koch*

Cite This: <https://dx.doi.org/10.1021/acs.jctc.0c00737>

Read Online

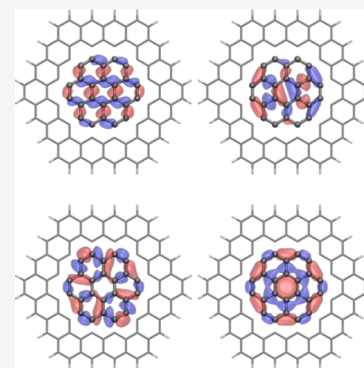
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: We present a novel energy-based localization procedure able to localize molecular orbitals into predefined spatial regions. The method is defined in a multiscale framework based on the multilevel Hartree–Fock approach. In particular, the system is partitioned into active and inactive fragments. The localized molecular orbitals are obtained maximizing the repulsion between the two fragments. The method is applied to several cases including both conjugated and non-conjugated systems. Our multiscale approach is compared with reference values for both ground-state properties, such as dipole moments, and local excitation energies. The proposed approach is useful to extend the application range of high-level electron correlation methods. In fact, the reduced number of molecular orbitals can lead to a large reduction in the computational cost of correlated calculations.



1. INTRODUCTION

Many processes in chemistry take place in a specific spatial region of a molecular system. To rationalize local phenomena, the concept of local occupied molecular orbitals (LMOs) is particularly useful in bridging chemical intuition and theoretical chemistry.¹ The LMOs are very convenient for describing electron correlation, as they can potentially reduce the computational cost of many-body methods.² Among the large variety of different localization procedures developed in the past,^{3–14} only a few are specifically designed to localize MOs into a specific spatial region of a molecular system.^{1,6,13–15} Among them, notable examples are given by the density matrix embedding theory^{16–22} and atomic valence active space.²³

However, such a localization procedure is important when dealing with phenomena taking place in a limited spatial location, for instance, local electronic excitations.^{15,24} This allows fragmentation of the target moiety in (at least) two different parts: the active, where the phenomenon takes place, and the inactive, that indirectly influences the active part. Such a partitioning defines the so-called focused models.²⁵ Most focused models are formulated in terms of quantum mechanical (QM)/classical approaches in which the active–inactive interaction is usually limited to electrostatics.^{26–30} In addition, several fragmentation approaches have been proposed in the last years, demonstrating their capability to treat large molecular systems.^{31–36} However, such approaches are commonly limited to the ground-state energy, although some of them have been extended to treat excitation energies,^{37–39} and the interaction between the monomers is usually treated at the electrostatic level.^{32,40} Such shortcomings can be overcome using quantum-

embedding approaches, such as frozen density embedding and projected-based methods.^{21–23,37,38,41–57}

In this paper, we are proposing a novel method which can provide localized occupied MOs in a specific spatial region of the system. The approach is based on multilevel Hartree–Fock (MLHF) theory,^{58,59} in which a molecular system is divided into two different fragments: A (active) and B (inactive). In particular, the density matrix of the whole system is partitioned into two fragment densities. Such partitioning is performed by selecting the number of electrons belonging to the active part, and consequently, the number of occupied orbitals is defined. This step is performed using a Cholesky decomposition of the initial density,^{58,60,61} and the active virtual orbitals are constructed by means of orthonormalized projected atomic orbitals (PAOs).^{62,63} The inactive density is then obtained by subtracting the active density from the total density. The inactive density is then kept frozen during the self-consistent field (SCF) cycles and enters the active Fock matrix as a one-electron term. This approach is similar to frozen density embedding methods;^{52–57} however, in MLHF the SCF procedure is solved in the active MO basis. Note that, in case the starting density is the full HF converged density, then the MLHF energy corresponds to the HF energy of the full system. The MLHF approach can therefore be defined as rigorous because all

Received: July 15, 2020

interactions between the two parts are retained at the HF level. If the decomposed density is not the full HF density, then the quality of the results depends on the chosen partitioning method, which however does not sacrifice the rigorosity of the approach.

As previously reported in ref 58, the MLHF method also differs from most projection-based approaches, either developed for density functional theory or HF frameworks. In fact, the starting point of projection-based approaches is usually a SCF calculation on the entire system, and the optimized MOs are then assigned to active and inactive parts using an a priori orbital assignment.^{38,42,46–51,64–66} The quantum embedding Hamiltonian is then constructed by including an exact or approximated embedding operator.^{42,46,49–51,66} However, we want to highlight that the selection of the active MOs is usually performed using a predefined threshold metric, which however may cause wrong MO selection, as reported by Kállay and co-workers.⁶⁵

In this paper, we provide a novel approach defined in the MLHF framework to localize the active MOs to a predefined fragment space. This is achieved by formulating a localization criterion entirely based on energetics. In particular, in our approach the repulsion between the two fragments is maximized in the space spanned by the occupied orbitals of the two fragments. In this way, the total energy of the system is unchanged. To the best of our knowledge, this is the first time that an energy-based variational method is used to localize MOs in a specific region of the molecule. Because the procedure is energy-guided, our approach differs conceptually from other MO localization procedures.^{1,3,4,23,38,47,48,67} Moreover, although the Cholesky decomposition of the starting density can be viewed as an approximation, it ensures the continuity of the potential energy surface (PES), which is not guaranteed for common localization procedures such as Boys.⁶⁷

In the following, we show that our novel energy-based localization procedure can provide MOs that are well localized on the specified active fragment, without carrying out an a priori orbital assignment. We apply our approach to ground-state properties and local excitation energies calculated at the coupled-cluster (CC2) level.⁶⁸ Within this scheme, CC2 ground- and excited-state calculations are performed using the MOs of the active fragment only; thus the intrinsic computational cost of high-level calculations is reduced, similarly to other multilevel methods.^{69,70} For the same reason, the accuracy of the computed local excitations depends crucially on the quality of the LMOs. In particular, the computational cost can be significantly reduced if the active MOs are correctly localized in the predefined spatial regions because a lower number of occupied orbitals is needed to achieve the desired level of accuracy. Also, we show that the method has a basis-set limit, which is guaranteed by the fact that the MO basis of both the active and the inactive fragments is defined in terms of the full atomic orbital (AO) basis.

The paper is organized as follows. In the next section, the MLHF theory is briefly summarized, and the energy-based orbital localization is discussed. Then, the computational procedure and the numerical applications are presented, with particular emphasis on the spread of the obtained localized MOs and on the accuracy of the novel approach in predicting local properties such as dipole moments and excitation energies. Summary and conclusions of the present work end the paper.

2. THEORY

Active–inactive partitioning in MLHF is realized by decomposing the density of the whole system into active and inactive densities ($\mathbf{D} = \mathbf{D}^A + \mathbf{D}^B$). The total HF energy can be written as

$$\begin{aligned} E &= \text{Tr } \mathbf{h}\mathbf{D} + \frac{1}{2}\text{Tr } \mathbf{D}\mathbf{G}(\mathbf{D}) + h_{\text{nuc}} \\ &= \text{Tr } \mathbf{h}\mathbf{D}^A + \frac{1}{2}\text{Tr } \mathbf{D}^A\mathbf{G}(\mathbf{D}^A) + \text{Tr } \mathbf{D}^A\mathbf{G}(\mathbf{D}^B) \\ &\quad + \text{Tr } \mathbf{h}\mathbf{D}^B + \frac{1}{2}\text{Tr } \mathbf{D}^B\mathbf{G}(\mathbf{D}^B) + h_{\text{nuc}} \end{aligned} \quad (1)$$

where \mathbf{h} and \mathbf{G} are the usual one- and two-electron matrices and h_{nuc} is the nuclear repulsion. The $\mathbf{G}(\mathbf{D})^X$ with $X = \{A, B\}$ matrix is defined as

$$G_{\mu\nu}(\mathbf{D}^X) = \sum_{\sigma\tau} D_{\sigma\tau}^X \left((\mu\nu|\sigma\tau) - \frac{1}{2}(\mu\tau|\sigma\nu) \right) \quad (2)$$

As can be seen from eq 1, the MLHF energy is formally equivalent to the HF energy if the decomposed density \mathbf{D} corresponds to the full HF density. However, in order to reduce the computational cost, in MLHF the density of fragment A is minimized in the field generated by the density B, which is kept fixed. This procedure is performed by minimizing the energy (see eq 1) in the MO basis of the active part, reducing the dimensionality of the problem. The MO basis is defined in terms of the full AO basis set, without performing any AO truncation, except linear dependencies. The quality of the MLHF results will therefore depend on both the accuracy of the decomposed density and the decomposition algorithm. In MLHF, the Fock matrix in AO basis is expressed by differentiating eq 1 with respect to \mathbf{D}^A

$$F_{\mu\nu} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{D}^A) + G_{\mu\nu}(\mathbf{D}^B) \quad (3)$$

In eq 3, the last term $G_{\mu\nu}(\mathbf{D}^B)$ is a one-electron contribution because the \mathbf{D}^B density is kept frozen during the SCF procedure.

