Multi-scale multireference configuration interaction calculations for large systems using localized orbitals: Partition in zones

Cristian Chang,[1] Carmen J. Calzado,[2] Nadia Ben Amor,[3,4,a)] Jose Sanchez Marin,[1] and Daniel Maynau[3,4]

[1]*Institut de Ciencia Molecular, Parc Científic, Universitat de València, Catedrático José Beltrán, 2. E-46980 València, Spain*

[2]*Departamento de Química Física, Universidad de Sevilla, c/ Prof. García González, s/n. E-41012 Sevilla, Spain*

[3]*CNRS, UPS, LCPQ (Laboratoire de Chimie et Physique Quantiques), IRSAMC, 118, rte de Narbonne F-31062 Toulouse Cedex, France*

[4]*Université de Toulouse, UPS, LCPQ (Laboratoire de Chimie et Physique Quantiques), IRSAMC, 118, rte de Narbonne F-31062 Toulouse Cedex, France*

A new multireference configuration interaction method using localised orbitals is proposed, in which a molecular system is divided into regions of unequal importance. The advantage of dealing with local orbitals, i.e., the possibility to neglect long range interaction is enhanced. Indeed, while in the zone of the molecule where the important phenomena occur, the interaction cut off may be as small as necessary to get relevant results, in the most part of the system it can be taken rather large, so that results of good quality may be obtained at a lower cost. The method is tested on several systems. In one of them, the definition of the various regions is not based on topological considerations, but on the nature, σ or π, of the localised orbitals, which puts in evidence the generality of the approach. © 2012 American Institute of Physics. [http://dx.doi.org/10.1063/1.4747535]

I. INTRODUCTION

Many theoretical studies nowadays deal with the description of systems involving multireference wavefunctions, with relevant examples on different fields such as the molecular transition-metal based magnets, biomimetic mixed-valence systems, organic biradicals, or spin crossover complexes, and many others. Among the available approaches, multiference configuration interaction (MRCI) and the complete active space second order perturbation (CASPT2)[1] methods are both well-adapted to describe these multireference systems.

In the MRCI method, the molecular orbitals of the system are divided into three different sets. One distinguishes the occupied orbitals which bear two electrons in each reference determinant, the virtual ones which are always vacant and, finally, the active orbitals that can have any occupation. In the MRCI of the single and double (MRSDCI) approach, all determinants obtained through single and double excitations from the reference determinantal space are included in the CI space.

This kind of approach is well-adapted to chemical systems where the physics is not described by a single reference wavefunction. Contrarily to other approaches that are also based on the calculation of the wavefunction, an advantage of MRCI is that the zero order wavefunction is not described by a single reference configuration such as in coupled cluster. Another method that starts from a multireference zero order wavefunction is the complete active space second order perturbation approach. Both MRCI and CASPT2 are well-adapted to spectroscopy calculations. Compared to CASPT2, the MRCI approach presents some advantages. It has no problem of intruder states, and smaller active spaces are required to obtain relevant results as it can be seen in the example of the relative stability of 1:1 Cu-O2 adducts, treated at CASPT2 level[2] (16 electrons in 15 orbitals in the CAS), restricted active space (RAS)PT2 (18 electrons in 21 orbitals in the RAS),[3] and difference dedicated configuration interaction[4] (DDCI) method[5,6] (6 electrons in 4 orbitals in the CAS). Another example concerning the determination of magnetic constant coupling compares CASPT2 and DDCI methods[7] with minimal active spaces for the latest. The main drawback is the computational cost of the calculation, and this paper aims to reduce it as much as possible. Compared to coupled custer or CASPT2, a major problem of the truncated configuration interaction method is the lack of size-extensivity, but several corrections have been proposed in the past.[8–12]

The DDCI method corresponds to a MRSDCI with a more reduced space of determinants: all double excitations from the occupied to the virtual orbitals are excluded. This method has several advantages. The most time consuming part of an SDCI calculation (processing of the doubly excited determinants) is eliminated. Based on perturbative considerations, one can demonstrate that it yields different total energies, compared to SDCI, but the energy differences are correct, provided that the same set of orbitals is used for all states. Finally, the method largely reduces the size-consistency error. DDCI is considered as the reference method in the study of magnetic problems. It is used in this paper for the study of a Cu(II) system.

a)Author to whom correspondence should be addressed. Electronic mail: nadia.benamor@irsamc.ups-tlse.fr.

However, even with a dependence on the N number of basis functions of N^5 instead of N^6 for SDCI, the DDCI approach remains demanding. During the last decade, the great step towards the reduction of the computational cost in the calculation of correlation energy is due to the use of localised orbitals. Since electron correlation is a local phenomenon, using local orbitals allows one to neglect long range interactions, so that all orbitals involved in an elementary operation must be in the same region of space to product a nonvanishing effect. In this way, an N-scaling behaviour may be obtained. Without being exhaustive, a lot of developments may be mentioned, concerning self-consistent field (SCF), second-order Moller-Plesset (MP2), coupled-cluster single double (triple) (CCSD(T)), or MP2 gradient,[13–18] and others.

Concerning CI, N-scaling codes were proposed by Schwee *et al.*[19–22] and by some of us in Refs. 23 and 24. Our previous papers[23,24] were based on the use of a topological matrix between local orbitals. According to the value of the exchange integral K_{ij}, orbitals i and j are considered as *interacting* or *not interacting* orbitals, and this information is stored in the topological matrix. According to this information, both the size of the CI matrix and the list of electron repulsion integrals *(ab|cd)* are reduced. The present paper proposes a new improvement of the method. It is based on the idea that, in a molecule, some regions or some orbitals are more important than others for the study of its properties. This idea has been largely used in the past, in particular, in the quantum mechanics/molecular mechanics developments. In the work presented here, the difference of treatment between two distinct regions is not so important. Only the threshold to which is compared the exchange integral will differ depending on the spatial region of the involved orbitals. As a consequence no acute problem of frontier between regions will appear.

The first part of the present paper gives a brief recall of the excitation selected CI (EXSCI) method introduced in Ref. 23 and of the way to obtain localized orbitals, and presents the multi-scale CI approach developed in this work. The second part is devoted to applications of the method. First, the method is calibrated on the study of dissociation of a [Ca^{2+}–atrazine model pesticide] molecule for which two zones are defined, one of them including Ca^{2+} and the weak bonds between Ca^{2+} and the atrazine model, while the other one consists of the remaining part of the atrazine model. In a second application, an alternative zoning choice is proposed in the study of the interaction between H_2 molecule and a carbon nanotube fragment, where the different regions are no more chosen according to their spatial position in the system but to the σ or π nature of local molecular orbitals. These two tests concern the calculation of a potential energy surface (PES), which corresponds to a difficult case for a truncated method. Finally, the third application concerns the multireferential study of the magnetic properties in a cubane system.

II. METHOD

The method presented in this paper is based on the use of localised molecular orbitals. The electron correlation is a local phenomenon and, therefore, using localised orbitals al-

lows one to neglect long range interactions and to reduce the computational cost.

The general road map of the method can be divided into three steps:

- To obtain well-localised molecular orbitals, both occupied and virtual.
- To set up rules concerning the truncation of the CI matrix and integral list, thanks to the use of these local orbitals and then to establish the corresponding MRCI program.
- Finally, to use more or less demanding truncation criteria in the various regions, considering that a molecule can be in general divided into regions of unequal importance according to the property of interest.

A. *A priori* localisation method

The localisation method is described in detail in Ref. 25. It differs from the various methods that can be found in the literature, for example, those by Boys,[26] or Pipek and Mezey.[27] In these approaches, canonical SCF or CASSCF orbitals are localised according to a given criterion, i.e., the orbitals are optimised in a first step and then localised, so that one can refer to them as *a posteriori* methods.

In the *a priori* approach, guess local orbitals are built in a first step, and optimised to reach the SCF or CASSCF condition in a second step. The guess local orbitals are not uniquely defined and can be chosen according to the physical problem under consideration. They can be bond orbitals, lone pairs, atomic or fragment orbitals.

There are two possibilities to optimize the guess localized orbitals.

1. Use a super-CI like method. In this approach, a contracted CI matrix of the single excitations on the CAS eigenvector Ψ_0 is constructed and then block diagonalised. The block diagonalisation preserves the local character of the guess orbitals in the pseudo-natural orbitals resulting from this operation, which can be used in turn as guess orbitals for a new iteration. At convergence, the iterative process gives a set of local CASSCF orbitals (i.e., one obtains the CASSCF energy). The method is presented in Ref. 28.

