Multi-d-occupancy as an alternative definition for the double d-shell effect

Matheus Morato F. de Moraes* and Yuri Alexandre Aoto*

Center of Mathematics Computing and Cognition, Federal University of ABF (UFABC), Santo André

E-mail: matheusmorat@gmail.com; yuri.aoto@ufabc.edu.br

Abstract

Despite of the prevalence of first-row transition metal containing compounds in virtually all areas of chemistry, the accurate modeling of these systems is a known challenge for the theoretical chemistry community. Such a challenge is shown in a myriad of facets, among them: difficulties in defining ground states multiplicities; disagreement in the results from methods considered highly accurate; and convergence problems in calculations for excited states. These problems cause a scarcity of reliable theoretical data for transition metal containing systems. In this work we explore the double d-shell effect, that plagues and makes difficult the application of multireference methods to this type of systems. We propose an alternative definition for this effect based on the mixing among d-occupancy configurations, or the multi-d-occupancy character of the wave function. Moreover, we present a protocol able to include this effect in multireference calculations using an active space smaller than the one usually used in the literature. A molybdenum-copper model system and its copper subsystem are used as example study cases, in particular, the molybdenum-copper charge transfer of the former and the electronic affinity of the latter. We have shown that our alternative definition can

be used to analyze their reference wave functions qualitatively. Based on this qualitative description it is possible to optimize an active space without a second d-shell, able to obtain relative energies accurately. Seeing the double d-shell effect through the lens of a multi-d-occupancy character it is possible to correctly describe the wave function, improve the accuracy of the relative energies, and cheapen the multireference calculations. That way, we believe that this alternative definition has the potential to improve the modeling of first-row transition metal containing compounds both for their ground and excited electronic structure.

Introduction

Photomagnetic switches form a class of compounds that upon irradiation drastically change their magnetic properties.[1] Although these compounds incite large interest in the materials field and several technological applications are foreseen,[2,3] computational simulations of these compounds lag behind. Only a hand full of data regarding the electronic structure of these systems can be found in the literature. Up to our knowledge, besides some experimental/theoretical hybrid works that use density functional theory,[4,5] there are only three works that employ theoretical methods considered highly-accurate,[6-8] conducted by Graaf et al. using a model system of the magnetic photoswitch synthesized by Herrera et al.[1]

One of the limitations when describing such series of compounds comes from the large system size. This is particularly the case when wave function based methods are necessary, due to their unfavorable scaling with the number of electrons and basis set.[9,10] However, these methods are often necessary to model the mechanisms where magnetic photoswitching processes take place. They can occur via two main routes: a spin-flipping transition (also known as Light-Induced Excited Spin-State Trapping, LIESST[11]) and/or a metal-metal charge transfer transition (MMCT).[1,3,5,12] These transition types are known to be difficult to calculate, either using density functional theory or wave function methods. However, spin-flipping processes have received substantially more attention than charge transfer processes, due to their relevance in many controversies regarding the ground state multiplicity of metal containing compounds (e.g. for the triplet/quintet states of iron-porphyrin[13-15]).[16-19]

As both mechanisms, LIESST and MMCT, are relevant for some photomagnetic switching properties,[12] they need to be modelled with the same accuracy if a reliable picture of the excited electronic states is desired. Both transitions are exceptionally expensive to model due to the large

number of low-lying energy states and orbitals. Moreover, they suffer, as all first-row transition metal containing compounds, from the *double d-shell effect* (or *3d double-shell effect*), that plagues multireference wave function based methods. This effect was first described by Andersson and Roos in 1992,[20] with the term coined in 1995.[21] To include this effect a double d-shell active space approach was developed to describe transition metal containing compounds based in two major works, from Fischer in 1977,[22] and Botch et al. in 1981.[23] Influenced by these works, two distinct, but related, definitions for the double d-shell effect can be formulated.

The first is the most applied and utilitarian definition that follows the general premise: *a system has a double d-shell effect if it requires two sets of d-orbitals in the active space to be correctly modeled.* Such definition has a strong resemblance to Fischer's work, where relative energies with good agreement with experimental values can be obtained by writting the atomic states wave functions as a linear combination of $3d$ to $4d$ excitations.[22] Extending this idea to multireference methods, it has been shown that for a better accuracy these $3d$ to $4d$ excitations are also necessary in the reference wave function. A complete explanation on the active space choice based on the ligand interaction with the metal orbitals can be found in Chapter III.B of reference.[24] A large drawback of this approach is that the minimum active space required for transition metal containing compounds must include at least 10 active orbitals, named by Pierloot in reference[25] as the "basic ten". In our point of view, however, this definition for the double d-shell effect has two problems. First, following this definition, systems with the double d-shell effect would always need a reference wave function based on large active spaces. Although such definition implies in a multireference character for the system, there are examples of systems that suffer from the double d-shell effect that can be better described by the canonical CCSD(T) method than by a single d-shell active space.[26,27] That way, a single-reference method can recover the double d-shell effect. Second, it suggests a single specific solution to account for this effect, without providing an explanation for its existence, what limits the development of alternative approaches.

The second usual definition for the double d-shell effect is grounded in Botch et al.[23] results. It states that the $3d$-electron correlation is dominated by a $3d \rightarrow 4d$ radial correlation.[28] In such a case, this effect should be accounted almost entirely by any post-CAS methods that considers $3d$ to $4d$ double excitations, what is not true even for small diatomic cases.[29] This inconsistency in this explanation is not ignored in the literature, for instance in Pierloot's words (reference[30] page 126): "[...] such an effect should be classified as dynamic rather than nondynamic, [...] this is one of the exceptional cases where an accurate description of dynamic correlation effects really benefits from a reoptimization of the orbitals involved, i.e., from a multireference treatment.". This justification of an exceptional type of dynamic correlation, that must be treated similarly as a nondynamic one, gives rise for more questions than answers. For instance, if an exceptional behavior exists in the first-row transition metals, what prevents it to affect heavier transition metals, such as $4d$ metals,

3

or other atoms as the lanthanides?