As stated above, eq 1 is formally equal to the full HF energy when \mathbf{D} is the converged HF density for the entire system. However, eq 1 does not have an apparent physical interpretation because the different energy terms are not assigned to the individual fragments. Such a physical insight can be achieved by dividing the one-electron term into the kinetic (T) and the electron–nuclear attraction operators for the two parts (\mathbf{V}^A and \mathbf{V}^B). Thus, eq 1 can be written as

$$\begin{aligned} E &= \underbrace{\text{Tr } \mathbf{h}^A\mathbf{D}^A + \frac{1}{2}\text{Tr } \mathbf{D}^A\mathbf{G}(\mathbf{D}^A)}_{E_A} + h_{\text{nuc}}^A \\ &\quad + \underbrace{\text{Tr } \mathbf{h}^B\mathbf{D}^B + \frac{1}{2}\text{Tr } \mathbf{D}^B\mathbf{G}(\mathbf{D}^B)}_{E_B} + h_{\text{nuc}}^B \\ &\quad + \underbrace{\text{Tr } \mathbf{V}^B\mathbf{D}^A + \text{Tr } \mathbf{V}^A\mathbf{D}^B + \text{Tr } \mathbf{D}^A\mathbf{G}(\mathbf{D}^B)}_{E_{AB}} + h_{\text{nuc}}^{AB} \end{aligned} \quad (4)$$

where h_{nuc}^A , h_{nuc}^B , and h_{nuc}^{AB} are nuclear repulsion terms; E_A and E_B are the energies of the two fragments, whereas E_{AB} is the interaction energy. The latter term is composed of the electron–nuclear attraction between A and B and vice versa and the coulomb and exchange interactions between the two fragments.

Although eq 4 is equivalent to eq 1, it permits the definition of our localization procedure in a fragment-based model such as MLHF. In fact, an additional SCF procedure can be performed to optimize the energy of part A and/or B, in the occupied space of both fragments, that is, without changing the total energy. The

procedure is general and can be performed on the basis of any density matrix \mathbf{D} that is decomposed into two densities belonging to two fragments.

Two alternatives can be defined. First, the energy term E_A (see eq 4) can be minimized (denoted MLHF-A). In such a case, the Fock matrix reads

$$F_{\mu\nu} = h_{\mu\nu}^A + G_{\mu\nu}(\mathbf{D}^A) \quad (5)$$

In the second approach (called MLHF-AB), the sum of E_A and E_B is minimized. From the computational point of view, $E_A + E_B$ can be rewritten by considering that the total density $\mathbf{D} = \mathbf{D}^A + \mathbf{D}^B$ remains constant during occupied–occupied rotations. This means that \mathbf{D}^B can be expressed in terms of it as $\mathbf{D}^B = \mathbf{D} - \mathbf{D}^A$. Therefore, the sum of A and B energies reads

$$\begin{aligned} E_A + E_B &= \text{Tr} \mathbf{h}^A \mathbf{D}^A + \frac{1}{2} \text{Tr} \mathbf{D}^A \mathbf{G}(\mathbf{D}^A) + \text{Tr} \mathbf{h}^B \mathbf{D} \\ &\quad - \text{Tr} \mathbf{h}^B \mathbf{D}^A + \frac{1}{2} \text{Tr} \mathbf{D} \mathbf{G}(\mathbf{D}) \\ &\quad + \frac{1}{2} \text{Tr} \mathbf{D}^A \mathbf{G}(\mathbf{D}^A) - \text{Tr} \mathbf{D}^A \mathbf{G}(\mathbf{D}) \\ &= \text{Tr}(\mathbf{V}^A - \mathbf{V}^B) \mathbf{D}^A + \text{Tr} \mathbf{D}^A \mathbf{G}(\mathbf{D}^A) \\ &\quad - \text{Tr} \mathbf{D}^A \mathbf{G}(\mathbf{D}) + \text{Tr} \mathbf{h}^B \mathbf{D} + \frac{1}{2} \text{Tr} \mathbf{D} \mathbf{G}(\mathbf{D}) \end{aligned} \quad (6)$$

where the last two terms depend on the total density only and are therefore constant energy terms. The first three terms are instead similar to the MLHF energy contributions (see eq 1) because they are characterized by one-electron and two-electron contributions involving the active density only. The third term $\text{Tr} \mathbf{D}^A \mathbf{G}(\mathbf{D})$ is instead the two-electron interaction between the active and the constant total density \mathbf{D} . The Fock matrix of the active part can be written as

$$\begin{aligned} F_{\mu\nu} &= V_{\mu\nu}^A - V_{\mu\nu}^B + G_{\mu\nu}(\mathbf{D}^A) - G_{\mu\nu}(\mathbf{D}^B) \\ &= V_{\mu\nu}^A - V_{\mu\nu}^B + 2G_{\mu\nu}(\mathbf{D}^A) - G_{\mu\nu}(\mathbf{D}) \end{aligned} \quad (7)$$

which is again in the same form as eq 3 because it is characterized by one-electron contributions ($V_{\mu\nu}^A - V_{\mu\nu}^B$), a two-electron contribution on the active density $2G_{\mu\nu}(\mathbf{D}^A)$, and a constant contribution due to the total density $G_{\mu\nu}(\mathbf{D})$.

Notice that minimizing the sum of A and B parts in MLHF-AB (see eq 6) is equivalent to maximizing the interaction energy E_{AB} . Physically, this means that the repulsion between the two parts is maximized, and the occupied orbitals obtained by this scheme are those maximally located in the two fragments. For this reason, MLHF-AB can be viewed as an extension to fragment-based methods of the Edmiston–Ruedenberg MO localization procedure;³ although our approach conceptually differs from the latter because it is energy-guided. Finally, notice that when \mathbf{D} is the HF converged density of the entire system, the MLHF-AB localization will provide MOs which are localized in the predefined fragments. In this case, the localization only depends on the number of electrons which are assigned to each fragment, but it is independent of the chosen decomposition algorithm.

3. COMPUTATIONAL PROCEDURE

The two approaches are implemented in a development version of the electronic structure program *e^T71* and follow the computation protocol graphically depicted in Figure 1:

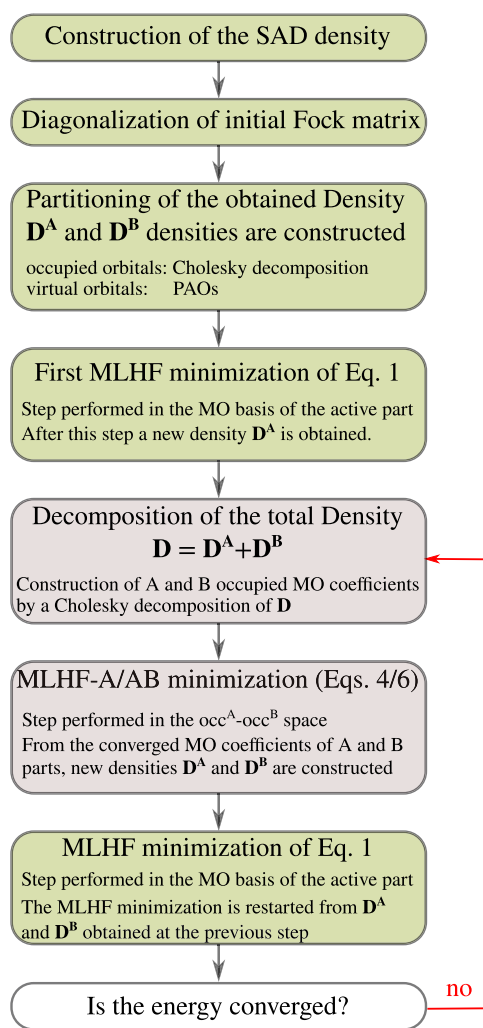


Figure 1. Graphical view of the computational procedure.

1. Construction of the initial density by means of superposition of atomic densities,⁷² followed by diagonalization of the initial Fock matrix.
2. Partitioning of the resulting density into A and B densities, using Cholesky decomposition for the active occupied orbitals and orthonormalized PAOs for active virtual orbitals.^{10,60,61,63,73,74} We note that the Cholesky decomposition of the total density \mathbf{D} into \mathbf{D}^A and \mathbf{D}^B is a mathematical method to decompose a matrix, which is unique if the same pivots are used. In this work, the Cholesky decomposition is performed by selecting the diagonals corresponding to the basis functions which are centered on the active atoms.^{58,61} In particular, \mathbf{D}^A is written in the AO basis $\{\alpha, \beta\}$ as⁷⁵

$$\mathbf{D}_{\alpha\beta}^A = \sum_{Ij} \mathbf{D}_{\alpha I} \tilde{\mathbf{D}}_{Ij}^{-1} \mathbf{D}_{j\beta} = \sum_I L_{\alpha I} L_{\beta I} \quad (8)$$

where I and J are the diagonal elements which are decomposed, $\tilde{\mathbf{D}}$ is the submatrix of \mathbf{D} containing the selected diagonal elements, and $L_{\alpha I}$ are the Cholesky orbitals. In particular, the number of \mathbf{D} diagonal elements which are selected corresponds to the correct number of occupied orbitals (n_o) of the active fragment, that is the largest n_o diagonals. As a result of the decomposition, the active Cholesky MOs are obtained and the active density

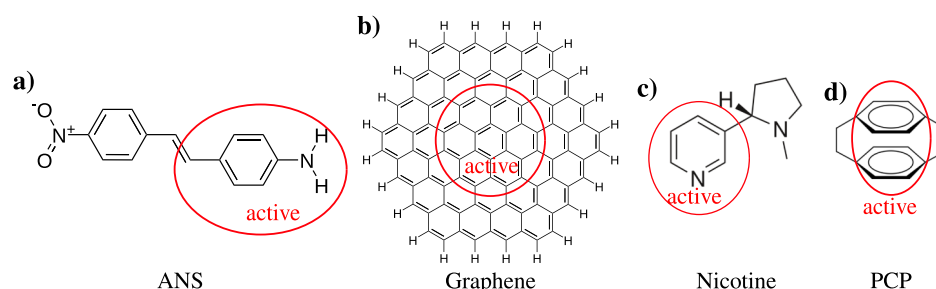


Figure 2. Molecular structures of ANS (a), graphene (b), nicotine (c), and PCP (d). The active parts used in MLHF calculations are highlighted.