2. Project each occupied (respectively, active, virtual) guess local orbital onto the space of occupied (respectively, active, virtual) canonical delocalised ones obtained by CASSCF calculation using a standard program. After orthonormalization, the method yields a set of well-localised occupied, active, and virtual orbitals, orthogonal and of CASSCF quality. This second approach has been used in all applications of this paper.

The *a priori* approach presents some drawbacks and some advantages as compared to the *a posteriori* one. The main defect is that the construction of the guess orbitals may be sometimes complicated. The main advantage is that one can easily get well-localised virtual orbitals, which it is not the case in many other approaches. In many situations, the drawbacks quoted above may be advantageous, since one can

build the local orbitals according to the physical problem under consideration. In both methods, the set of localized orbitals is not unique, but not for the same reasons. In the *a posteriori* approach, this is because SCF or CASSCF orbitals are not unique. In the *a priori* one, the final local orbitals depend on the construction of the guess, for which different choices may be done. One may notice however that, when the optimized local orbitals are obtained through a projection onto an SCF or CASSCF solution, even if the set of CASSCF orbitals is not unique, there is only one CASSCF solution (corresponding to the lowest energy), i.e., only one set of occupied, active, and virtual spaces. As a consequence, the final local orbitals, after projection onto CASSCF spaces, do not depend on the reference CASSCF orbitals.

B. Reducing the CI size and the list of bielectronic repulsion integrals: EXSCI program

After obtaining a set of localised occupied, active, and virtual orbitals, one retains only the excitations among "interacting local orbitals," using a threshold value for the corresponding exchange integral. Two orbitals a and b are *interacting orbitals* when the corresponding exchange integral: $K_{ab} = (ab|ba)$ is larger than a certain threshold *Th*. A detailed description of the procedure and its applications can be found in Refs. 23 and 24, but a short overview is given below.

In the Davidson[29] process of diagonalisation of the CI matrix, the most time consuming step is by far the product $H\Psi$ of the CI matrix H by the current vector Ψ. If one considers, in a CI of the single and double excitations (SDCI), the interaction of two doubly excited determinants $ij \rightarrow ab$ and $ij \rightarrow cd$ (they must have at least two common indices, here i and j, otherwise they do not interact), one performs the product,

$$\Delta(H\Psi)_{ij \rightarrow cd} = (ac|bd) \cdot \Psi_{ij \rightarrow ab},$$

which means that the coefficient of $H\Psi$ corresponding to the determinant $ij \rightarrow cd$ is modified by a quantity equal to the value of the product of the two-electron repulsion integral $(ac|bd)$ times the coefficient of Ψ corresponding to $ij \rightarrow ab$. Since six different orbitals i, j, a, b, c, d are present in this product and each of them can take N values, the computational cost grows as N^6 (N is the total number orbitals; all orbitals may be chosen independently).

With local orbitals, the value of the integral $(ac|bd)$ vanishes if the indices are not in the same region of space. In the same way, the weight of the $ij \rightarrow ab$ and $ij \rightarrow cd$ doubly excited determinants is expected to be almost zero if i or j does not belong to this region. Finally, as all orbitals must be close together, the N-dependence is no longer N^6, but $N_n^5.N$, where N_n is the average number of orbital neighbours for a given orbital. N_n is independent on the size of the molecule and, therefore, a N-scaling behaviour is obtained.

1. Brief recall of the EXSCI program

The EXSCI method has been initially proposed by Bories *et al.*[23] and recently improved by Ben Amor *et al.*[24] The rules that decide whether a determinant or a two-electron integral

can be eliminated (from the CI space, or from the integral list) are detailed in Ref. 24. A brief summary is given here.

A file containing the one and two electron integrals in a basis of localised orbitals is given as input. After defining the active space, some information concerning the construction of the interacting orbital topological matrix must be yielded. As said above, orbitals i and j are considered as *interacting orbitals* if the exchange integral K_{ij} is larger than a certain threshold, and this information is stored in a Boolean matrix. Two Boolean matrices are used, one concerns the determinants and the other one the integrals. If different thresholds are given for the determinant and for the integrals, the two Boolean matrices are not identical. These two matrices fully determine the truncated lists of integrals and determinants.

Concerning **integrals**, the condition that determines if one of them is kept or eliminated is that all orbitals involved must be connected trough the *interacting* condition, such as depicted in the scheme below.
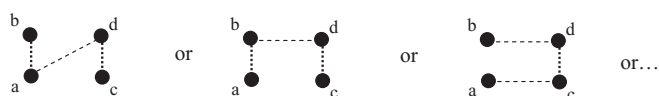
The rules to include or not a **determinant** in the CI space are a bit more complicated.

– All active orbitals are considered as *interacting orbitals*. Therefore, all reference determinants are present in the CI space

– An excited determinant is kept if all nonactive involved orbitals interact either with an active orbital or with another involved nonactive one. In this later case, the two orbitals form a hole-particle pair. For example, the rules to include a doubly excited determinant are illustrated by the following scheme.
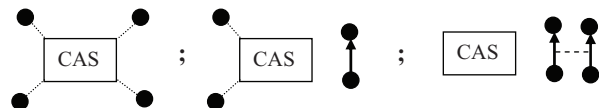
– However, further refinements must be added. In the last diagram corresponding to a double nonactive excitation, the dashed line means that the two pairs are not too distant. If they are distant, such as in the following case,

the diagram should correspond to a dispersive interaction. Depending on the nature of the physical problem under study, one can decide to include these dispersion effects or not. Therefore, a second threshold was added to build the determinant basis set.

To summarize, three thresholds can be set, two for the determinants, (namely, a principal Th_1, and the additional Th_2 for the dispersive effects), and the third one, Th_i, for the integrals. When using the three thresholds, a quasi-linear scaling

is obtained. Reference [24] present a detailed study of this behaviour, based on several test examples.

C. Multi-scale EXSCI

In the method presented above, all parts of the molecule are treated on an equal foot. However, in most systems, some parts play a more important role than others. For example, in an organometallic system, one can expect that the region near the metal atom(s) should require to be treated in a more accurate way than more distant ligand or ligand parts. In the same way, in any conjugated hydrocarbon molecule, the π system plays a larger role than the σ skeleton in the description of the spectrum. The new code allows one to treat at different level of accuracy orbitals of unequal importance in the system.

In the new version of the program, the molecule can be divided in zones according to their relevance to the problem of interest. A zone is defined by a set of atoms or orbitals. Each orbital belongs to a given zone, depending on its importance in the molecule, and this choice is made by the user of the program, according to physical considerations. In each zone, the values of the three thresholds Th_1, Th_2, and Th_i are different. The smaller their value, the larger the number of determinants and integrals. Hence the Th values should be set smaller in the most relevant parts of the system.

Once fixed these thresholds, a new topological matrix is obtained, and the rest of the procedure remains unchanged. If two orbitals belong to the same zone, they are considered as *interacting* or *not interacting* according to the thresholds of the zone. If two orbitals belong to different zones, several solutions have been envisaged. A first possibility is to use the lowest of the two threshold values. One can decide, on the contrary, to use the largest one. Finally, a third possibility is to use the average value between the two considered zones. The average solution appears to be a good compromise, the most relevant and easy to use and has been chosen as the default.

III. APPLICATIONS

A. Complexation reaction of an atrazine metabolite with Ca^{2+}

The 2-chloro-1,3,5-triazine-4,6-diamine molecule (see Figure 1) is a metabolite of the atrazine pesticide. This pesticide reacts with soil cations to give complexes. Then, the study of these complexations matters to understand the adsorption and desorption processes of pesticides in soil. The complex of the atrazine model with Ca^{2+}, which has large
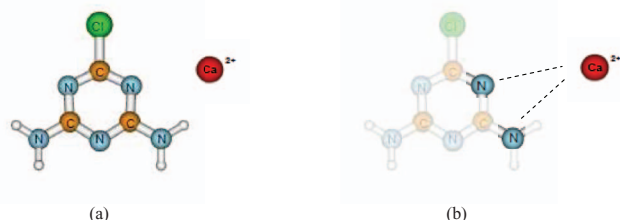
complexation energy,[30] has been studied in a previous paper[24] to evaluate the ability of selected SDCI calculations to correctly describe a reaction path. In this kind of study, the property that mainly changes from a distance to another is the intermolecular interaction, in particular, dispersion and double excitation charge transfer involving the Ca^{2+} ion and the atrazine model. The use of thresholds in such a case is challenging since these small interactions should not be neglected. Small thresholds were then necessary to describe correctly the potential energy curve: $Th = 0.0003$ a.u. (Th means that $Th_1 = Th_2 = Th_i$) to get an error of 1.3 kcal mol^{-1} or $Th = 0.0001$ a.u. to reduce this error to 0.5 kcal mol^{-1}. The CI space was reduced at least to 30% from the standard SDCI calculation. Even though the reduction can be considered as quite important, it is not satisfactory for a quasi-linear method. A solution could come from treating in a zone the atoms involved in these intermolecular interactions. In the present work, the Ca^{2+} and the two nearest nitrogen atoms (Figure 1) are then assigned to the same zone (zone 1), with small thresholds ($Th = 0.0001$ a.u.), while the remaining atrazine model atoms (zone 2) are treated at lower precision, with larger thresholds ($Th = 0.003$ a.u. or $Th = 0.001$ a.u.). With this partition, the interaction energy should be obtained precisely enough at a lower cost than using small thresholds on the whole complex.