Using Graaf et al. molybdenum-copper model complex[6] as a study case, in this article we propose that there is an unified definition for the double d-shell effect that contemplates the results for both single and multireference methods: *The double d-shell effect is the mixing of two or more configurations with distinct d-occupancies in a wave function.* It is the wave function's multi-d-occupancy character. Similarly to the multireference character, the multi-d-occupancy character is necessary to qualitatively describe the wave function. However, the latter does not require a multi determinantal expansion, differently from the former, because the multi-d-occupancy character can be encoded in the molecular orbitals. This encoding was discussed in our previous work[31] and will be further explored in the present article. We point out that hints of the multi-d-occupancy character in the first-row transition metal containing compounds electronic structure have been present in the literature for a long time. For instance Chong et al. in 1986 explored the first-row transition metal hydrides ground states based on their $3d^n4s^2$ and $3d^{n+1}4s$ character mixing.[32] Moreover, Bauschlicher Jr. in 1988[33] has shown, for TiH, that when compared with full configuration interaction (FCI) results (for a basis set that contains three d-shells), CASSCF/MRCISD+Q calculations using a single d-shell active space have a better performance for the pure d^2-occupancy state than the single reference SCF/CPF. On the other hand, the single reference approach had better agreement with the FCI values for a mixed $3d^24s^2/3d^34s$ state.

Based on our alternative definition for the double d-shell effect, the apparent need of a double d-shell active space can be explained. As it is known since the late 70's, electronic configurations of first-row transition metal atoms with distinct d-occupancies lead to small variations of d-orbital sizes that drastically affect their energies.[22,34] Therefore, if one aims to minimize the total energy of two configurations with distinct d-occupancies, the space must be flexible enough to optimize d-orbitals with sizes optimal for both d-occupancies. Such a flexibility can be obtained with a double d-shell active space. However, for a qualitative description of wave functions, it is not necessary to minimize the total energy of all states, but only to have a set of states described with similar accuracy. This is, for instance, the primary idea of the state-averaged complete active space self-consisten field method (SA-CASSCF).[35,36] Thus, to consider the multi-d-occupancy character (or double d-shell effect) it is only necessary that the configurations with distinct d-occupancies are described with similar accuracy, not necessarily via individual energy minima. We have shown that such similar accuracy is possible using a single d-shell active space if only two d-occupancies are necessary to describe the system.[29,31]

Similar results were obtained in previous calculations for the molybdenum-copper model system,[8] where the authors have shown that a single d-shell active space can be used to obtain good results, while the MMCT state could not been found using a double d-shell active space. However, a deeper explanation for the quality of these results is lacking. For instance, it is unclear whether

4

the results rely on some sort of fortuitous error cancellation, or if their approach can be extended to other studies. In either case, it becomes apparent that the double d-shell effect may be treated, or at least the problems associated to it mitigated, without the second set of d-orbitals. This is consistent with our attempt to redefine the double d-shell effect without considering one specific solution.

In this article, we will present and discuss results for the molybdenum-copper model originally proposed by Graaf et al.[6] using our approach, the *partially fixed reference space* (PFRS).[31,37] It is a protocol for orbital optimization designed to impose that the d-orbitals converge to a size that is able to describe, simultaneously and with similar accuracy, configurations with distinct occupancies. This is made in a two-step procedure, where the ligand orbitals are optimized without allowing changes in the size of the metal d-orbitals, that have been previously optimized. This is in line with a recent series of articles that report good results when alternative orbital optimization schemes are used in a smaller CASCI expansion, when compared to the usual SA-CASSCF approach, as compiled by Levine et al.[38] In particular it has potential to fill the gap that exists for orbitals to model charge transfer processes in transition metal containing compounds, as pointed out by the mentioned compilation work. We will show that this procedure, together with our proposed definition for the double d-shell effect, can explain Graaf et al. results,[8] based on the multi-d-occupancy character of the wave function.[31,37] For this purpose, as the double d-shell effect is substantially larger for the first-row transition metals,[30] we will first focus on the electron affinity of the copper subsystem. This reduces the computational cost but maintais the difficulties due to the $3d$ transition metal. The molybdenum-copper system will then be analysed and discussed under the light of what has been learned from the copper subsystem.

Methods

Despite of the large flexibility and applicability of the standard SA-CASSCF method, considered a non-black-box method,[38] in some situations the user control over the orbitals convergence is quite limited. For instance, in larger molecules often only a hand full of active orbitals can be included in the calculations, due to computational limitations. Such a small active space bias the orbitals convergence towards the ones that optimize states, and/or configurations, with the smallest dynamic correlation energies. Good examples are the isolated transition metals that, regardless of the ground state, have a tendency to converge to the state with the lowest d-occupancy configurations, that is, $4s^2 3d^n$, when a single reference wave function[23] or single d-shell (inner-valence) active space is used.[39]

A possibility to increase control over orbital convergence, as we have previously proposed,[31,37]

5

is to divide the orbitals in groups, to be optimized in different steps. Although this division may be arbitrary, we propose it to be between metal centers and ligands. The flowchart in Figure 1 summarizes the proposed optimization. It starts with the division of the atoms between metal centers and ligands. Then the isolated metal orbitals are optimized, as shown in box I. A single set of d-orbitals is optimized in a way that it is able to describe two distinct d-occupancies with similar accuracy. The metal basis set is then tailored by removing the optimized virtual orbitals. This is made to prevent the d-orbital size to be altered during the following steps, when the ligands molecular orbitals will be optimized. The isolated ligands, that is, without metal atoms, may have their orbitals pre-optimized as well, and then merged with the tailored metal ones, as described in the box II. Although in some cases this merged initial guess may improve the molecular orbitals convergence, other automatic initial guesses can be employed. The molecular orbitals are optimized, but with the restriction that the size of the metal orbitals are kept as previously optimized because metal virtual orbitals are absent. This set of molecular orbitals, optimized with restrictions, is then projected onto the complete basis set of the metal, that is, the untailored one, to recover the complete virtual space. Finally, these orbitals are used to form a CASCI reference wave function for post-CASCI calculations, as shown in box III. Once this optimization fixes the orbitals centered at the metal and, consequently, part of the spawned reference space, independently of the system geometry, it was named *partially fixed reference space* (PFRS).[37]