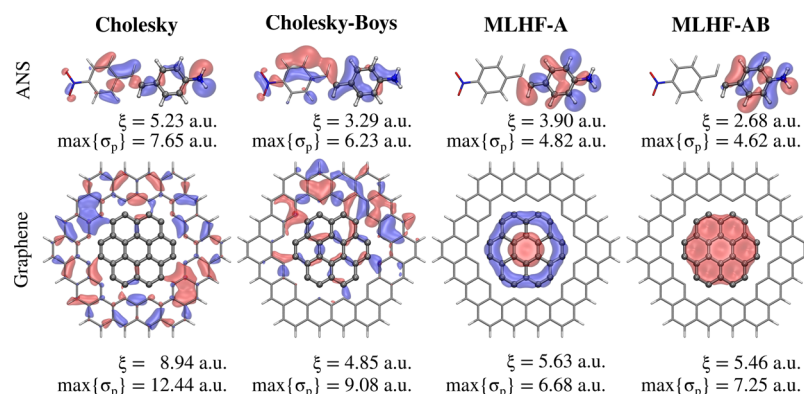


Figure 3. Graphical depiction of the most delocalized MOs of ANS (top) and graphene (bottom) calculated by MLHF (Cholesky), Cholesky–Boys, MLHF-A, and MLHF-AB methods. Computed ξ and maximum MO spread for all methods are also given. ANS is described with aug-cc-pVDZ basis set. The active part of the graphene sheet is described using cc-pVTZ basis set, whereas its inactive part with cc-pVDZ basis set.

- matrix \mathbf{D}^A is trivially constructed (see eq 8). As stated above, the active virtual orbitals are constructed by means of PAOs obtained from the AOs centered in the active fragment. The active and inactive occupied orbitals are projected out from them. The obtained PAOs are defined in terms of the full AO basis set and are orthonormalized through the Lowdin procedure. The threshold for removing the linear dependencies is set to 10^{-6} . The inactive density \mathbf{D}^B is instead obtained as a difference between the total density \mathbf{D} and the active one \mathbf{D}^A .
- The energy defined in eq 1 is minimized in the MO basis of the active part. In this way, the dimensionality of the problem is reduced, because, although the active MO coefficients are defined in the whole AO basis set, their number corresponds to the selected active orbitals only. Therefore, all AO matrices defined in eq 1 can be transformed in the active MO basis using the active MO coefficients.
 - The total density \mathbf{D} is constructed by summing the MLHF converged density \mathbf{D}^A and the inactive density \mathbf{D}^B , and active/inactive occupied orbitals are obtained by a Cholesky decomposition. Again, the total density \mathbf{D} is Cholesky decomposed by considering the diagonals belonging to the active atoms. The inactive density $\mathbf{D}^B = \mathbf{D} - \mathbf{D}^A$ is then Cholesky decomposed by considering the diagonals belonging to the inactive atoms. From the two Cholesky decompositions, active and inactive MOs are obtained and the occupied–occupied space is defined.
 - The energy of A (in MLHF-A) and B (in MLHF-AB) are minimized (eq 4) in the MO space defined by the occupied orbitals of the active and inactive parts in an SCF procedure. Also, in this case, the equations defining the two approaches can be reduced to the occupied–

occupied space changing the AO basis through the occupied active and inactive MO coefficients.

- From MLHF-A/AB occupied MO coefficients, active and inactive densities are constructed and a new MLHF calculation is restarted from point 3 until convergence is reached. For all results reported in this paper, three macrocycles MLHF–MLHF-A/AB are sufficient to reach full convergence of the energy. It is worth noticing that since the MLHF calculation is restarted from the MO coefficients obtained at the 5th step, the total computational cost of MLHF-A/AB is only twice that of a standard MLHF calculation.

4. NUMERICAL APPLICATIONS

The capabilities of MLHF-A/AB are illustrated for four different molecular moieties that have previously been studied theoretically and experimentally.^{76–83} Those are 4-amino-4'-nitrostilbene (ANS), a part of a graphene sheet (which is indeed a graphene quantum dot), (*S*)-nicotine (in its most stable conformer⁷⁶), and [2,2]paracyclophane (PCP) (see Figure 2 for the molecular structures). Molecular geometries of ANS and nicotine are optimized at the B3LYP/aug-cc-pVDZ by using Gaussian 16 package.⁸⁴ The graphene sheet is constructed by setting the C–C distance to 1.42 Å and the C–H distance to 1.07 Å.⁸⁵ The PCP geometry is taken from ref 82. Graphene and ANS are chosen because they are conjugated systems. The conjugation is broken by our definition of the active regions as depicted in Figure 2a,b (in both cases, the bonding electrons are assigned to the inactive part). In the case of nicotine and PCP, single covalent bonds are cut and the bonding electrons are assigned to the active fragment (see Figure 2c,d). Hereby, we demonstrate the generality of our procedure, which can be

MLHF-AB

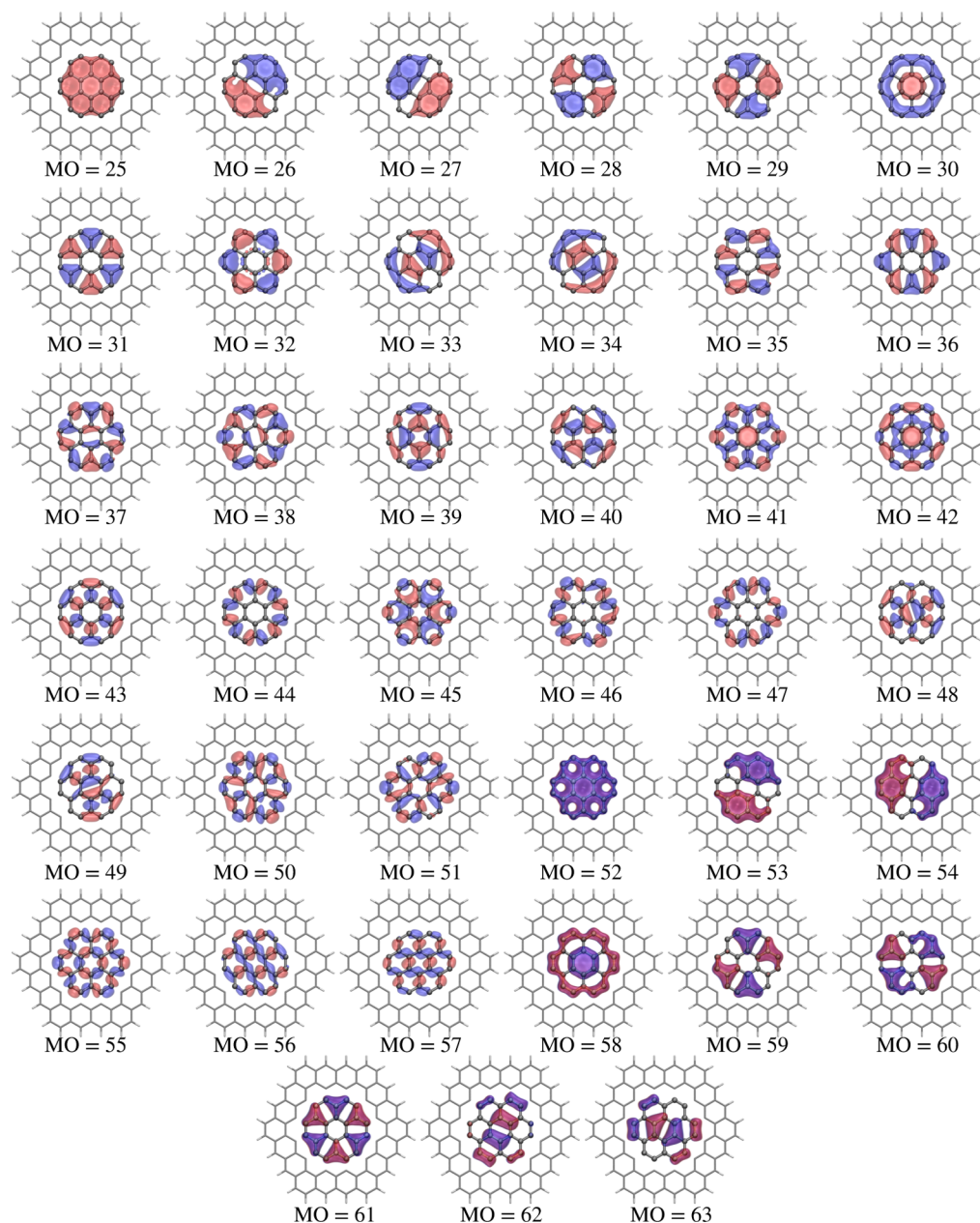


Figure 4. Graphical depiction of the active MOs of graphene as predicted by MLHF-AB. Basis sets: cc-pVTZ (active) and cc-pVDZ (inactive).

applied to different cases (single/double bond cutting) and to different definitions of the active region.