With respect to the previous paper, the same geometry[30] and double-zeta plus polarization basis sets (standard Pople 6-31G* basis set for C, N, H, and the corresponding Francl for Cl[31] and Blaudeau for Ca[32]) are used. The distance has been varied from the equilibrium situation (2.372 Å to the closest cycle nitrogen) up to 12.372 Å, keeping the N-Ca^{2+}-N angle constant. The largest distance corresponds to a total elongation of 10 Å. As the total complexation energy is large (121.07 kcal mol^{-1} at the standard SDCI level) compared with the error, it is convenient to look at the relative error to SDCI calculations, taking the longest 10.0 Ångströms distance as reference point. In Figure 2 are drawn the previous results obtained with several Th values without using

FIG. 1. (a) Ca^{2+}–Atrazine model at the equilibrium geometry. (b) Two zones are defined: zone 1 in bold ($Th = 0.0001$ a.u.) and zone 2 in watermark (various thresholds are used).
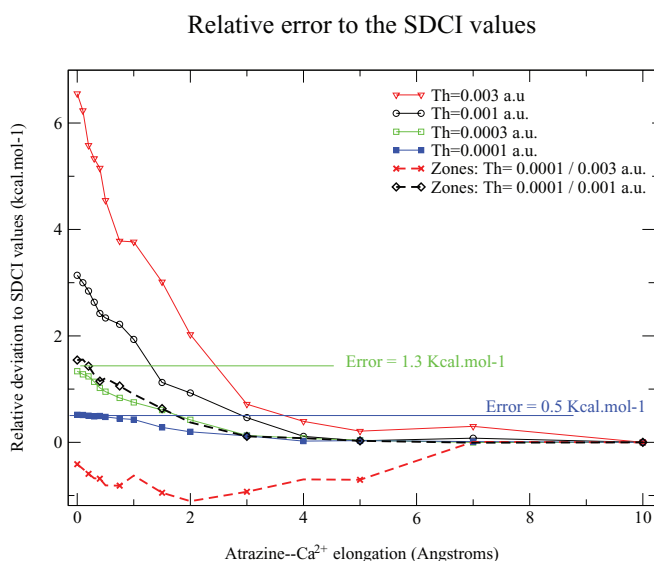

FIG. 2. Atrazine–Ca^{2+} elongation (Ångströms). Comparison of selected CI with vs without zones. Relative error to the SDCI values.

TABLE I. Ca^{2+}–Atrazine model. Dimensions of the calculations with and without partition into zones as a function of integral exchange thresholds (a.u.). Average thresholds are used between the two zones.

Zone 1 (Ca^{2+} --- N) ⟍ N	Zone 2 (remaining atrazine atoms)	CPU/it (s)	CI size at 0.0 Angströms	CI size at 10.0 Angströms (% compared with 0.0 Angströms)	Error on the dissociation energy compared to SDCI (kcal mol^{-1})
0.0001	0.003	134	1 297 713	830 365 (64%)	−0.4
0.0001	0.001	423	2 773 530	2 097 227 (76%)	1.5
	0.0003	797	4 093 170	3 391 558 (83%)	1.3
	0.0001	2662	7 573 990	5 855 697 (77%)	0.5

the zones approach. Two new curves have been added, with a two zones approach. In both cases, $Th = 0.0001$ a.u. for the zone 1 but different thresholds have been used for the zone 2: $Th = 0.001$ a.u. and $Th = 0.003$ a.u. In this test, average thresholds are employed for terms that involve the two zones. The error on the dissociation energy relative to SDCI, obtained with zones, is about 1.5 kcal mol^{-1} for the lowest threshold ($Th = 0.001$ a.u.) on zone 2, while it is around –0.4 kcal mol^{-1} but more noisy for the largest one, for which the maximal error is larger (–1.1 kcal mol^{-1}) than the error on the dissociation energy. Despite a global better behaviour for the largest thresholds, the experience shows that it is preferable to diminish the thresholds until the disappearance of the noise. The involved errors are all quite small in the present case.

Compared to the result without partition into zones ($Th = 0.0003$ a.u.), for a result of similar quality (1.3 kcal mol^{-1}), the CI matrix size and the central processing unit (CPU) time per iteration are significantly lowered (see Table I). One can observe that the percentage of determinants at 10.0 Ångströms compared to the equilibrium distance varies from 64% (calculation with a threshold $Th = 0.003$ a.u. on zone 2) up to 83%. This point will be discussed later.

1. Interzone threshold

In a next step, a test of the different ways to treat the determinants or the integrals that involve orbitals in dif-

ferent zones is investigated. The corresponding threshold will be called hereafter "interzone threshold." As mentioned above, three choices are considered: maximal, average, or minimal threshold. For example, if the following thresholds $Th = 0.0001$ a.u. for zone 1 and $Th = 0.003$ a.u. for zone 2 are chosen, then, between these two zones, if **maximal** interzone threshold is selected, the exchange integral that connects two orbitals belonging to different zones must be lower than the maximal threshold of the two zones, i.e., $Th = 0.003$ a.u. of zone 2. On the contrary, if **minimal** is chosen, the threshold of zone 1 is relevant. Finally, for **average** interzone threshold choice, the exchange integral must be lower than the average value of the two thresholds, i.e., 0.00155 a.u. Dimensions of the corresponding calculations are shown in Table II and results are depicted in Figure 3. The time per iteration is drastically reduced when a threshold of 0.003 a.u. is used for zone 2 and interzone. As one can see in Figure 3, all results are quite reasonable. The deviation to the SDCI calculation ranges from –1.1 to 1.6 kcal mol^{-1}.

2. Discontinuities and topological matrix

One can notice that discontinuities appear, particularly important for the 0.003 a.u. threshold. This is due to the fact that orbitals which are interacting at a certain distance can be noninteracting when distance varies even for small distance

TABLE II. Ca^{2+}–Atrazine model. Dimensions of the calculations as a function of the interzone threshold options (a.u.).

Zone 1 (Ca^{2+} --- N) ⟍ N	Zone 2 (remaining atrazine atoms)	Inter zone	CPU/it (s)	CI size at 0.0 Angströms
0.0001	0.003	Max (0.003)	84	971 221
0.0001	0.003	Aver (0.00155)	134	1 297 713
0.0001	0.003	Min (0.0001)	843	3 763 160
0.0001	0.001	Max (0.001)	283	2 245 776
0.0001	0.001	Aver (0.00055)	423	2 773 530
0.0001	0.001	Min (0.0001)	1222	4 782 697
	0.0003		797	4 093 170
	0.0001		2662	7 573 990
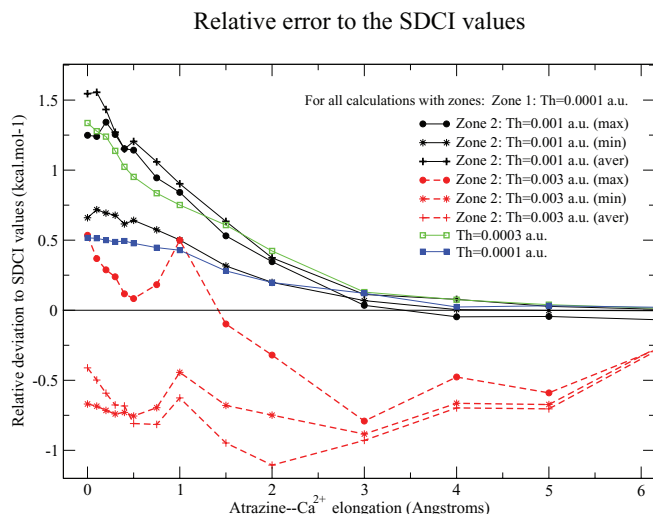
FIG. 3. Atrazine–Ca^{2+} elongation (Ångströms). Interzone thresholds. Relative error to the SDCI values.