The main difference, and advantage, of our proposed approach in comparison with the usual SA-CASSCF is the need to define two weights for the states averaging. The first weight, denoted by m, is associated to the steps in box I of Figure 1 and can be used to control the weight of different states of the metal center when optimizing its orbitals. Since the electronic states of the isolated metals have well-defined d-occupancy, in this step the user can optimize the metal orbitals, in particular its d-orbitals, in accordance to the d-configurations that are relevant to the system/process to be described. The second weight, denoted by w and associated to the CASSCF step in box III, can be used to control the optimization of the ligands in the presence of the (already pre-optimized) metals. The removal of metal virtual orbitals in this step prevents the size change of d-orbitals, and they stay as coming from the initial optimization. Such a division enables one to study the importance of the weight for each group of orbitals for the process of interest. On the other hand, it increases the dependence of the user choice and, consequently, may result in unreliable values if the weights are not well chosen. An in depth discussion regarding the development, properties and limitations of this approach can be found in reference.[37]
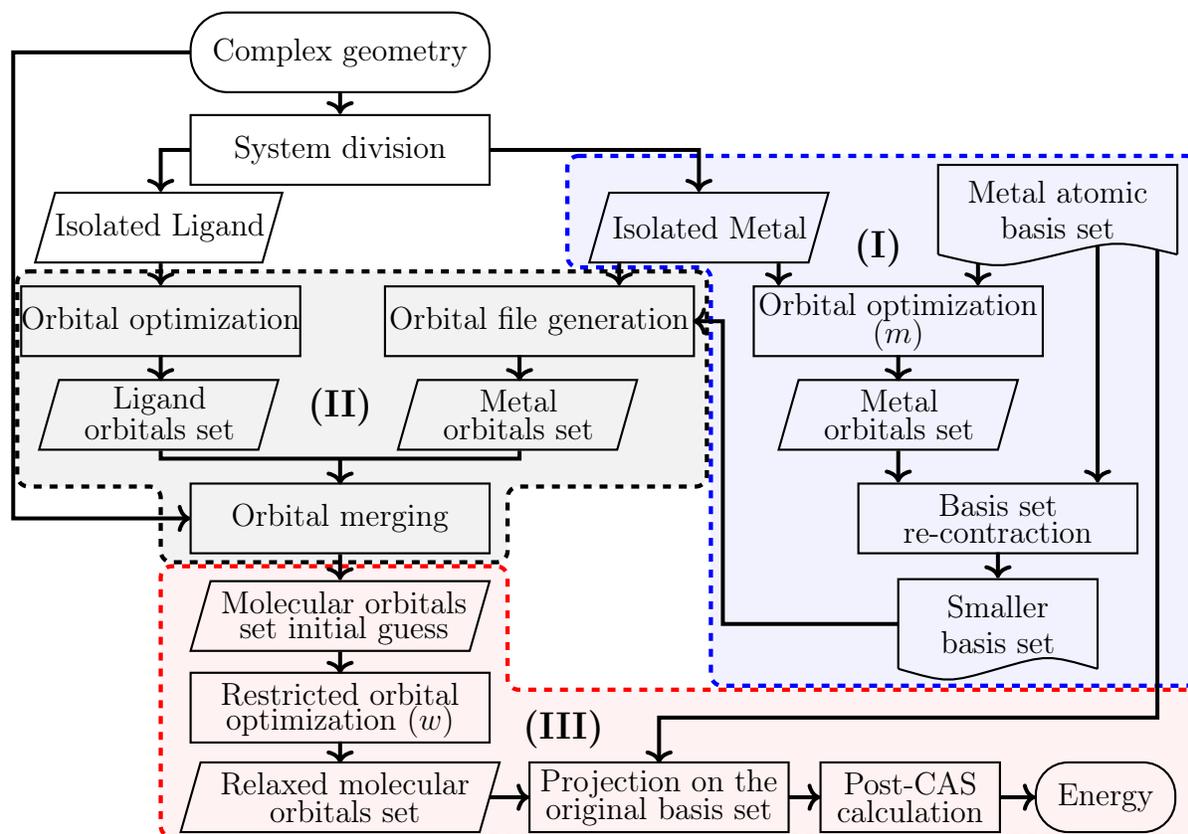
6

Figure 1: Flowchart describing the PFRS procedure. It assumes a system containing a single transition metal center for simplicity. In the case of multiple centers the rightmost branch starting from "Isolated Metal" (box I) is done for each center.

Technical details

All calculations were made using the MOLPRO package versions 19 and 21.[40–42] The molecular geometries are the ones described in reference[8] and reproduced in this article Support Information. The same basis set family used in reference[6,8] was used as well, namely def2-TZVPP for the metal atoms (Cu and Mo) and def2-SVP for the representative ones (H, C and N).[43] The active spaces for the copper subsystem and the molybdenum-copper were CAS(1,1) and CAS(3,3), respectively.

Post-CASCI calculations were performed using the n-electron valence perturbation theory of second order (NEVPT2)[44–46] level of theory, where core orbitals were not correlated ($1s$ for the representative atoms, from $1s$ to $3p$ for copper and $1s$ to $4p$ for molybdenum). For the molybdenum-copper model system, results from the partially contract version, pc-NEVPT2, presented in some punctual cases large total energy variations with the orbitals choice. On the other hand, due to the use of a single Slater determinant reference wave function, the strongly and partialy contracted versions of NEVPT2 for the copper subsystem result in the same total energies. Thus, although conclusions based on the pc-NEVPT2 results, after ignoring these non-physical values, are the same as for the sc-NEVPT2, the results for NEVPT2 calculations shown and discussed in this text are for the strongly contracted version, sc-NEVPT2. Results using both contractions can be found in the Support Information.

All fitting have been done using gnuplot version 5.4.[47] The fitting residual here considered is the square root of the weighted sum-of-squares residual divided by the number of degrees of freedom. Therefore, the residuals have the same unit as the fitted function.