Nicotine and ANS calculations are performed by combining MLHF(/CC2) with aug-cc-pVDZ basis sets. The active part of the graphene sheet is described using cc-pVTZ basis set, whereas its inactive part with cc-pVDZ basis set. The PCP MLHF/CC2 calculations are performed with the triple-zeta quality 6-311G(d,p) basis set.⁸³ Notice that the different basis sets are chosen so to demonstrate the reliability of our approach in combination with diffuse/polarization functions.

The orbital second central moment (orbital variance) is used to quantitatively characterize orbital locality. The second central moment μ_2^p of an MO φ_p is defined as⁶⁷

$$\mu_2^p = \langle \varphi_p | \mathbf{r}^2 | \varphi_p \rangle - \langle \varphi_p | \mathbf{r} | \varphi_p \rangle^2 \quad (9)$$

The orbital spread σ_p is defined as the square root of μ_2^p . We also defined ξ as the average value of σ_p , that is, ξ is a measure of the mean locality of the considered set of MOs.¹ In this paper, MLHF-A/AB MOs are compared with canonical MLHF ones (named Cholesky because they are obtained through a Cholesky decomposition of the initial density matrix) that are also localized with the Boys procedure (Cholesky–Boys).⁶⁷ Notice that in Boys localization, the sum over p of μ_2^p in eq 9 is minimized,⁶⁷ and the obtained MOs can therefore be used as a reference for both MLHF-A and MLHF-AB approaches.

4.1. MLHF-A/AB Localized MOs. The most delocalized MOs of ANS and graphene are depicted in Figure 3, and the value of ξ and the maximum σ_p are also reported (see Sections S1.2 and S1.3 in the Supporting Information for the spreads of all occupied valence orbitals). First, we notice that in both ANS and graphene, Cholesky orbitals have the largest spread on

average (ξ) and the largest maximum MO spread ($\max\{\sigma_p\}$). As expected, both parameters are reduced by Cholesky–Boys. The MOs calculated by both methods are delocalized over the whole molecule (for both ANS and graphene).

A completely different picture arises when MLHF-A/AB methods are employed. The MOs obtained by both the latter approaches are well-localized on the active part only, and the values of ξ and $\max\{\sigma_p\}$ are reduced compared to the corresponding Cholesky counterparts. It is also worth noticing significant differences between MLHF-A and MLHF-AB, in particular for ANS. In fact, the most diffuse MLHF-A MO has a tail connecting active and inactive fragments, which should be absent since bonding electrons are assigned to the inactive part. Such a tail is completely absent in the case of MLHF-AB. From a physical point of view, this is not surprising. In fact, in the MLHF-AB procedure (see eq 4), the occupied orbitals of the active and inactive fragments are rotated in order to minimize the sum of the two energies. As stated above, such a rotation corresponds to maximizing the interaction energy between the two parts, that is, to maximizing the repulsion between them. As a consequence, the active occupied orbitals calculated by MLHF-AB are more localized on the active part.

The same does not apply to MLHF-A where the active energy is minimized in the occupied–occupied space. Thus, no constraints are imposed either on the inactive energy or on the interaction energy. However, notice that a few MOs have tails in both methods (see Sections S1.2 and S1.3 in the Supporting Information) since MLHF-A/AB orbitals are orthogonal.¹ The tails can be reduced by further localizing MLHF-A/AB orbitals using standard localization procedures. Notice also that the MLHF-AB ξ and $\max\{\sigma_p\}$ for ANS computed by using cc-pVTZ and aug-cc-pVDZ give very similar results, thus showing the consistency of our approach when diffuse functions are included (see Table S2 in the Supporting Information). The observations for ANS and graphene also apply to nicotine and PCP, whose MOs and corresponding spreads are reported in Sections S1.4 and S1.5 in the Supporting Information. For the latter systems, the ξ for Cholesky–Boys are lower than the corresponding MLHF-A/AB counterparts, but the MOs also spread in the inactive region. To illustrate the robustness of our approach, a different definition of active/inactive parts of ANS is also investigated (see Section S1.2.2 in the Supporting Information). The calculated results confirm the findings discussed here.

In Figure 4, we report the local MOs belonging to the active fragment of graphene as calculated by using the MLHF-AB method. All MLHF-AB local MOs are well-confined in the active part, and the symmetry of each orbital is evident. As a final comment, it is worth noticing that MLHF-AB orbitals may be further confined by using common localization procedures, such as Boys. The resulting local MOs will be more localized on the active atoms than those obtained by performing a hypothetical localization on the MOs resulting from SCF procedure of the entire system.

4.2. Ground-State Dipole Moments. The MLHF-A and MLHF-AB methods are also applied to calculate ground-state properties. We study the dipole moments of the active and inactive regions, together with the total dipole moments predicted by Cholesky, MLHF-A, and MLHF-AB. We are not reporting the results using Cholesky–Boys because a rotation among the active occupied orbitals does not change the active density and the density-related properties, such as dipole moment. The numerical values of active, inactive, and total

dipole moments for both nicotine and ANS are reported in Figure 5 together in a graphical representation of active (blue)

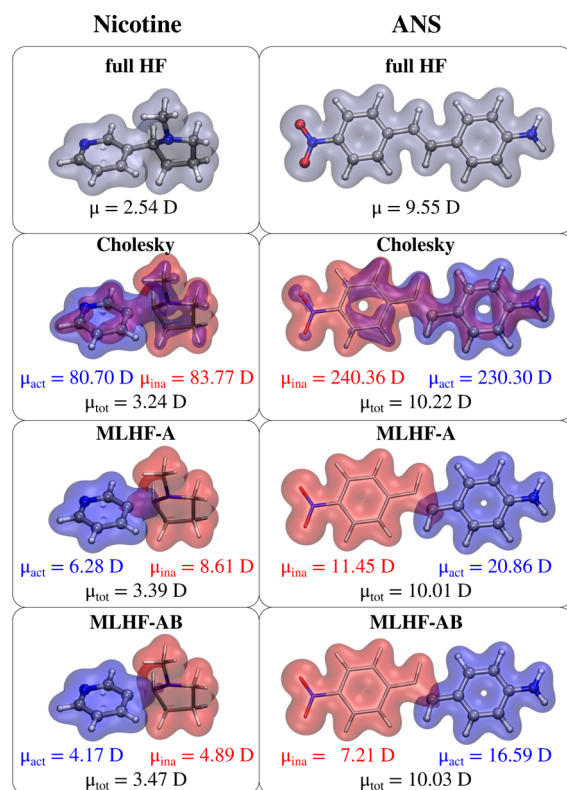


Figure 5. Nicotine (left) and ANS (right) MLHF (Cholesky)/MLHF-A/MLHF-AB dipole moments. For MLHF calculations, active (blue) and inactive densities (red) and their corresponding dipole moments are also given. Full HF dipole moments and molecular density are also shown. Basis-set: aug-cc-pVDZ.

and inactive (red) densities. Full HF densities and ground-state dipole moments for both molecules are also given and used as a reference. Dipole moments of both graphene and PCP structures are not reported because they are zero due to symmetry.

The MLHF (Cholesky) predicts large active/inactive dipole moments for both systems; for nicotine, they are almost 80 D, whereas for ANS almost 230 D. Such large dipole moments can be explained by investigating the spatial extension of active and inactive densities, which in both molecules are overlapping. This is due to the fact that the initial Cholesky decomposition defines an inactive density that overlaps with the active part and vice versa. Such issues are solved by MLHF-A/AB. Although both methods start from the same densities as those obtained by MLHF (Cholesky), the occupied–occupied rotations make the active and inactive densities more confined in their specific spatial regions, with a partial overlap limited to the bonding regions. As a consequence, the calculated dipole moments are very similar, in particular for ANS.

In both molecules, the MLHF-AB dipole moments are much lower than the corresponding ones for MLHF-A (see Figure 5). This is due to the maximization of the active–inactive repulsion in MLHF-AB. Thus, further confinement of the two densities in their specific spatial regions takes place. This can be appreciated by inspecting the bonding regions in both nicotine and ANS (Figure 5), showing that the overlap between active and inactive densities is lower in MLHF-AB than in MLHF-A. Notice also

that the bonding electrons of nicotine are assigned to the active part. Therefore, the active (blue) density in Figure 5 defines the bond, whereas for ANS the opposite applies. In the case of both nicotine and ANS, the total dipole moments are in very good agreement with the full HF reference value, with the largest discrepancy given by MLHF-AB for nicotine (error = 37%). The numerical value of the total dipole moments can be improved by using a different initial guess density, as, for instance, superposition of molecular densities (obtained by means of molecular fractionation with conjugate caps⁸⁶).