FIG. 4. Atrazine–Ca^{2+} elongation (Ångströms) with transfer of the topological matrix. Relative error to the SDCI values.

changes. That is, the number of interacting orbitals decided by the exchange integral cut-off varies along the reaction path.

A solution can come from fixing these interactions at a given distance and to transfer the corresponding topological matrix to other geometries. The most judicious choice is to transfer the topological matrix of the shortest distance, which corresponds to the maximal number of interactions, to the other geometries. The size of the CI matrix is then the same all along the reaction path. This way of doing is also employed in the following test applications presented in this paper. As soon as reduction of the computational cost is obtained through truncations based on cut off criteria, discontinuity phenomena appear. The general solution is to transfer the truncation of a particular point, in general, the point with minimal truncation like it is done here, or, if no point plays a particular role, to use the union of all determinant spaces for all distances. The limitation of this kind of solution is that the orbitals should be of the same nature for all points of the PES.

The corresponding potential energy curves are presented on Figure 4. The improvement is impressive. Only the case
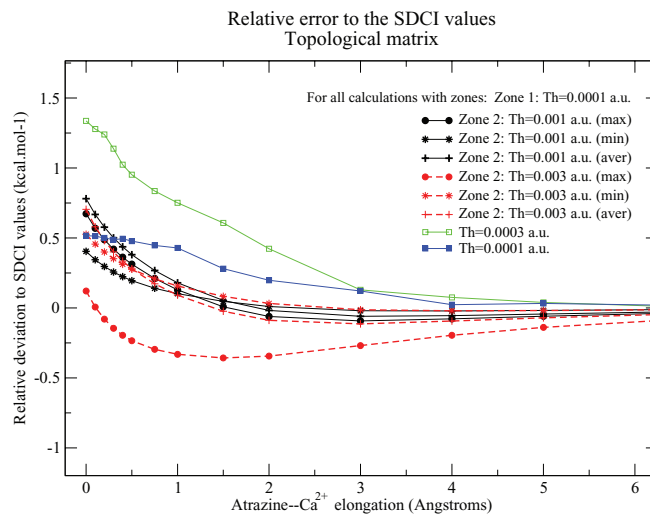
with $Th = 0.003$ a.u. and maximal interzone threshold moves away from the others, but even in this case, the error on the interaction energy is only –0.4 kcal mol^{-1}. The use of topological matrix not only makes smooth the potential energy curve but also gives better complexation energy by reducing the error at dissociation. In Table III, we compare the maximal error on the interaction energy with and without using the topological matrix.

One can conclude that: (1) the use of average thresholds for the interzone threshold part is relevant and already save a lot of CPU time compared to a calculation without zones; (2) for reaction pathways, the use of the same topological matrix along the pathway is suitable and gives excellent results. Compared to the CISD interaction energy (121.07 kcal mol^{-1}) the error is lower than 1%.

B. Axial approach of a hydrogen molecule to a nanotube fragment

The nanosystems have been a subject of great interest during the last decade. In particular, the approach of small

TABLE III. Ca^{2+}–Atrazine model. Topological matrix transferred. Maximal error on the potential energy curve compared to standard SDCI as a function of exchange integral thresholds (a.u.) with and without zones.

Zone 1 (Ca^{2+} --- N) Th (a.u.)	Zone 2 (remaining atrazine atoms) Th (a.u.)	Interzone Th (a.u.)	Maximal error compared to SDCI (kcal mol^{-1})	Maximal error compared to SDCI (topological matrix of the equilibrium distance) (kcal mol^{-1})
0.0001	0.003	Max (0.003)	−0.8	−0.4
0.0001	0.003	Aver (0.00155)	−1.1	0.7
0.0001	0.003	Min (0.0001)	−0.9	0.5
0.0001	0.001	Max (0.001)	1.3	0.7
0.0001	0.001	Aver (0.00055)	1.6	0.8
0.0001	0.001	Min (0.0001)	0.7	0.4
	0.0003		1.3	
	0.0001		0.5	

FIG. 5. Schematic view of the axial approach of a hydrogen molecule to a (0,9) nanotube fragment.

FIG. 6. H_2 approach to a nanotube fragment: σ and σ^* localized orbitals.

molecules or atoms to a nanosurface or nanotube has been largely studied.[33–38] These systems have extended sizes, and it is therefore relevant to address them with the methods developed in this paper. No extensive study is done here since the goal is to illustrate the application of the present methods based in the CI techniques in this particular example of an extended system.

The nanotube fragment selected for the test calculation is shown in Figure 5. It contains 54 carbon atoms and the study corresponds to the determination of the potential energy curve as a function of the distance of a hydrogen molecule moving along the cylindric symmetry axis. The carbon nanotube fragment belongs to a zigzag (0,9) nanotube and has been thoroughly selected so that both the diameter and length allow for a significant and representative enough interaction with the axially approaching H_2 molecule. The edge C valences had been saturated with dangling single C–H bonds. In all calculations, the core $1s$ orbitals are frozen at the CI level. The basis sets used are atomic natural orbitals (ANOS),[39] double zeta for all orbitals of the nanotube and double zeta plus polarisation for H_2. In the following, for sake of simplicity, the ANO corresponding to the $2s$ and $2p$ orbital of the carbon atoms will be called "valence" orbitals, and the second ANOs of the double zeta basis set, $3s$ and $3p$, "diffuse" orbitals. The symmetry group used in the calculation is C_s, with the H_2 molecule contained in the symmetry plane. The origin of distances is placed at the centre of the nanotube fragment. Geometries are not relaxed during the approach process.

1. Localisation scheme

This system is rather large, and the goal of this test is to establish some rules that would make possible the study of even larger systems. H_2 mainly interacts with the π cloud of the nanotube. Then, one can decide to separate the σ and π parts and to treat them in different zones. Since these two sets of orbitals do not belong to different symmetries and are not defined by an axis direction as in planar systems, it is not straightforward to distinguish between them. This condition will be obtained by first localising the σ part and then building local π orbitals that are forced to be locally orthogonal to the tangential plane. The "diffuse" functions ($3s$ and $3p$) are atomic orbitals; therefore they are already well-localised. As only the π system must be treated accurately, the number of orbitals predominantly involved in the dynami-
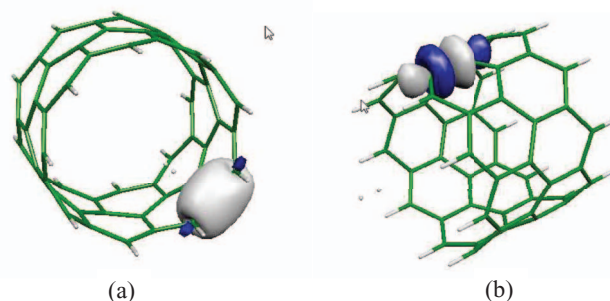
cal correlation calculation remains small, even for larger nanotube fragments. As a consequence, the computational cost should not grow fast with the size of the nanotube. Furthermore, if localised π orbitals are used, the scaling should be linear.

The localisation of the orbitals is straightforward for H_2, for the $1s$ core of carbon atoms and for the C–C σ and σ^* orbitals. By way of example, a $\sigma_{C–C}$ and a $\sigma_{C–C}^*$ orbitals are represented on Figure 6. Both orbitals appear as well-localised.