Results and discussion

The copper subsystem

To model the metal-metal charge transfer (MMCT) in the molybdenum-copper model we will divide it in two subsystems and explore them first individually, with focus on the copper containing one, $[Cu(NCH)N(CH_2CH_2N)_3]^{2+}$, that we will denote as S^{2+}. Therefore, we will look at the reduction process of the copper subsystem,

$$S^{2+} + e^- \rightarrow S^+ \,, \tag{1}$$

that corresponds to the electron gain of the MMCT process in the model system. We will thus calculate the electronic affinity of S^{2+}:

$$\Delta E_{ea} = E(S^{2+}) - E(S^+) \,, \tag{2}$$

8

where $E(X)$ is the ground state energy of specie X. The electronic structure calculations needed are relatively straightforward, because they are for ground states of species with different number of electrons. Thus, their wave functions do not mix, what facilitates the convergence and analysis.

Starting with the reduced specie (S^+), its Hartree-Fock (HF) wave function results in a d^{10}-occupancy configuration, associated to a Cu^+ system, that is a formal charge in accordance with the expected chemical intuition. Moreover, when the coupled-cluster wave function with singles and doubles $(CCSD)^{48}$ is calculated on top of this reference wave function, the single excitations amplitudes are very small. The T_1 and D_1 diagnostics are 0.017 and 0.063, respectively, and the largest single excitation amplitude is smaller than 0.05. This indicates that the ground state is fairly well described by a state-specific single Slater determinant wave function and by a single d-occupancy configuration.
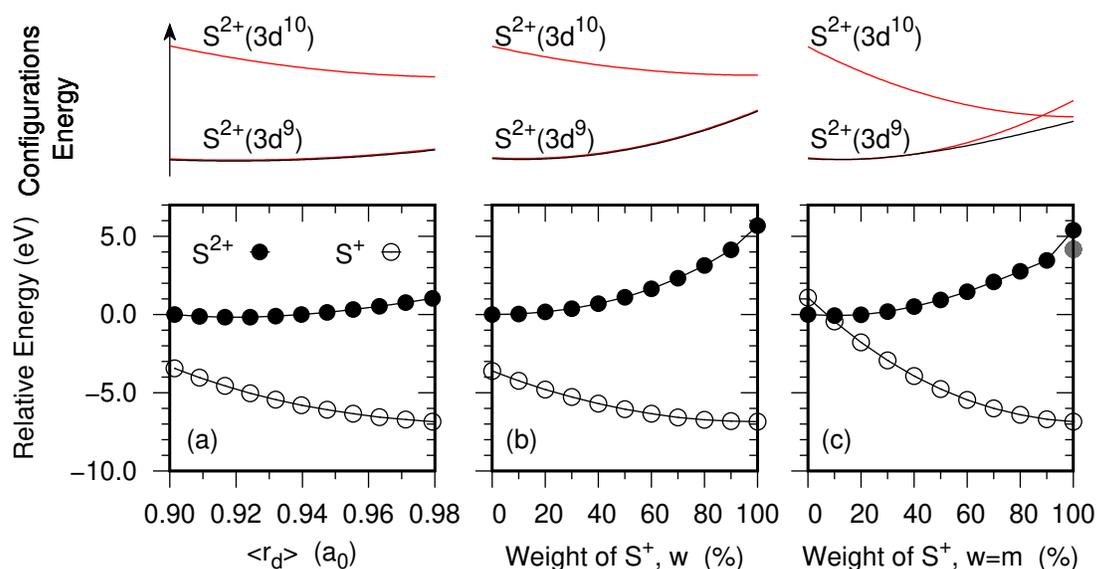

Figure 2: Total energy variation at the CASCI level of theory for the copper model subsystem, S^{2+}, ground state and its reduced form, S^+. **(a)** Varying the copper d-orbitals (m weight) and optimizing the ligand ones for each state. **(b)** Varying the weight between states (w weight), while the d-orbital size is fixed (0.90 a_0 for S^{2+} and 0.98 a_0 for S^+). **(c)** Using the same weights for both the metal and ligand orbitals ($m = w$). **At top:** Qualitative representation of the S^{2+} d-occupancy configurations relative energies (in red) and its ground state (in black).

Despite of the good description using optimal molecular orbitals, we can obtain a more complete picture of the electronic structure using our protocol to control the optimization of metal and ligand orbitals separately, and see how they affect the energy. The dependence of the S^+ ground state total energy (using a single Slater determinant wave function) with the optimization parameters is shown in Figure 2. We will analyse first the energy dependence with the d-orbital size, that

9

can be controlled with the weight used in the box I in the flowchart of Figure 1. Such a control of the d-orbital size is possible because there is an almost affine dependence between the weight and the orbital size, as shown in reference[37] in the neutral iron atom case and specifically for the copper case in Figure S1 at the Supporting Information. Thus, using our proposed optimization we can measure the subsystem energy variation from the orbitals having optimal size for a d^{10} configuration, of around 0.98 a_0, to optimal d^9 orbitals, with size of 0.90 a_0. Figure 2 a) shows that this rather small change in the d-orbital size impacts the S^+ ground state total energy by almost 3.5 eV.[i] This energy change can be qualitatively explained by the increase in the d-d repulsion with the orbital shrinkage, and it is analogous to what is known to happen for the atom. The optimal size of the d-orbital for isolated first-row transition metal atoms are strongly dependent on the state d-occupancy.[29,34,49] From scandium to copper neutral atoms, the magnitude of $3d$-orbital size variation decreases, whereas this size effect over the relative energy among distinct d-occupancies drastically increases. For instance, this leads to a variation in relative energies of up to 8.7 eV between the neutral copper atomic states with d^9 and d^{10}-occupancies[37] and almost 10 eV for the $Cu^{2+} \to Cu^+$ electronic affinity (as shown in Figure S1).

Similarly, using our approach it is possible to fix the $3d$-orbital, as the optimal for a d^{10} configuration, and vary the ligand orbitals by changing the relative weight (w) between the S^+ and S^{2+} ground states during the orbital optimization, that is represented in box III of Figure 1. Doing this we can measure how the interaction among ligand electrons and the copper center is impacted if the ligand orbitals are optimized around a Cu^{2+} center or a Cu^+ center. The energy variation of the S^+ ground state with the ligand orbitals is shown in Figure 2 b) and it is similar to the one seen for the d-orbital size change in Figure 2 a). Therefore, going from ligand orbitals shaped around the Cu^+ center to those shaped around Cu^{2+} provokes the same energy variation on the S^+ ground state as changing the internal metal orbitals from an optimal $3d^9$ to an optimal $3d^{10}$.