4.3. MLHF-AB Basis Set Dependence. In this section, we investigate the dependence of the MLHF-AB localization scheme on the chosen basis set. To this end, we have selected acetone and benzene. Acetone geometry was optimized at the B3LYP/aug-cc-pVDZ level using Gaussian 16,⁸⁴ whereas the benzene geometry was obtained from ref 87. The active part of acetone is defined as the C=O group including or discarding the bonding electrons, resulting in 16 and 12 active electrons, respectively. In the benzene moiety, the active part is constituted by half the molecule, and the bonding electrons are assigned to the inactive part. Therefore, 18 electrons are retained in the active fragment. Notice that the benzene moiety is selected to demonstrate the reliability of the MLHF-AB in treating highly symmetric systems.

First, we report a qualitative investigation of the localization of the active/inactive densities in their predefined spatial regions. In particular, we used cc-pVTZ, aug-cc-pVTZ, and d-aug-cc-pVTZ to study the effect of increasing the number of diffuse functions in the basis set. The densities obtained by using the three basis sets are plotted in Figure 6. Clearly, all active densities (blue) are well-confined in the active spatial region, and the basis-set does not affect such a result.

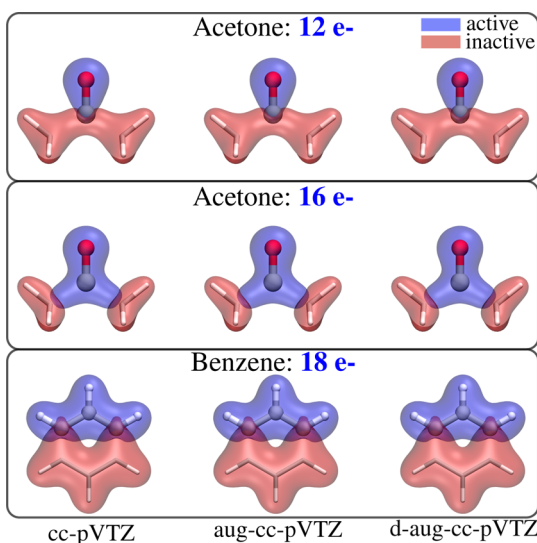


Figure 6. MLHF-AB/cc-pVTZ (left), MLHF-AB/aug-cc-pVTZ (middle), and MLHF-AB/d-aug-cc-pVTZ (bottom) active (blue) and inactive (red) densities of acetone (top, middle) and benzene (bottom). The number of active electrons are also reported.

As a second test case, we report the CCSD correlation energies of the studied moieties as a function of the cardinal number in the correlation-consistent cc-pVXZ basis sets. In particular, we performed the CCSD calculations in the active MO space as defined by the MLHF-AB procedure using the cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pVSZ basis sets. The

computed energies are then fitted by using the X^{-3} approach, to extrapolate the asymptotic convergence of the correlation energies.^{88,89} The results for the three studied systems are reported in Figure 7. For all three moieties, the CCSD correlation energies show an asymptotic convergence, thus demonstrating that our approach indeed has a basis set limit. This is not surprising and directly follows from the fact that the basis set is not truncated in the definition of the MO basis in the MLHF-AB approach.

4.4. MLHF-AB Versus Projection-Based Approaches. In this section, the MLHF-AB model is compared to HF projection-based approach (called projected-HF). In particular, we compute ground-state energies and dipole moments of the active fragments of nicotine and ANS molecules as a function of the elongation of the active–inactive covalent bond, which is cut by the partitioning into two fragments. The projected-HF results are obtained by first performing a SCF calculation on the entire structure. Then, the SCF MOs are localized by Boys localization procedure, and the MOs are assigned to the active and the inactive part. In order to be comparable with MLHF-AB results, the number of occupied MOs belonging to the active fragment is calculated by setting the number of the electrons in the active part ($n_o = n_{el}/2$) similarly to MLHF-AB calculations. The MOs belonging to the active fragment have to be selected on some mathematical criterion. Here, we calculate the percentage (p_i^A) of the i -th MO in the active part A as

$$p_i^A = \frac{\sum_{\mu \in A} C_{i\mu}^2}{\sum_{\mu \in A, B} C_{i\mu}^2} \cdot 100 \quad (10)$$

where, the $C_{i\mu}$ is the MO coefficient of the i -th MO in AO basis $\{\mu\}$. The n_o active MOs in projected-HF calculations are then selected as those having the highest percentage in the active atoms. It is worth pointing out that the active MOs in projected-HF models can also be selected as those having a percentage $\geq 50\%$, instead of fixing the number of active MOs to n_o . However, when applied to PES studies, such a choice leads to unavoidable PES discontinuities because a different number of active MOs may be selected depending on the active–inactive distance. Also, different methods to calculate the MO percentage in A can be arbitrarily chosen; thus the results are not unique. For these reasons, we prefer to keep the number of active MOs fixed to n_o . We notice that such arbitrariness is almost absent in MLHF-AB calculations, which only depend on the active–inactive partitioning of the electrons in the studied system.

In projected-HF method, once the active MO coefficients are selected, the active density is constructed ($D_{\mu\nu}^A = \sum_{ij} C_{i\mu} C_{j\nu}$), and the active energy is calculated as the E^A term in eq 4.

In Figure 8, MLHF-AB and projected-HF methods are applied to the calculation of ground-state energies and dipoles of the active fragment of nicotine as a function of the elongation of the single covalent bond connecting the active and inactive parts. In particular, two different basis sets, aug-cc-pVDZ (panel a) and 6-31+G* (panel b), are used. The equilibrium distance is 1.51 Å, and the covalent bond is varied from 1.00 to 2.1 Å. In both approaches, the number of active occupied MOs n_o is fixed to 21 because, as stated above, the bonding electrons are assigned to the active fragment.

The results reported in Figure 8a clearly show that by using the aug-cc-pVDZ basis set both MLHF-AB and projected-HF do not display any PES discontinuities (left panel). Also, the energy difference between the two approaches rapidly decreases

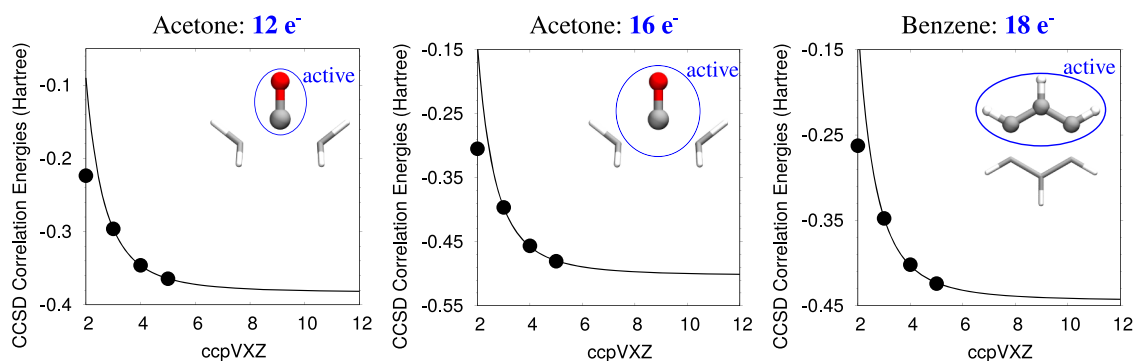


Figure 7. Convergence of CCSD correlation energies of acetone (left, middle) and benzene (right) as a function of the cardinal number in cc-pVXZ basis sets. The computed correlation energies (dots) are given together with the fitted curve. The number of active electrons is also reported.