The localisation of the π part is less straightforward. One could be tempted to use a Kékulé description (Fig. 7(a)). There are several possibilities, but none respects the symmetry of the problem. A solution is to centre each π local orbital on a carbon atom. Each C_6 ring can bear 3 occupied orbitals centred on 3 nonadjacent carbon atoms (see Fig. 7(b)). This is a quite unusual and surprising proposition, since the carbon atoms seem to be nonequivalent, and the charges over them to be alternatively 0 and 2. One must notice, however, that this is only the guess definition. After
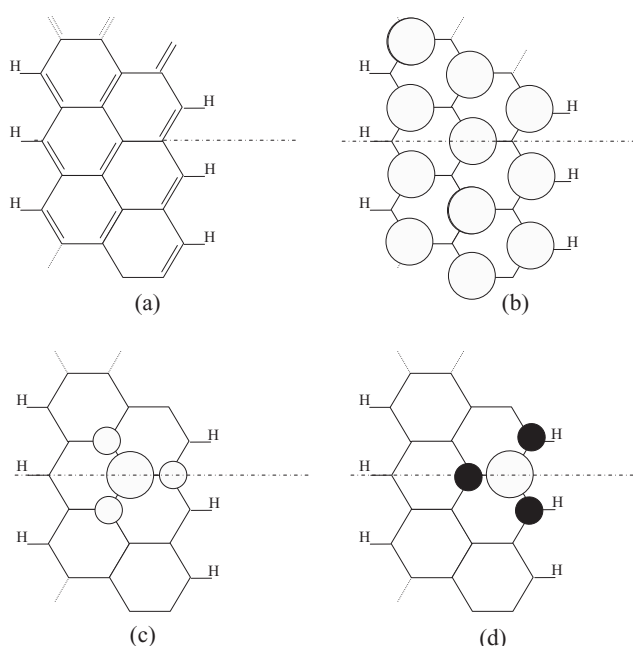

FIG. 7. Surface of the nanotube. Kékulé π orbitals, and atom centered π and π^* orbitals as they are used in the localisation procedure.
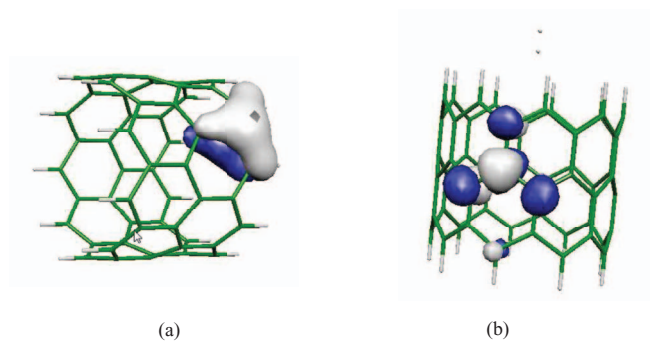
(a)   (b)

FIG. 8. H_2 approach to a nanotube fragment: π and π^* localized orbitals.

FIG. 9. H_2 approach to a nanotube fragment. Taking account of the σ skeleton through variation of thresholds. Since the total energies obtained are different in the various cases, all energies are set to zero at 15 Angströms, in order to see the deviation from the reference curve.

projecting these orbitals on the correct SCF solution and after orthonormalising them, orthogonalisation tails will appear on neighbour atoms, as shown in Fig. 7(c), in such a way that the electronic charge will be equivalent on all carbon atoms, as expected.

In the same way, the corresponding π^* empty orbitals (we can call them antibonding, because they play this role and show nodal planes along skeletal sigma bonds) are positioned on the unoccupied carbons of the ring, showing also tails on the neighbours, but with opposite phase (see Fig. 7(d)). The shape of these π and π^* orbitals as they appear in the calculation are plotted on Figs. 8(a) and 8(b). These calculated orbitals are not very different from those of the corresponding model on Figure 7. They satisfy the two conditions of separating the σ and π systems and having a limited spatial extension.

2. Zones

The problem is now to find a procedure to treat the σ and π parts at different levels of precision. A first possibility is to freeze the whole σ core. In a more sophisticated treatment, the σ and π systems should be allowed to interact. This means that, if one considers a local $\pi \rightarrow \pi^*$ excitation, which corresponds to a charge fluctuation, the underlying σ skeleton must have the capacity to respond to it through repolarisation effects. In this test study, three different zones have been considered:

- Zone 1: π, $\pi*$ orbitals of the nanotube and all H_2 orbitals.
- Zone 2: "diffuse" $3s$ and $3p$ atomic orbitals on carbon atoms of the nanotube.
- Zone 3: σ_{CC}, σ_{CH}, $\sigma*_{CC}$, and $\sigma*_{CH}$ orbitals.

For the (π, H_2) zone 1, thresholds are set to low values ($Th_1 = 0.002$ a.u.; $Th_2 = 0.001$ a.u.) and never changed. Concerning the two other zones, various thresholds are used in order to analyse the relative role of the σ skeleton and of the diffuse functions on the interaction energy of the H_2 molecule with the nanotube during its approach and also to find the higher thresholds allowed without precision loss.

To avoid an additional parameter, the thresholds on the bielectronic integrals are unchanged for all calculations: $Th_i = 0.0001$ a.u. for zone 1, $Th_i = 0.0003$ a.u. for zone 2,
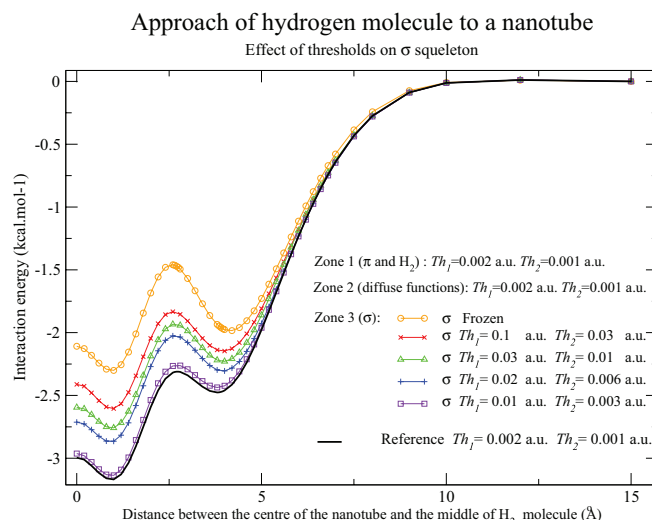
$Th_i = 0.001$ a.u. for zone 3. For the reference calculation, the lowest integrals threshold is used ($Th_i = 0.0001$ a.u.).

The average threshold choice, as defined above, has been used to rule the interactions when orbitals of two zones are involved. To avoid discontinuities, the topological matrices that give the lists of determinants and of integrals at the distance corresponding to the central position have been kept along the pathway reaction.

3. Results

In the reference calculation, all orbitals are considered as belonging to the same zone, with a low thresholds set ($Th_1 = 0.002$ a.u.; $Th_2 = 0.001$ a.u.), which gives a reference potential curve (black-bold on Figures 9 and 10). The dimension of the SDCI spaces and their relative CPU times in the calculations using different threshold sets are given in Table IV. The reduction of the calculation cost is important, even for the lowest thresholds for which the potential curve almost overlaps with the reference one.

a. Role of σ skeleton. In the following calculations, the role of the bonding and antibonding σ_{C-H} and the σ_{C-C} orbitals is analysed by assigning them to zone 3 and performing various tests: (1) these orbitals are frozen (equivalent to $Th_1 = Th_2 = \infty$ in zone 3, with maximal interzone thresholds). This corresponds to the less time consuming calculation; (2) keeping the thresholds on the two other zones unchanged, the thresholds are lowered for the σ zone (i.e., zone 3) from $Th_1 = 0.1$ a.u./$Th_2 = 0.03$ a.u. until $Th_1 = 0.01$ a.u./$Th_2 = 0.003$ a.u. (Figure 9). The first case, with the σ skeleton frozen, gives the worst result. The deviation is rather small (0.9 kcal mol^{-1}) but represents 30% of the total interaction energy taken as the difference between placing the H_2 molecule far from the nanotube and placing it at the center of the fragment.
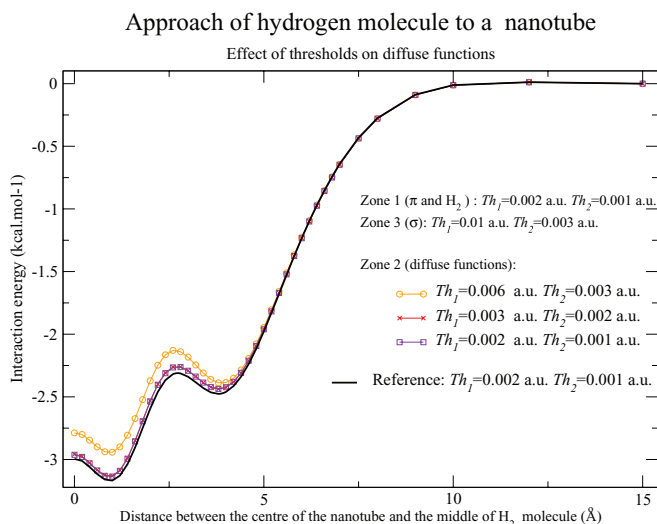
FIG. 10. H$_2$ approach to a nanotube fragment. Taking account of the "diffuse" functions through variation of thresholds.