Lastly, because these energy variations arise from distinct sets of orbital, their interaction must be considered as well. That way, varying both ligand and metal orbitals the energy change is of around 8.0 eV, Figure 2 c), that is almost 1 eV larger than the sum of the energy variation from ligands and metal. Thus, these energy variations basically states the obvious, an optimal set of orbitals minimizes the energy in an energy minimization-based method. However, the important point here is: even with the destabilization of the configuration via the orbitals change there is no qualitative change in the wave function. That is, regardless of how well this state is described by the orbitals, its reference wave function remains as a pure d^{10}-occupancy state. Therefore, in agreement with the CCSD results, we can define this state to be of a single reference and single d-occupancy character. In other words, its double d-shell effect is small and a double d-shell active

---

[i]We reinforce here that the ligand orbitals are relaxed in the presence of the fixed d-orbitals size in this case via a state specific calculation.

space is not necessary to describe this state using the usual SA-CASSCF approach.

At first the S^{2+} ground state follows a similar behavior, but associated to a d^9 configuration. Its HF wave function converges to a d^9-occupancy, that is, a Cu^{2+} center, in agreement with what is chemically expected. However, when part of the correlation energy is added via the CCSD method, the painted picture changes. The final wave function has large contribution of single excitations, with T_1 and D_1 diagnostics of 0.023 and 0.18, respectively. Two amplitudes are particularly large, with absolute values of 0.20 and 0.05, and excite a ligand electron to the singly occupied d-orbital, originating d^{10} configurations. The only other amplitude higher than 0.05 in absolute value (considering both singles and doubles) is of a $3d{\rightarrow}3d$ excitation. That is, differently from the HF wave function, that describe the S^{2+} ground state as a pure d^9-occupancy state, the CCSD wave function can be better described as a multi-d-occupancy wave function composed of a d^9 and d^{10} mixture.

Following an analysis for the orbital variation similar as done for S^+, the S^{2+} dependence with the d-orbital size is relatively small (around 1 eV), as shown in Figure 2 a). We notice that such a result has precedent, for instance, that for the isolated iron atom the d^6-occupancy state has a smaller total energy dependence with the d-orbital size than the d^7 one.[37] On the other hand, the change in the ligand orbitals has a large impact over the S^{2+} ground state energy, as shown in Figure 2 b). With the variation of the ligand orbitals to better stabilize a Cu^+ center, the S^{2+} ground state is destabilized in almost 6 eV. This destabilization shows that the $Cu^{2+}(3d^9)$ configuration has a larger dependence on the optimization of ligand orbitals than on the metal orbitals. This contrasts to what happens for the d^{10} configuration, for which the dependence on ligands and metal orbitals are almost the same. Lastly, when both metal and ligand weights are gradually changed towards optimal orbitals for the $S^+(d^{10})$ ground state, the S^{2+} ground state is also highly destabilized as shown in Figure 2 c).[ii] However, if compared with the previous graph, (b), this destabilization is slightly smaller. In other words, using the least optimal ligand and metal orbitals for the $S^{2+}(3d^9)$ wave function leads to a lower total energy for this state than if a semi-optimal set, composed of an optimal d^9-orbital size with a set of ligand orbitals optimized around a Cu^+ center, is used.

Such a result may appear contradictory, as the interaction between two effects that destabilize the states leads to a relative stabilization. However, this stabilization is caused by the relative energies between distinct d-occupancies and a qualitative change in the state wave function. Figure 2 top schemes show, qualitatively, how the distinct d-occupancies configurations of the S^{2+} ground states are stabilized depending on the orbitals. For instance, in all cases, the orbitals change toward optimal for $S^+(3d^{10})$ destabilizes the expected $S^{2+}(3d^9)$ configuration, that is a Cu^{2+} interacting with a neutral ligand ($[Cu^{2+}L]^{2+}$), and stabilizes the more energetic $S^{2+}(3d^{10})$ configuration, a Cu^+

[ii]We point out that the discontinuity from 90% to 100% is because the wave function optimized at this point has all orbitals double occupied. That way, active and closed orbitals can mix without affect the S^+ energy, but it affects the energy of the S^{2+} ground state.

11

center interacting with a positively charged ligand ($[Cu^+L^+]^{2+}$). However, in the two first cases, Figure 2 a) and b), this reduction in the relative energy gap is not enough to cause a mixing between configurations. As a consequence, in both cases the final ground state wave function (black line in the qualitative scheme) is always a pure $S^{2+}(3d^9)$ configuration. On the other hand, when the effect of changing both groups of orbitals is considered, the gap is reduced at the point where there is a mixing between d-occupancies, what leads to an additional stabilization in graph (c). Such a mixing can be seen as the representation, in the multi-d-occupancy definition, of the backdonation effect recovered in calculations with the usual double d-shell active space approach.[25]
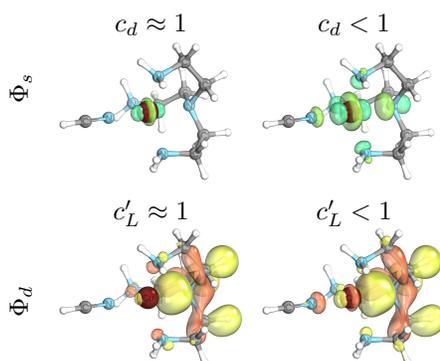

Figure 3: Copper subsystem molecular orbitals. The S^{2+} singly and one of the doubly occupied orbitals at top and bottom rows, respectively, as deffined in equations 3 and 4. At the left column, orbitals optimized for the S^{2+} ground state ($m = w = 0\%$) and, at the right one, an intermediate optimization ($m = w = 70\%$).