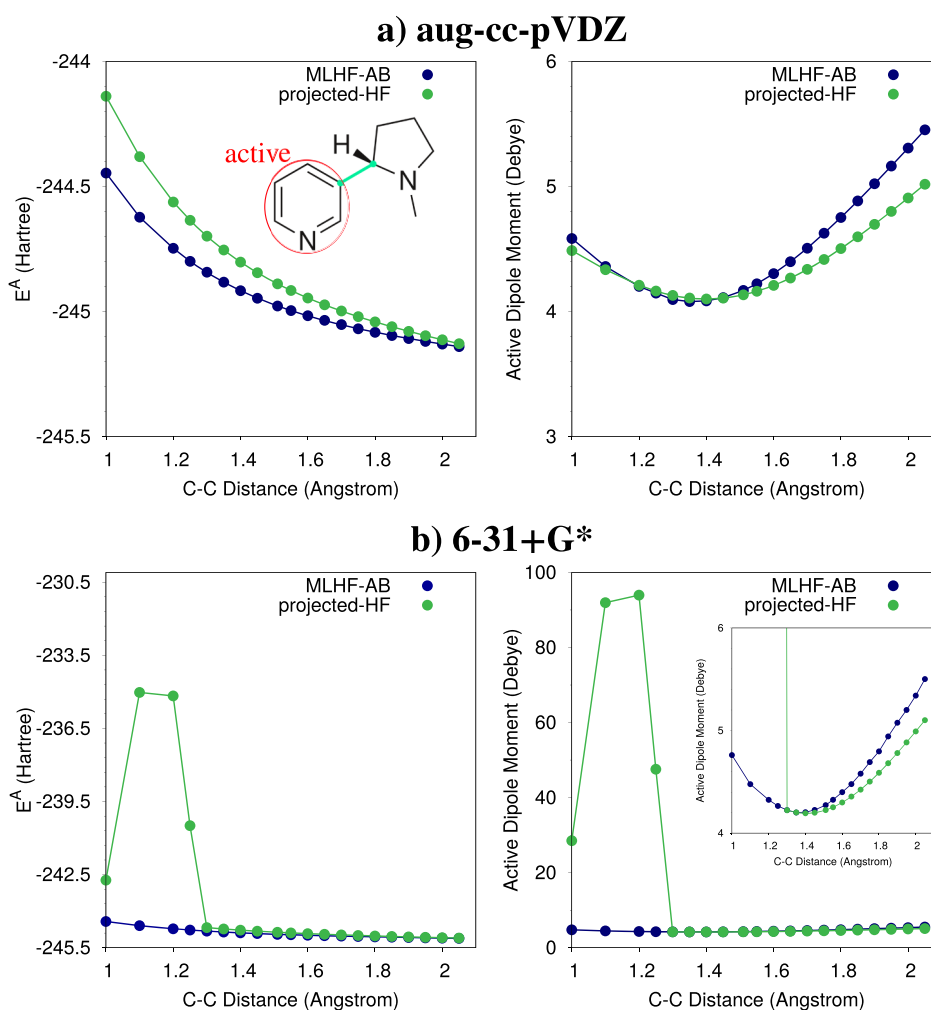


Figure 8. MLHF-AB and projected-HF active ground-state energy E^A (left) and dipole moment of the active part (right) of nicotine as calculated by using aug-cc-pVDZ (a) or 6-31+G* (b) basis sets. In panel (b), the y ranges are set so as to highlight the error due to the discontinuities obtained using Boys localization in projected-HF results.

as the active-inactive distance is elongated. At the equilibrium geometry, the MLHF-AB-projected-HF energy difference is of about 0.1 hartree, with the MLHF-AB energy that is lower than the projected-HF one at all considered distances. This is not surprising and results from the minimization procedure in MLHF-AB (see eq 6). The dipole moment of the active part is reported in the right panel of Figure 8a. Also in this case, the curves obtained by using both approaches do not display any

discontinuities, and a difference of about 0.3 debye is reported at the equilibrium geometry.

A different picture arises when the 6-31+G* basis set is used (Figure 8b). In this case, the projected-HF PES clearly displays a large discontinuity at small active-inactive distances, both in the ground-state energy (left) and in the dipole moment of the active part (right). Such a discontinuity reflects a discontinuity in the Boys space, which is common in MO localization procedures as it has been reported in different contexts.⁹⁰ The

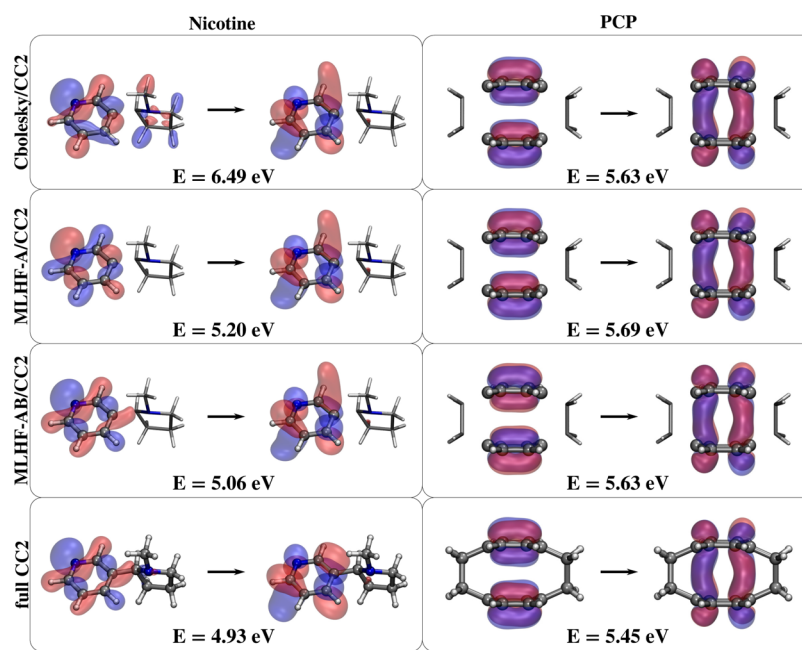


Figure 9. Nicotine (left) and PCP (right) MLHF (Cholesky)/MLHF-A/MLHF-AB and full CC2 excitation energies for the depicted transitions. Basis sets: aug-cc-pVDZ (nicotine) and 6-311G(d,p) (PCP).

discontinuity is completely absent in MLHF-AB. However, in the proximity of the equilibrium geometry, both approaches do not display any discontinuities. At the equilibrium geometry, the MLHF-AB energy is lower than the projected-HF by about 0.1 hartree, and the MLHF-AB-projected-HF active dipole moment difference is about 0.05 debye. The present analysis shows that the MLHF-AB PES is always continuous, whereas the projected-HF PES can display some discontinuities depending on the selected basis set. Notice that the results discussed for nicotine also apply to the case of ANS molecule (see Figure S27 given as the [Supporting Information](#)). For both nicotine and ANS, the average spread (ξ) of the LMOs used in projected-HF is smaller than that in MLHF-AB [1.71 a.u. vs 2.11 a.u. (nicotine) and 1.70 a.u. vs 2.68 a.u. (ANS)]. This is expected as the Boys localization minimizes the orbital spread ξ and therefore provides the LMOs with the lowest ξ . Also, the MLHF-AB is not intended to provide the most localized MOs overall but the most localized MOs in a specific spatial region. In passing, we note that MLHF-AB PES can display discontinuities depending on the initial Cholesky decomposition. However, this can be avoided by selecting the same pivots during the Cholesky decomposition of the initial density.

4.5. Absorption Energies. As a final application of MLHF-A and MLHF-AB, we select two local transitions, that is, occurring in the selected active parts, exhibited by nicotine and PCP (for which we investigate a through-space charge-transfer excitation,⁸¹ see [Figure 9](#)).

Local excitations are a perfect test case for demonstrating the capabilities of both approaches proposed here. In fact, the quality of the localized orbitals is crucial for obtaining reliable excitation energy. In this work, the excitation energies are computed using CC2⁶⁸ for the active MOs.

The MLHF (Cholesky) and MLHF-A/AB results are compared with full CC2 reference excitation energies. From the inspection of [Figure 9](#), it is clear that MLHF-A and MLHF-AB transition energies are in reasonable agreement with the reference values, particularly in the case of MLHF-AB. For

nicotine, MLHF (Cholesky) is completely unable to reproduce the full CC2 excitation energy because of the non-locality of the occupied MO. For PCP, all investigated methods give similar excitation energies, as the occupied orbitals are similarly reproduced by all approaches. We note a relatively large difference between MLHF-A/AB and full CC2 excitation energies. This discrepancy can be reduced by systematically increasing the number of atoms in the active region.⁶¹ In this case, such an increase will rapidly result in the whole system being treated as active. As previously reported,⁸¹ this system is highly correlated, and although the chosen active–inactive selection is justified by chemical intuition, the entire system needs to be correlated to achieve a correct reproduction of the excitation energies. However, we point out that the results reported here are chosen only to demonstrate the improved representation given by MLHF-A/AB as compared to MLHF (Cholesky).

4.6. Summary and Conclusions. To summarize, we have presented a novel energy-based criterion to localize MOs in specific spatial regions of a molecular system. In particular, this approach is based on a MLHF partitioning of the system. The method maximizes the repulsion between two predefined fragments and therefore can provide MOs which are mostly localized in the predefined fragment regions. The prospects of our approach are demonstrated for six selected systems characterized by both conjugated and non-conjugated skeletons. In particular, we have shown that MLHF-AB approach has the potentialities to provide continuous PES, thus solving the discontinuity issues that can arise by exploiting common localization procedures in projection-based approaches. Also, we have shown the stability of the proposed energy-based MO localization when diffuse functions are used. The basis set convergence is shown for two small systems, acetone and benzene.

The accuracy of our approach is then shown for ground-state properties (dipole moments) and excitation energies calculated at the full CC2 level. The computational cost is reduced due to

the partitioning of the system in active and inactive fragments. Both MLHF-A and MLHF-AB are able to reduce the discrepancy between MLHF and reference full CC2 values, in this way demonstrating their reliability in describing local excitations. Notice that in the present study, the procedure is applied to relatively small molecules in order to allow a direct comparison with full CC2 results. However, the model has the potential to be applied to very large systems. A detailed benchmark of the performances of MLHF-A/AB on excitation energies will be the topic of future communications.

To conclude, the MO localization provided by our approach can have different applications, ranging from those illustrated in this work (i.e., local ground-state properties^{91,92} and local excitations^{24,93,94}) to the accurate calculation of interaction and reaction energies of molecular systems in large biological matrices or adsorbed on nanomaterials.^{6,95,96} In addition, the local MOs obtained through our procedure may be used to define the different fragment densities in fragmentation approaches,³¹ and different boundaries in the cap regions for QM/MM approaches when covalent bonds are cut.⁹⁷

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.0c00737>.