All potential energy curves present two minima (0.99 Å and 3.80 Å) and two maxima (at 0.0 Å and 2.75 Å). When the H$_2$ molecule is approaching the edge of the nanotube, the potential energy curve presents a minimum (3.80 Å). It corresponds to the position where the first hydrogen atom of the H$_2$ molecule reaches the dangling hydrogens ring of the nanotube. This minimum is not mainly due to dispersive effects since it exists also for the frozen σ zone. However, only the lowest thresholds (0.01/0.003 a.u.) give a correct interaction energy and a correct position of the minimum compared to the reference. The second extremum encountered is then a maximum, at 2.75 Å, when H$_2$ reaches the ring of C–H bonds. The position of this maximum is found, as previously, with an error of 0.15 Å for all the higher thresholds and only the lowest ones give correct result (Table V). The same observation can be done for the interaction energy (Table VI). When

TABLE IV. H$_2$ approach to a nanotube. Dimensions of the calculations as a function of Th_1 / Th_2 / Th_i (a.u.) thresholds. For zone 1 (π, H$_2$), the same thresholds were used for all calculations: $Th_1 = 0.002$ a.u., $Th_2 = 0.001$ a.u., and $Th_i = 0.0001$ a.u.

Thresholds on zone 2 (a.u.) (diffuse)	Thresholds on zone 3 (a.u.) (σ squeleton)	CPU/it (s)	CI size
σ squeleton (Figure 9)			
0.002/0.001/0.0003	Frozen	210	1 065 969
0.002/0.001/0.0003	0.1/0.03/0.001	211	1 066 368
0.002/0.001/0.0003	0.03/0.01/0.001	389	1 715 232
0.002/0.001/0.0003	0.02/0.006/0.001	419	1 937 824
0.002/0.001/0.0003	0.01/0.003/0.001	733	3 138 426
Diffuse (Figure 10)			
0.006/0.003/0.0003	0.01/0.003/0.001	453	2 214 659
0.003/0.002/0.0003	0.01/0.003/0.001	598	2 756 337
0.002/0.001/0.0003	0.01/0.003/0.001	733	3 138 426
Reference			
	0.002/0.001/0.0001	2248	5 767 286
Non selected SDCI			
		X	2 696 779 081

TABLE V. H$_2$ approach to a nanotube. Position of the extrema (Å) as a function of the thresholds of the zone 3 (σ skeleton).

	Th_1/Th_2 (a.u.)					
	Frozen	0.1/0.03	0.03/0.01	0.02/0.006	0.01/0.003	Reference
Min1	0.99	0.99	0.99	0.99	0.99	0.99
Min2	4.19	3.99	3.99	3.99	3.80	3.80
Max	2.60	2.60	2.60	2.60	2.79	2.75

the H$_2$ molecule is near the centre of the nanotube (0.0 Å and 0.99 Å), all thresholds give the correct distance for the extrema (Table V). The interaction energy is not so satisfactory (Table VI): the error is larger than 10% except for the lowest threshold ($Th_1 = 0.01$ a.u., $Th_2 = 0.003$ a.u.) for which it is only 1%.

One can conclude that the σ skeleton cannot be neglected, and the more convenient thresholds of zone 3 are around $Th_1 = 0.01$ a.u., $Th_2 = 0.003$ a.u. Note that (i) these values are larger than expected; (ii) these are much higher values that those selected for zones 1 and 2. Using these values, one takes into account rather long range interactions between π and diffuse orbitals, and only very short range interactions in the σ system. For the interaction between the two sets (π and diffuse with σ), average thresholds are used, so that this interaction is treated at a higher level than the σ–σ interactions.

b. Zone 2, role of "diffuse" orbitals. In a further step the thresholds of zone 2, containing the "diffuse" 3s and 3p atomic orbitals of the carbon atoms, are modified, while the thresholds of the two other zones remain unchanged ($Th_1 = 0.002$ a.u. and $Th_2 = 0.001$ a.u. for zone 1 and $Th_1 = 0.01$ a.u. and $Th_2 = 0.003$ a.u. for zone 3). The results are depicted in Figure 10. Only the highest thresholds ($Th_1 = 0.006$ a.u. and $Th_2 = 0.003$ a.u.) give a non-negligible error on the interaction energy. In all cases, the positions of the extrema are correct according to the reference calculation.

In conclusion, we can separate π, diffuse and finally σ orbitals into three zones treating them at decreasing levels of precision and to reach a precision that can be compared to the reference calculation. This allows one to find the correct

TABLE VI. H$_2$ approach to a nanotube. Interaction energy at the extrema (kcal mol^{-1})) as a function of the thresholds of the zone 3 (σ skeleton).

	Th_1/Th_2 (a.u.)					
	Frozen	0.1/0.03	0.03/0.01	0.02/0.006	0.01/0.003	Reference
Min1 at 0.99 Å	−2.30	−2.60	−2.76	−2.86	−3.14	−3.17
Min2 at 3.80 Å	−1.98	−2.15	−2.23	−2.30	−2.44	−2.48
0.0 Å	−2.11	−2.41	−2.60	−2.71	−2.96	−2.995
Max at 2.75 Å	−1.46	−1.83	−1.93	−2.03	−2.26	−2.31

position and depth of the wells, as well as the barriers. For zone 1, the thresholds used are $Th_1 = 0.002$ a.u., $Th_2 = 0.001$ a.u. The adequate highest thresholds finally found for each the other zones are: (i) zone 2: $Th_1 = 0.003$ a.u., $Th_2 = 0.002$ and (ii) zone 3: $Th_1 = 0.01$ a.u., $Th_2 = 0.003$. With these values, the dimension of the CI matrix is reduced by half while the CPU time is only 26% of the reference calculation. One can notice that the dimension of the reference calculation is only 0.2% of the nonselected SDCI one (more than 2×10^9 of determinants).

SCHEME 1.

C. Cu_4O_4 cubane magnetic system

The third test concerns a tetranuclear Cu(II) complex with an open cubane-like Cu_4O_4 core structure, with four short and two long Cu–Cu distances, synthesized by Shi Tan *et al.*[40] The system of formula $[Cu_4(hpda)_4][ClO_4]$ H_2O with hpda = N-(2-hydroxyethyl)-1,3-propane-diamine (Figure 11) has been characterized as having a magnetic ground state. The magnetic coupling constants of this compound has been recently evaluated by some of us by means of the EXSCI approach and its susceptibility vs temperature curve simulated with the parameters resulting from these calculations.[41] While all previous EXSCI calculations have been done by imposing the same thresholds to the entire molecule; here we are going to explore the impact of the zones approach on the amplitude of the extracted magnetic terms. All the details regarding the procedure employed to extract these parameters as well as the
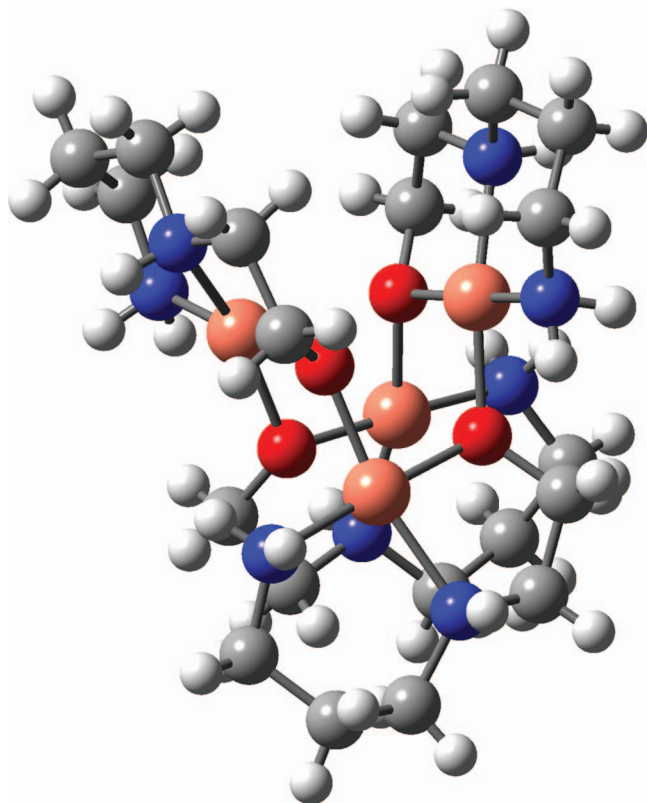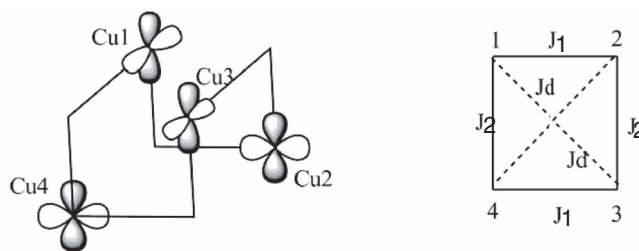

FIG. 11. $[Cu_4(hpda)_4][ClO_4]$ H_2O with hpda = N-(2-hydroxyethyl)-1,3-propane-diamine molecule. Pink, red, blue, gray, and white balls represent Cu, O, N, C, and H atoms, respectively.

discussion and analysis of the so-obtained results are provided in Ref. 41. Only a brief recall of the main features of the system and the procedure is given below.