Although the CASCI results discussed up to this point are from single Slater determinant wave functions, they are able to recover the mixing between d-occupancies because it can be encoded within the S^{2+} singly occupied orbital. This orbital, denoted here as ϕ_s, is a mixture of the copper $3d_{z^2}$-orbital, in the CN bridge direction, with a mainly nitrogen ligand orbital ϕ_L (as shown in the top row of Figure 3):

$$\phi_s \approx c_d 3d_{z^2} - c_L \phi_L . \tag{3}$$

When the weight of S^+ is zero ($w = 0\%$), the contribution of ϕ_L is essentially zero. Thus, it is an essentially pure d orbital for the optimal orbitals of S^{2+}, leading to a $3d^{10}$ configuration for S^+ and a $3d^9$ configuration for S^{2+} for this choice of weight. However, when the weight of S^+ increases, as in the results shown in Figure 2 c), the c_L/c_d ratio also increases. As a balance, the doubly occupied orbital ϕ_d (Figure 3 bottom row),

$$\phi_d \approx c_d' 3d_{z^2} + c_L' \phi_L , \tag{4}$$

has an increased contribution of the $3d_{z^2}$ orbital. If we consider the localized orbitals $3d_{z^2}$ and ϕ_L to write the wave function, we can expand the S^{2+} single Slater determinant in a multi-d-occupancy

wave function:

$$|\{core\}\phi_d^2\phi_s\rangle \approx c_9|\{core\}\phi_L^2 3d_{z^2}\rangle + c_{10}|\{core\}\phi_L 3d_{z^2}^2\rangle, \tag{5}$$

where $\{core\}$ represents all other doubly occupied orbitals, including the other four $3d$ ones. The coefficients c_9 and c_{10}, that depend on the contribution of $3d_{z^2}$ and ϕ_L to the molecular orbitals, give the contribution of d^9 and d^{10} configurations to S^{2+}. As a consequence of this wave function expansion, the effect of the metal and ligand mixture in the singly occupied molecular orbital is that the S^{2+} is not described as a pure d^9 configuration, but has a multi-d-occupancy character.

To secure that such a mix in the singly occupied orbital, Equation (3), is not an artifact caused by the restriction of the active space, we increased the active space to also include the orbital describe in Equation (4), but without re-optimization. More specifically, a CAS(39,20), due to the limit of 20 active obitals in the code used. The result was the same single determinant wave function and, consequently, total energy. The only exception is the $m = w = 100\ \%$ case, were the additional active orbitals fixed the discontinuity of the total energy, whose energy is marked by the grey point in Figure 2 c).

Therefore, we describe the S^{2+} ground state as a single-reference system, but with a multi-d-occupancy character. Despite of this single-reference character, a single Slater determinant wave function does not converge to the correct d-occupancy mixing unless a state-average calculation that destabilizes the d^9-occupancy configuration and stabilizes the d^{10} one is used. Therefore, this approach does not optimize the states energy-wise, but allows the optimization of relative energies among distinct occupations. As a consequence, both (i) the *relative energies* among states with distinct d-occupancy profiles, and (ii) the *mixing among occupancies* within a given state, can be obtained with higher accuracy.

This conclusion is reinforced by and give us insights to explain the rather large T_1 and D_1 diagnostics, presented previously, that are however still below the loose threshold proposed for transition metal containing systems.[50] Although the S^{2+} HF reference wave function used in the CCSD calculation is biased toward the d^9-occupancy, such an orbital bias can be partially fixed via the single excitations, what leads to these high diagnostics. However, as observed previously at the MRCISD level for the iron atom excitation energy,[37] the bias fixed via single excitations can be propagated to the correlation energy via double excitations. In this case, the CCSD method is able to mitigate this bias onto the correlation energy due to the presence of triple excitations. These higher excitations recover the correlation for both the reference d^9-occupancy and d^{10}-occupancy configurations formed by single excitations, that is, for instance, missed at the CISD level of theory. Similar results can be seen for the $Cu^{2+} \rightarrow Cu^+$ transitions, as shown in Figure S2. Thus, despite of the need of further wave function analysis, the copper subsystem results similarities with the atomic ones indicate that these diagnostics may measure simultaneously both the multi-reference

13

and multi-d-occupancy characters indistinctly.

Because the multi-d-occupancy character can be measured, the proposed definition of the double d-shell effect is more quantitative than the usual definition, although it is still somewhat arbitrary as it depends on how the d-orbitals are localized. Here we will use an indirect, but straightforward, procedure to measure the d-occupancy mixing, based on the metal $3d$ population (d_{pop}). Starting from Equation (5) for the division of the wave function in the contribution of two occupancies, the $3d$ population may be written as

$$d_{pop} = 10c_{10}^2 + 9c_9^2,\tag{6}$$

as it has the contribution of 10 electrons from configuration d^{10} and 9 electrons from configuration d^9. Assuming that the wave function is normalized ($c_9^2 + c_{10}^2 = 1$), we have

$$\begin{cases} c_9^2 = 10 - d_{pop} \\ c_{10}^2 = d_{pop} - 9 \end{cases}.\tag{7}$$

From this point onward a myriad of population analysis may be used, but a comparison among them goes beyond the scope of this study. Therefore, we will discuss how the multi-d-occupancy character can be identified in the Mulliken population analysis. Despite of the many limitations and caveats associated to this measurement,[51–53] its cheapness, simplicity and wide implementation among quantum chemistry codes make any possible application, even if semi-quantitative, valuable for large systems. Figure 4 shows how the Mulliken d population varies with the w and m weights for the S^{2+} and S^+ ground states, at the CASCI level of theory using the PFRS approach. Starting with the latter, as previously discussed, the wave function has a pure d^{10}-occupancy character independently of the orbitals. However, the Mulliken population values falls-down to 9.65, that way, the value per se is not quantitatively reliable. On the other hand, we can trace a qualitative dependence between these values and the weights: the increase of the weight w leads to a reduction of the d_{pop}, as shown by any of the full lines; and when the weight m increases, d_{pop} also increases, as shown by the any vertical set of points. As a consequence, when these opposite effects are applied at the same time, as for $m = w$ (filled points), the d_{pop} variation is relatively small and between 10.0 and 9.9, in line with the previous discussion for this state.