Cartesian coordinates of the studied molecules; parameters of the calculations; and data related to Figures 3–5 and 9 (PDF)

Molecular geometry of acetone (XYZ)

Molecular geometry of graphene (XYZ)

Molecular geometry of ANS (XYZ)

Molecular geometry of nicotine (XYZ)

Molecular geometry of benzene (XYZ)

Molecular geometry of PCP (XYZ)

■ AUTHOR INFORMATION

Corresponding Authors

Tommaso Giovannini – Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway; orcid.org/0000-0002-5637-2853;

Email: tommaso.giovannini@ntnu.no

Henrik Koch – Scuola Normale Superiore, 56126 Pisa, Italy;

orcid.org/0000-0002-8367-8727; Email: henrik.koch@sns.it

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jctc.0c00737>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge Sarai Dery Folkestad and Ida-Marie Høyvik (NTNU) for discussions on technical aspects of the implementation. We acknowledge Chiara Cappelli (SNS) for computer resources. We acknowledge funding from the Marie Skłodowska-Curie European Training Network “COSINE—Computational Spectroscopy in Natural Sciences and Engineering”, grant agreement no. 765739, and the Research Council of Norway through FRINATEK projects 263110 and 275506.

■ REFERENCES

- (1) Høyvik, I.-M.; Jørgensen, P. Characterization and generation of local occupied and virtual Hartree–Fock orbitals. *Chem. Rev.* **2016**, *116*, 3306–3327.
- (2) Ma, Q.; Werner, H.-J. Explicitly correlated local coupled-cluster methods using pair natural orbitals. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1371.
- (3) Edmiston, C.; Ruedenberg, K. Localized atomic and molecular orbitals. *Rev. Mod. Phys.* **1963**, *35*, 457.
- (4) Boughton, J. W.; Pulay, P. Comparison of the boys and Pipek–Mezey localizations in the local correlation approach and automatic virtual basis selection. *J. Comput. Chem.* **1993**, *14*, 736–740.
- (5) Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. Analysis of charge transfer effects in molecular complexes based on absolutely localized molecular orbitals. *J. Chem. Phys.* **2008**, *128*, 184112.
- (6) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. Unravelling the origin of intermolecular interactions using absolutely localized molecular orbitals. *J. Phys. Chem. A* **2007**, *111*, 8753–8765.
- (7) Aquilante, F.; Bondo Pedersen, T.; Sánchez de Merás, A.; Koch, H. Fast noniterative orbital localization for large molecules. *J. Chem. Phys.* **2006**, *125*, 174101.
- (8) Høyvik, I.-M.; Jansik, B.; Jørgensen, P. Orbital localization using fourth central moment minimization. *J. Chem. Phys.* **2012**, *137*, 224114.
- (9) Jansik, B.; Høst, S.; Kristensen, K.; Jørgensen, P. Local orbitals by minimizing powers of the orbital variance. *J. Chem. Phys.* **2011**, *134*, 194104.
- (10) Høyvik, I.-M.; Kristensen, K.; Kjærgaard, T.; Jørgensen, P. *Thom H. Dunning, Jr.*; Springer, 2015; pp 287–296.
- (11) Høyvik, I.-M.; Jansik, B.; Jørgensen, P. Trust region minimization of orbital localization functions. *J. Chem. Theory Comput.* **2012**, *8*, 3137–3146.
- (12) Ziolkowski, M.; Jansik, B.; Jørgensen, P.; Olsen, J. Maximum locality in occupied and virtual orbital spaces using a least-change strategy. *J. Chem. Phys.* **2009**, *131*, 124112.
- (13) Gianinetti, E.; Raimondi, M.; Tornaghi, E. Modification of the Roothaan equations to exclude BSSE from molecular interaction calculations. *Int. J. Quantum Chem.* **1996**, *60*, 157–166.
- (14) Stoll, H.; Wagenblast, G.; Preuß, H. On the use of local basis sets for localized molecular orbitals. *Theor. Chim. Acta* **1980**, *57*, 169–178.
- (15) Li, W.; Ni, Z.; Li, S. Cluster-in-molecule local correlation method for post-Hartree–Fock calculations of large systems. *Mol. Phys.* **2016**, *114*, 1447–1460.
- (16) Zhang, X.; Carter, E. A. Subspace Density Matrix Functional Embedding Theory: Theory, Implementation, and Applications to Molecular Systems. *J. Chem. Theory Comput.* **2018**, *15*, 949–960.
- (17) Govind, N.; Wang, Y. A.; Carter, E. A. Electronic-structure calculations by first-principles density-based embedding of explicitly correlated systems. *J. Chem. Phys.* **1999**, *110*, 7677–7688.
- (18) Yu, K.; Carter, E. A. Extending density functional embedding theory for covalently bonded systems. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E10861–E10870.
- (19) Huang, C.; Pavone, M.; Carter, E. A. Quantum mechanical embedding theory based on a unique embedding potential. *J. Chem. Phys.* **2011**, *134*, 154110.
- (20) Libisch, F.; Huang, C.; Carter, E. A. Embedded correlated wavefunction schemes: Theory and applications. *Acc. Chem. Res.* **2014**, *47*, 2768–2775.
- (21) Knizia, G.; Chan, G. K.-L. Density matrix embedding: A simple alternative to dynamical mean-field theory. *Phys. Rev. Lett.* **2012**, *109*, 186404.
- (22) Knizia, G.; Chan, G. K.-L. Density matrix embedding: A strong-coupling quantum embedding theory. *J. Chem. Theory Comput.* **2013**, *9*, 1428–1432.
- (23) Sayfutyarova, E. R.; Sun, Q.; Chan, G. K.-L.; Knizia, G. Automated construction of molecular active spaces from atomic valence orbitals. *J. Chem. Theory Comput.* **2017**, *13*, 4063–4078.