Each Cu(II) atom in the complex presents a slightly distorted square-planar coordination with the two nearest O and N atoms and it bears one unpaired electron on the $3dx^2\text{-}y^2$ orbital, in agreement with the square-planar ligand field. Due to the co-ordination imposed by the hpda ligand, the $3dx^2\text{-}y^2$ orbitals on two close Cu atoms (short Cu–Cu distances) are practically orthogonal, while for two distant Cu atoms (long Cu–Cu distances) the $3dx^2\text{-}y^2$ orbitals are placed on parallel planes, but shifted along the x (or y) axis (Scheme 1).

The magnetic susceptibility χ vs T curve has been fitted by Shi Tan *et al.*[40] to a Heisenberg model of the form $H = -\sum J_{ij} S_i S_j$, with only two coupling constants $J = 89.8$ cm^{-1} ($J = J_{12} = J_{23} = J_{34} = J_{14}$) and $J_d = -32.6$ cm^{-1} ($J_d = J_{24} = J_{13}$) (see Scheme 1). While the leading ferromagnetic J interaction is consistent with the relative orientation of the Cu $3dx^2\text{-}y^2$ orbitals in two bridged Cu atoms (Scheme 1), the proposed large antiferromagnetic nature of the coupling between unbridged Cu atoms J_d is inconsistent with the geometry of the system, as in fact, it has been shown by our calculations.

The extraction of the set of parameters, both at experimental and theoretical level, starts by defining a spin model, i.e., by deciding the interactions playing a role in the magnetic behaviour of the system. In complete generality, three different types of magnetic interactions can be distinguished in the cubane: the nearest and next-nearest neighbours J and J_d interactions, which are two-body terms, and the four-body terms, J_r, which produce the simultaneous permutation of the four spin on the cubane. Taking into account the C_2 symmetry of the system, some simplifications can be done. In fact, there are only two different first-neighbour interactions J_1 ($J_1 = J_{12} = J_{34}$) and J_2 ($J_2 = J_{23} = J_{14}$) and two next-nearest neighbour interactions J_{d1} ($J_{d1} = J_{13}$) and J_{d2} ($J_{d2} = J_{24}$). There are three different four-body terms, J_{r1}, J_{r2}, and J_{r3} (see Ref. 41), all of them are expected to be small, but they can be evaluated without any additional cost. The extraction of this set of parameters, without any *a priori* approximation, is a laborious task, which requires the evaluation of the six low-lying magnetic states of the system (the quintet state (Q), three triplet states (T$_1$, T$_2$, T$_3$) and two singlet states (S$_1$, S$_2$)), and their mapping onto the eigenvalues of the model spin Hamiltonian:

$\lvert a\bar{b}c\bar{d}\rvert$	$\lvert \bar{a}b\bar{c}d\rvert$	$\lvert ab\bar{c}\bar{d}\rvert$	$\lvert \bar{a}\bar{b}cd\rvert$	$\lvert a\bar{b}\bar{c}d\rvert$	$\lvert \bar{a}bc\bar{d}\rvert$
$J_1 + J_2$	$-J_{r1}/2$	$-J_2/2 + A$	$-J_2/2 + A$	$-J_1/2 + B$	$-J_1/2 + B$
	$J_1 + J_2$	$-J_2/2 + A$	$-J_2/2 + A$	$-J_1/2 + B$	$-J_1/2 + B$
		$J_2 + (J_{d1} + J_{d2})/2$	$-J_{r2}/2$	$-J_{d2}/2 + C$	$-J_{d1}/2 + C$
			$J_2 + (J_{d1} + J_{d2})/2$	$-J_{d1}/2 + C$	$-J_{d2}/2 + C$
				$J_1 + (J_{d1} + J_{d2})/2$	$-J_{r3}/2$
					$J_1 + (J_{d1} + J_{d2})/2$

where $\{a,b,c,d\}$ are the four magnetic Cu $3dx^2$-y^2 orbitals and the $S_z = 0$ determinants $\{\lvert a\bar{b}c\bar{d}\rvert, \lvert \bar{a}b\bar{c}d\rvert, \lvert ab\bar{c}\bar{d}\rvert, \lvert \bar{a}\bar{b}cd\rvert, \lvert a\bar{b}\bar{c}d\rvert, \lvert \bar{a}bc\bar{d}\rvert\}$ constitute the model space. This mapping fixes the amplitudes of the integrals appearing in the model spin Hamiltonian and allows one to verify whether non-negligible additional interactions are present.

The energy of the magnetic states has been evaluated by means of DDCI calculations, using the EXSCI methodology; otherwise the diagonalization of the full DDCI space is completely unfeasible, since it contains more than 870×10^6 of determinants. The active space is composed of 4 electrons in the 4 magnetic orbitals. It is used to obtain the localized CASSCF orbitals and also for the MR-DDCI calculations. Several tests have been performed to analyze the impact of the EXSCI selection criteria on the values of the magnetic terms, the results being collected in Table VII. Tests A and B apply the same thresholds on the whole molecule, while the rest distinguish different regions as explained in Sec. II B.

Table VII shows the different selection thresholds (Th_1, Th_2, and Th_i), the CPU time (in minutes) per root and iteration of the Davidson diagonalization, the number of determinants included in the CI matrix and the resulting two-body magnetic coupling terms. Regarding the four-body terms, they are null as expected for all the tests considered.

For test A we have dealt with around 100×10^6 of determinants once that the excitations that do not match the selection criteria ($Th_1 = 10^{-3}$ a.u., $Th_2 = 10^{-4}$ a.u., and $Th_i = 10^{-4}$ a.u.) have been eliminated of the DDCI space. The first-neighbour interactions J_1 and J_2 are both ferromagnetic with close values of 24 cm^{-1} and 29 cm^{-1}, respectively, which agree with the relative Cu–Cu distance (larger for J_1 than J_2). The second-neighbour J_{d1} and J_{d2} interactions are also ferromagnetic although with a negligible absolute value.

Increasing the number of determinants in the CI space by diminishing the Th_1 and Th_2 thresholds as in test B practically does not affect the amplitude of any of these terms. As discussed in Ref. 36, the absolute values of these terms are slightly enhanced when size-consistency errors are corrected, but this is a point that we are not going to address here.

Now it is possible to use the zone-based selection to (i) reduce the computational cost (tests C, D, and E) or (ii) improve the description of the active zone diminishing the Th_1 threshold (tests F, G, and H).

For tests C, D, and E a reduction of the CPU time is obtained by using the threshold values of test A only for the active zone, and increasing the threshold values for the rest of the molecule. This produces a significant reduction of the number of determinants involved in the CI space, and consequently of the computational cost. The amplitudes of the magnetic terms are not sensitive to this topological selection and then values similar to those resulting from test A are obtained for all the parameters.

TABLE VII. Magnetic coupling terms (in cm^{-1}) for the Cu$_4$O$_4$ cubane obtained from EXSCI calculations with different thresholds, CPU time in minutes per iteration and root, and number of determinants in the CI space. In parenthesis, the corresponding percentage of the DDCI space included in the calculations is shown. The full DDCI space contains 8743×10^6 determinants.