For the S^{2+} wave function, using small metal weights, in particular $m = 0\%$ and 10%, the same qualitative pattern can be seen. That is, a decreasing in d_{pop} with w. With the increase of m, the w dependency shows an inflection point. This indicates that, differently from S^+, an additional effect that also increases the d_{pop} is present in the S^{2+} ground state. This additional effect is the d-occupancy mixing, that is even clear following the points where $m = w$. To reinforce this tendency, similar calculations were made using the intrinsic basis bonding analysis (IBBA) program,[40] based

14

on the intrinsic atomic orbitals (IBO), that has a smaller basis set dependency.[53] Although the values for both states overshoot the $3d$ population in around 0.3, the d_{pop} dependency follows the multi-d-occupancy character of each state. Numerically, for the S^+ ground state the variation with the weights is smaller than 0.1. On the other hand, for the S^{2+} the occupancy dependency is around 0.3 as shown in Figure S3.
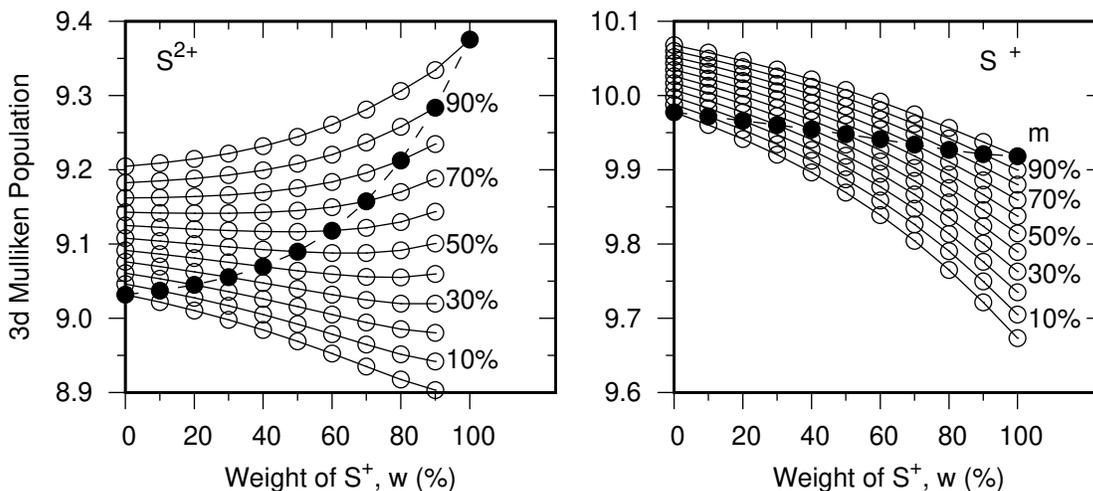

Figure 4: Mulliken population for the copper d orbital for S^{2+} and S^+ ground state wave functions varying the weights w and m during the PFRS approach. Full lines indicate points with the same m (marked in labels at the graphs) and the filled points the results for $m = w$.

The results up to this point reinforce that the multi-d-occupancy character can be taken into account in single d-shell calculations, and can be semi-qualitatively measured. However, this character is missed if orbitals are not properly optimized. Once the configurations with lower number of d-electrons have a smaller d-d correlation,[23] they tend to be more stabilized in calculations that do not include d-d correlation, in comparison to those configurations with higher d-occupancies. That way, HF or single d-shell CASSCF calculations are prone to converge toward the smaller d-occupancy configurations. This causes the low-lying states to preferentially converge to lower occupancy configurations, even if they have large contribution of higher occupancy configurations. This was previously explored by us for the cobalt monohydride diatomic system, for which the usual single d-shell CASSCF calculation converges to a pure d^7-occupancy, while methods with higher accuracy converge into almost pure d^8-occupancy ground state.[29] A more interesting example is the cobalt monofluoride diatomic system, where the consideration of the multi-d-occupancy character of the states dictates whether the ground state is triplet or quintet.[31]

15

Relative energy

Once we have an in depth description of the states total energies, the next step is to explore how the relative energy depends on the orbitals choice. Figure 5 shows the electron affinity dependence with both the metal and ligand weights at the CASCI and NEVPT2 levels of theory. The huge energy variation at the CASCI level indicates the large difference between optimal sets of orbitals for each state and, consequently, between their wave functions. That way, the closer to $(m, w) = (0, 0)$, the lower is the S^{2+} state total energy and the higher the S^+ one. However, the smaller is the d-occupancy mixing of the former. Moreover, due to the orbitals being closer to the optimal for a d^9-occupancy, the S^+ state energy is over-corrected at the NEVPT2 level of theory. This provokes an inverted behavior in how the electron affinity depends with the weights.
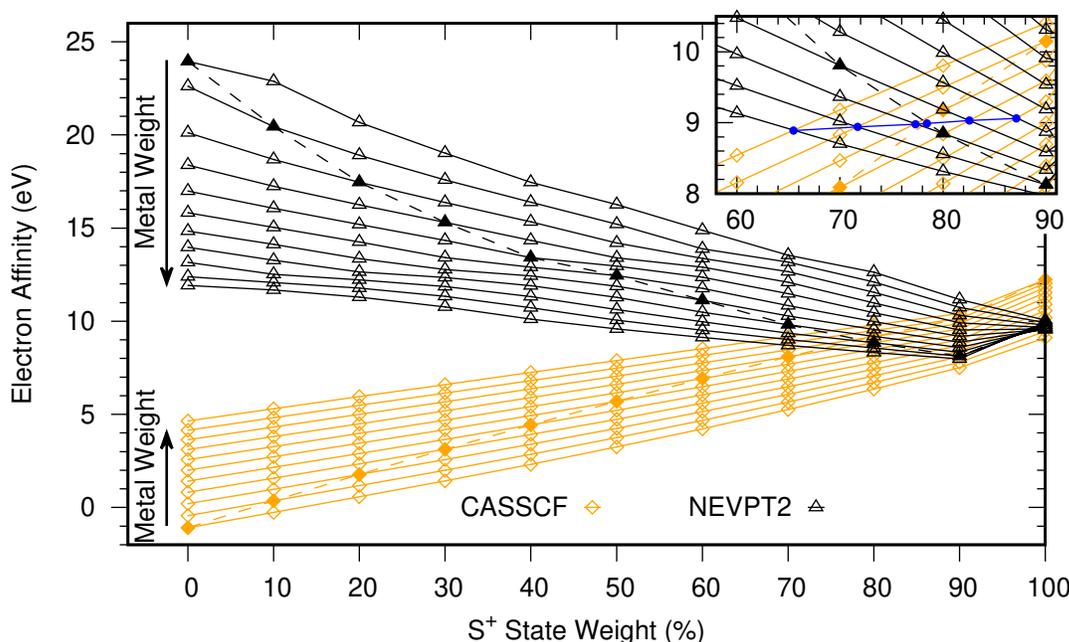