- (24) Azarias, C.; Russo, R.; Cupellini, L.; Mennucci, B.; Jacquemin, D. Modeling excitation energy transfer in multi-BODIPY architectures. *Phys. Chem. Chem. Phys.* **2017**, *19*, 6443–6453.
- (25) Mennucci, B.; Corni, S. Multiscale modelling of photoinduced processes in composite systems. *Nat. Rev. Chem.* **2019**, *3*, 315–330.
- (26) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (27) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.
- (28) Giovannini, T.; Egidi, F.; Cappelli, C. Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864–22879.
- (29) Mennucci, B. Polarizable Continuum Model. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 386–404.
- (30) Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQF μ model. *J. Chem. Theory Comput.* **2019**, *15*, 2233–2245.
- (31) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* **2012**, *112*, 632–672.
- (32) Collins, M. A.; Bettens, R. P. A. Energy-based molecular fragmentation methods. *Chem. Rev.* **2015**, *115*, 5607–5642.
- (33) Pruitt, S. R.; Bertoni, C.; Brorsen, K. R.; Gordon, M. S. Efficient and accurate fragmentation methods. *Acc. Chem. Res.* **2014**, *47*, 2786–2794.
- (34) Collins, M. A.; Cvitkovic, M. W.; Bettens, R. P. A. The combined fragmentation and systematic molecular fragmentation methods. *Acc. Chem. Res.* **2014**, *47*, 2776–2785.
- (35) Pruitt, S. R.; Addicoat, M. A.; Collins, M. A.; Gordon, M. S. The fragment molecular orbital and systematic molecular fragmentation methods applied to water clusters. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7752–7764.
- (36) Khaliullin, R. Z.; Head-Gordon, M.; Bell, A. T. An efficient self-consistent field method for large systems of weakly interacting components. *J. Chem. Phys.* **2006**, *124*, 204105.
- (37) Ding, F.; Manby, F. R.; Miller, T. F., III Embedded mean-field theory with block-orthogonalized partitioning. *J. Chem. Theory Comput.* **2017**, *13*, 1605–1615.
- (38) Wen, X.; Graham, D. S.; Chulhai, D. V.; Goodpaster, J. D. Absolutely Localized Projection-Based Embedding for Excited States. *J. Chem. Theory Comput.* **2020**, *16*, 385–398.
- (39) Bennie, S. J.; Curchod, B. F. E.; Manby, F. R.; Glowacki, D. R. Pushing the limits of EOM-CCSD with projector-based embedding for excitation energies. *J. Phys. Chem. Lett.* **2017**, *8*, 5559–5565.
- (40) Chen, X.; Gao, J. Fragment Exchange Potential for Realizing Pauli Deformation of Inter-Fragment Interactions. *J. Phys. Chem. Lett.* **2020**, *11*, 4008.
- (41) Fertitta, E.; Booth, G. H. Energy-weighted density matrix embedding of open correlated chemical fragments. *J. Chem. Phys.* **2019**, *151*, 014115.
- (42) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F., III A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.
- (43) Fornace, M. E.; Lee, J.; Miyamoto, K.; Manby, F. R.; Miller, T. F., III Embedded mean-field theory. *J. Chem. Theory Comput.* **2015**, *11*, 568–580.
- (44) Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. Accurate first principles model potentials for intermolecular interactions. *Annu. Rev. Phys. Chem.* **2013**, *64*, 553–578.
- (45) Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. The effective fragment potential: a general method for predicting intermolecular interactions. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177–193.
- (46) Sun, Q.; Chan, G. K.-L. Quantum embedding theories. *Acc. Chem. Res.* **2016**, *49*, 2705–2712.
- (47) Chulhai, D. V.; Goodpaster, J. D. Projection-based correlated wave function in density functional theory embedding for periodic systems. *J. Chem. Theory Comput.* **2018**, *14*, 1928–1942.
- (48) Chulhai, D. V.; Goodpaster, J. D. Improved accuracy and efficiency in quantum embedding through absolute localization. *J. Chem. Theory Comput.* **2017**, *13*, 1503–1508.
- (49) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III Density functional theory embedding for correlated wavefunctions: Improved methods for open-shell systems and transition metal complexes. *J. Chem. Phys.* **2012**, *137*, 224113.
- (50) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III Accurate and systematically improvable density functional theory embedding for correlated wavefunctions. *J. Chem. Phys.* **2014**, *140*, 18A507.
- (51) Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller, T. F., III Exact nonadditive kinetic potentials for embedded density functional theory. *J. Chem. Phys.* **2010**, *133*, 084103.
- (52) Zhang, K.; Ren, S.; Caricato, M. Multi-state QM/QM Extrapolation of UV/Vis Absorption Spectra with Point Charge Embedding. *J. Chem. Theory Comput.* **2020**, *16*, 4361–4372.
- (53) Ramos, P.; Papadakis, M.; Pavanello, M. Performance of frozen density embedding for modeling hole transfer reactions. *J. Phys. Chem. B* **2015**, *119*, 7541–7557.
- (54) Pavanello, M.; Neugebauer, J. Modelling charge transfer reactions with the frozen density embedding formalism. *J. Chem. Phys.* **2011**, *135*, 234103.
- (55) Wesolowski, T. A.; Shedde, S.; Zhou, X. Frozen-density embedding strategy for multilevel simulations of electronic structure. *Chem. Rev.* **2015**, *115*, 5891–5928.
- (56) Jacob, C. R.; Neugebauer, J.; Visscher, L. A flexible implementation of frozen-density embedding for use in multilevel simulations. *J. Comput. Chem.* **2008**, *29*, 1011–1018.
- (57) Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050–8053.
- (58) Sæther, S.; Kjærgaard, T.; Koch, H.; Høyvik, I.-M. Density-Based Multilevel Hartree–Fock Model. *J. Chem. Theory Comput.* **2017**, *13*, 5282–5290.
- (59) Høyvik, I.-M. Convergence acceleration for the multilevel Hartree–Fock model. *Mol. Phys.* **2020**, *118*, 1626929.
- (60) Aquilante, F.; Boman, L.; Boström, J.; Koch, H.; Lindh, R.; de Merás, A. S.; Pedersen, T. B. *Linear-Scaling Techniques in Computational Chemistry and Physics*; Springer, 2011; pp 301–343.
- (61) Sánchez de Merás, A. M. J.; Koch, H.; Cuesta, I. G.; Boman, L. Cholesky decomposition-based definition of atomic subsystems in electronic structure calculations. *J. Chem. Phys.* **2010**, *132*, 204105.
- (62) Pulay, P. Second and third derivatives of variational energy expressions: Application to multiconfigurational self-consistent field wave functions. *J. Chem. Phys.* **1983**, *78*, 5043–5051.
- (63) Saebo, S.; Pulay, P. Local treatment of electron correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.
- (64) Culpitt, T.; Brorsen, K. R.; Hammes-Schiffer, S. Communication: Density functional theory embedding with the orthogonality constrained basis set expansion procedure. *J. Chem. Phys.* **2017**, *146*, 211101.
- (65) Hégyely, B.; Nagy, P. R.; Ferenczy, G. G.; Kállay, M. Exact density functional and wave function embedding schemes based on orbital localization. *J. Chem. Phys.* **2016**, *145*, 064107.
- (66) Huzinaga, S.; Cantu, A. A. Theory of separability of many-electron systems. *J. Chem. Phys.* **1971**, *55*, 5543–5549.
- (67) Boys, S. F. Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Rev. Mod. Phys.* **1960**, *32*, 296.
- (68) Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.
- (69) Myhre, R. H.; Koch, H. The multilevel CC3 coupled cluster model. *J. Chem. Phys.* **2016**, *145*, 044111.

- (70) Folkestad, S. D.; Koch, H. Multilevel CC2 and CCSD Methods with Correlated Natural Transition Orbitals. *J. Chem. Theory Comput.* **2020**, *16*, 179.
- (71) Folkestad, S. D.; Kjønstad, E. F.; Myhre, R. H.; Andersen, J. H.; Balbi, A.; Coriani, S.; Giovannini, T.; Goletto, L.; Haugland, T. S.; Hutchesson, A.; Høyvik, I.-M.; Moitra, T.; Paul, A. C.; Scavino, M.; Skeidsvoll, A. S.; Tveten, Å. H.; Koch, H. eT 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods. *J. Chem. Phys.* **2020**, *152*, 184103.
- (72) Lehtola, S. Assessment of initial guesses for self-consistent field calculations. Superposition of atomic potentials: Simple yet efficient. *J. Chem. Theory Comput.* **2019**, *15*, 1593–1604.
- (73) Koch, H.; Sánchez de Merás, A.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* **2003**, *118*, 9481–9484.
- (74) Christiansen, O.; Manninen, P.; Jørgensen, P.; Olsen, J. Coupled-cluster theory in a projected atomic orbital basis. *J. Chem. Phys.* **2006**, *124*, 084103.
- (75) Myhre, R. H.; Sánchez de Merás, A. M. J.; Koch, H. Multi-level coupled cluster theory. *J. Chem. Phys.* **2014**, *141*, 224105.
- (76) Egidi, F.; Segado, M.; Koch, H.; Cappelli, C.; Barone, V. A benchmark study of electronic excitation energies, transition moments, and excited-state energy gradients on the nicotine molecule. *J. Chem. Phys.* **2014**, *141*, 224114.
- (77) Marder, S. R.; Beratan, D. N.; Cheng, L.-T. Approaches for optimizing the first electronic hyperpolarizability of conjugated organic molecules. *Science* **1991**, *252*, 103–106.
- (78) Grigorenko, A. N.; Polini, M.; Novoselov, K. S. Graphene plasmonics. *Nat. Photonics* **2012**, *6*, 749.
- (79) Gibson, S. E.; Knight, J. D. [2.2] Paracyclophane derivatives in asymmetric catalysis. *Org. Biomol. Chem.* **2003**, *1*, 1256–1269.
- (80) Gleiter, R.; Hopf, H. *Modern Cyclophane Chemistry*; John Wiley & Sons, 2006.
- (81) Grimme, S. On the Importance of Electron Correlation Effects for the π - π Interactions in Cyclophanes. *Chem.—Eur. J.* **2004**, *10*, 3423–3429.
- (82) Demissie, T. B.; Dodziuk, H.; Waluk, J.; Ruud, K.; Pietrzak, M.; Vetokhina, V.; Szymański, S.; Jazwiński, J.; Hopf, H. Structure, NMR and Electronic Spectra of [m.n]Paracyclophanes with Varying Bridges Lengths (m, n = 2–4). *J. Phys. Chem. A* **2016**, *120*, 724–736.
- (83) Bachrach, S. M. DFT Study of [2.2]-, [3.3]-, and [4.4]-Paracyclophanes: Strain Energy, Conformations, and Rotational Barriers. *J. Phys. Chem. A* **2011**, *115*, 2396–2401.
- (84) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, Revision A.03; Gaussian Inc., Wallingford CT, 2016.
- (85) Castro Neto, A. H.; Guinea, F.; Peres, N. M. R.; Novoselov, K. S.; Geim, A. K. The electronic properties of graphene. *Rev. Mod. Phys.* **2009**, *81*, 109.
- (86) Zhang, D. W.; Zhang, J. Z. H. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein–molecule interaction energy. *J. Chem. Phys.* **2003**, *119*, 3599–3605.
- (87) Gauss, J.; Stanton, J. F. The equilibrium structure of benzene. *J. Phys. Chem. A* **2000**, *104*, 2865–2868.
- (88) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
- (89) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N₂, and H₂O. *Chem. Phys. Lett.* **1998**, *286*, 243–252.
- (90) Russ, N. J.; Crawford, T. D. Potential energy surface discontinuities in local correlation methods. *J. Chem. Phys.* **2004**, *121*, 691–696.
- (91) Giovannini, T.; Lafiosca, P.; Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854–4870.
- (92) Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective yet Reliable Computation of Hyperfine Coupling Constants in Solution by a QM/MM Approach: Interplay Between Electrostatics and Non-electrostatic Effects. *J. Chem. Phys.* **2019**, *150*, 124102.
- (93) Høyvik, I.-M.; Myhre, R. H.; Koch, H. Correlated natural transition orbitals for core excitation energies in multilevel coupled cluster models. *J. Chem. Phys.* **2017**, *146*, 144109.
- (94) Giovannini, T.; Ambrosetti, M.; Cappelli, C. Quantum Confinement Effects on Solvatochromic Shifts of Molecular Solutes. *J. Phys. Chem. Lett.* **2019**, *10*, 5823–5829.
- (95) Su, P.; Li, H. Energy decomposition analysis of covalent bonds and intermolecular interactions. *J. Chem. Phys.* **2009**, *131*, 014102.
- (96) Boulanger, E.; Thiel, W. Toward QM/MM simulation of enzymatic reactions with the drude oscillator polarizable force field. *J. Chem. Theory Comput.* **2014**, *10*, 1795–1809.
- (97) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.