Test		Zones	$Th1$	$Th2$	Thi	N^o det	CPU	J_1	J_2	J_{d1}	J_{d2}
A	1	All	10^{-3}	10^{-4}	10^{-4}	99.9×10^6 (11%)	417	23.7	29.1	0.6	0.6
B	1	All	5×10^{-4}	0	10^{-3}	204.1×10^6 (23%)	536	25.1	31.6	0.2	0.6
C	1	Cu, O	10^{-3}	10^{-4}	10^{-4}	54.7×10^6 (6%)	142	26.0	31.4	0.8	0.7
	2	Rest	10^{-2}	10^{-3}	10^{-3}						
D	1	Cu, O, N	10^{-3}	10^{-4}	10^{-4}	95.6×10^6 (10%)	423	25.0	29.9	0.7	0.6
	2	Rest	10^{-2}	10^{-3}	10^{-3}						
E	1	Cu,O	10^{-3}	10^{-4}	10^{-4}	53.6×10^6 (6.1%)	117	25.9	31.5	0.8	0.7
	2	N	10^{-2}	10^{-4}	10^{-3}						
	3	Rest	10^{-1}	10^{-3}	10^{-2}						
F	1	Cu,O	10^{-4}	10^{-4}	10^{-4}	262.9×10^6 (30%)	1705	24.9	30.9	0.6	0.9
	2	Rest	10^{-3}	10^{-3}	10^{-3}						
G	1	Cu,O	10^{-4}	10^{-4}	10^{-4}	230.1×10^6 (26.3%)	1249	25.3	31.6	0.6	1.0
	2	C1-C4,N2,N4	10^{-3}	10^{-3}	10^{-3}						
	3	Rest	10^{-2}	10^{-2}	10^{-2}						
H	1	Cu,O	10^{-4}	10^{-4}	10^{-4}	173.7×10^6 (19.9%)	747	26.5	33.6	0.8	1.1
	2	Rest	10^{-2}	10^{-2}	10^{-2}						

Comparing C and D we see that including the nitrogen atoms on zone 1 does not change the amplitude of the terms, but increases substantially the computational times.

Tests F, G, and H represent an improvement of the description of the active zone with respect to test A, as far as a smaller threshold Th_1 value has been employed. This produces a dramatic increase of the size of the CI space and the CPU time (test F is particularly expensive), but practically any noticeable change on the parameters.

Since it is not possible to perform the calculation of the complete DDCI space, we can consider as effective magnetic parameters for this system those resulting for the largest available CI space (test F):

$$J_1 = 24.9 \text{ cm}^{-1}, \quad J_2 = 30.9 \text{ cm}^{-1}, \quad J_{d1} = 0.6 \text{ cm}^{-1},$$

$$J_{d2} = 0.9 \text{ cm}^{-1}, \quad J_{r1} = J_{r2} = J_{r3} = 0.0 \text{ cm}^{-1},$$

which are in pretty good agreement with the less expensive set of parameters (test E).

IV. CONCLUSION

The new version of the quasi-N-scaling CI code based on the use of localised orbitals has been extended to the possibility of treating various regions of a molecule at different levels of precision. This improvement permits to address larger systems at the MRCI level. The method has been explored in a set of three different test studies.

The separation of the set of local orbitals into different zones can be made according to different criteria. A first criterion, the most intuitive one, is to divide the system in different geographical regions, considering, for example, that the active part of the molecule on one hand, and the ligands on the other hand play unequal roles. But it is also possible to distinguish orbitals in the same region of space with dissimilar importance, such as σ or π in one of the examples presented in the paper.

The test calculations presented above show that this approach reduces the computational cost for a comparable precision of the results, and then it is aimed to problems that could not be treated without the introduction of these approximations, at this level of calculation. Beyond this conclusion, another aspect – less tackled in this work – is to take benefit from the possibility to deal with various partitions as an analysis tool. For example, freezing the sigma core in the nanotube fragment and treating it at different levels of accuracy in a second step gives an idea of the importance of its particular role in the interaction studied in Sec. III. The possibility to divide a system into zones not only allows one to treat larger systems, it may be used as a theoretical analysis tool.

ACKNOWLEDGMENTS

[1] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).
[2] C. J. Cramer, J. R. Gour, A. Kinal, M. Włoch, P. Piecuch, A. R. M. Shahi, and L. Gagliardi, J. Phys. Chem. A **112**, 3754 (2008).
[3] P.-Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, and L. Gagliardi, J. Chem. Phys. **128** 204109 (2008).
[4] J. Zapata-Rivera, R. Caballol, and C. J. Calzado, J. Comput. Chem. **32**, 1144 (2011).
[5] J. Miralles, J. P. Daudey, and R. Caballol, Chem. Phys. Lett. **198**, 555 (1992).
[6] J. Miralles, O. Castell, R. Caballol, and J.-P. Malrieu, Chem. Phys. **172**, 33 (1993).
[7] N. Queralt, D. Taratiel, C. de Graaf, R. Caballol, R. Cimiraglia, and C. Angelin, J. Comput. Chem. **29**, 994 (2008).
[8] R. Gdanitz and R. Ahlrichs, Chem. Phys. Lett. **143**, 413 (1988).
[9] P. J. A. Ruttink, J. H. van Lenthe, R. Zwaans, and C. Groenenboom, J. Chem. Phys. **94**, 7212 (1991).
[10] A. Venkatnathan, A. B. Szilva, D. Walter, R. J. Gdanitz, and E. A. Carter, J. Chem. Phys. **120**, 1693 (2004).
[11] (a) P. G. Szalay and R. J. Bartlett, Chem. Phys. Lett. **214**, 481 (1993); (b) J. Chem. Phys. **103**, 3600 (1995); (c) J. Phys. Chem. **100**, 6288 (1996).
[12] N. Ben Amor, D. Maynau, J. P. Malrieu, and A. Monari, J. Chem. Phys. **129**, 64112 (2008).
[13] A. El Azahary, G. Rauhut, P. Pulay, and H.-J. Werner, J. Chem. Phys. **108**, 5185 (1998).
[14] M. Schütz and H.-J. Werner, J. Chem. Phys. **113**, 9986 (2000).
[15] G. E. Scuseria and P. Y. Ayala, J. Chem. Phys. **111**, 8330 (1999).
[16] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
[17] S. Li, J. Ma, and Y. Jiang, J. Comput. Chem. **23**, 237 (2002).
[18] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[19] T. S. Chwee, A. B. Szilva, R. Lindh, and E. Carter, J. Chem. Phys. **128**, 224106 (2008).
[20] T. S. Chwee and E. Carter, J. Chem. Phys. **132**, 74104 (2010).
[21] T. S. Chwee and E. Carter, Mol. Phys. **108**, 2519 (2010).
[22] T. S. Chwee and E. A. Carter, J. Chem. Theory Comput. **7**, 103 (2011).
[23] B. Bories, D. Maynau, and M. L. Bonnet, J. Comput. Chem. **28**, 632 (2007).
[24] N. Ben Amor, F. Bessac, S. Hoyau, and D. Maynau, J. Chem. Phys. **135**, 014101 (2011).
[25] D. Maynau, S. Evangelisti, N. Guihery, J. P. Malrieu, and C. J. Calzado, J. Chem. Phys. **116**, 10060 (2002).
[26] S. F. Boys, in *Quantum Theory of Atoms, Molecules and the Solid State*, edited by P. O. Löwdin (Academic, New York, 1966), p. 253.
[27] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
[28] C. Angeli, S. Evangelisti, R. Cimiraglia, and D. Maynau, J. Chem. Phys. **117**, 10525 (2002).
[29] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).
[30] F. Bessac and S. Hoyau, Comput. Theor. Chem. **966**, 284–298 (2011).
[31] M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. Defrees, and J. A. Pople, J. Chem. Phys. **77**, 3654 (1982).
[32] J. P. Blaudeau, M. P. McGrath, L. A. Curtiss, and L. Radom, J. Chem. Phys. **107**, 5016 (1997).
[33] J. S. Arellano, L. M. Molina, A. Rubio, and A. J. Alonso, J. Chem. Phys. **112**, 8114–8119 (2000).
[34] G. Mpourmpakis, G. E. Froudakis, G. P. Lithoxoos, and J. Samios, J. Chem. Phys. **126**, 144704 (2007).
[35] G. P. Miller, J. Kintigh, E. Kin, P. F. Weck, S. Berber, and D. Tománek, J. Am. Chem. Soc. **130**, 2296–2303 (2008).
[36] F. F. Contreras-Torres, A. F. Jalbout, I. Jiménez-Fabián, O. F. Amelines, and V. A. Basiuk, J. Phys. Chem. C **112**, 2736–2742 (2008).
[37] A. Lakhlifi and J. P. Killingbeck, Surf. Sci. **604**, 38–46 (2010).
[38] H. Kyakuno, K. Matsuda, H. Yahiro, Y. Inami, T. Fukuoka, Y. Miyata, K. Yanagi, Y. Maniwa, H. Kataura, T. Saito, M. Yumura, and S. Iijima, J. Chem. Phys. **134**, 244501 (2011).
[39] K. Pierloot, B. Dumez, P.-O. Widmark, and B. O. Roos, Theor. Chim. Acta **90**, 87 (1995).
[40] X. Shi Tan, Y. Fujii, R. Nubada, M. Mikuriya, and Y. Nakano, J. Chem. Soc. Dalton Trans. 2415–2416 (1999).
[41] C. J. Calzado and D. Maynau, J. Chem. Phys. **135**, 194704 (2011).