Figure 5: S^{2+} electron affinity calculated using our proposed step-wise protocol. Each point corresponds to a set of (m, w) coordinates, where m is the weight assigned to the metal orbitals optimization and w the one assigned to the molecule states. Both vary from 0 to 100% in increments of 10%, the latter is marked in the x-axis and the former is represented by the full lines, where the increase direction is represented by the arrows. In orange diamond points are those at the CASCI level of theory and in black triangles those from NEVPT2. The filled points and dashed lines represent the points for which $m = w$. The smaller top right plot is a zoom of the larger one. In this plot the blue points mark the weights in which CASCI and NEVPT2 electron affinities are equal.

With such a wide energy variation, it becomes difficult to decide when a description with same accuracy for both sates is achieved. We thus used the following criterion: We considered that

both species have been described with the same accuracy at the CASCI level when their recovered correlation energy at the NEVPT2 level are the same:

$$E^{\text{NEVPT2}}(S^+; m, w) - E^{\text{CAS}}(S^+; m, w) = E^{\text{NEVPT2}}(S^{2+}; m, w) - E^{\text{CAS}}(S^{2+}; m, w), \qquad (8)$$

The reasoning behind is that, if the reference wave functions for the two species are described with similar accuracy, the remaining total energies will be similar. This criterion implies that the electron affinity, Equation (2), calculated with CASCI and NEVPT are the same:

$$\Delta E_{ea}^{\text{CAS}}(m, w) = \Delta E_{ea}^{\text{NEVPT2}}(m, w). \qquad (9)$$

The points where this occur are marked in blue at the inset of Figure 5. The relative energies are consistently between 8.9 and 9.1 eV what is close to our most accurate result of 8.6 eV at the CCSD(T) level of theory (for these points, in the range $65 < w < 98$, $\Delta E \approx 0.008w + 8.369$ with residual of 0.003 eV). Therefore, we are able to obtain a relative energy close to the CCSD(T) state-specific calculation using a state-average NEVPT2 approach for a large range of weights. In other words, it is possible to reproduce accurate relative energies using a single set of orbitals and smaller computational resources. Moreover, the weights w and m for these points correlate almost linearly, with the increase of one causing the decrease of the other ($w \approx 221 - 1.8m$ with residual of 1.1%). This behavior can be interpreted based on the multi-d-occupancy as a balance between orbitals that stabilize the $3d^{10}$-occupancy or stabilize the $3d^9$ one. For instance, when m is large the $3d$-orbitals are biased toward the $3d^{10}$-occupancy, while the ligand ones are biased toward the $3d^9$-occupancy. That way, although this good agreement can be in part via a bias cancellation, when both weights are the same $m = w = 78.45\%$ both groups of orbitals should have the lowest bias toward any occupancy.

With an approach similar to the one used by Graaf et al.,[8] that is, the electron affinity calculated by a single d-shell active space and a single set of orbitals, both states behaves as our protocol without our protocol restriction. In this case, the optimization is made by an usual SA-CASSCF, where the orbitals have more freedom to be optimized. As a consequence, the user has less control over them, with only the weight among states as a parameter. The energy behavior and range of variation is the same as using our protocol, as shown in Figure 6 a). Moreover, the point where both the relative energies are the same is also in the range of energies of our protocol, around 8.9 eV for 70% weight of the S^+ ground state.

Analyzing the optimized orbitals using the usual SA-CASSCF protocol, we observe the same behavior as the PFRS results. That is, the singly occupied active orbital goes from an almost pure
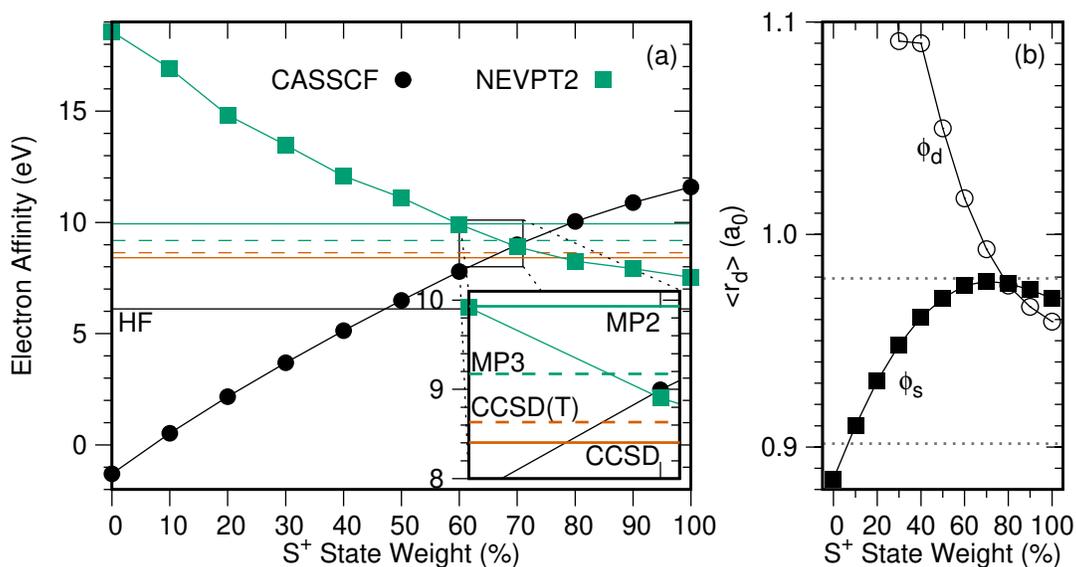
Figure 6: Results for the copper complex subsystem using the usual SA-CASSCF approach as reference wave function. **(a):** Electron Affinity dependence with the orbital optimization at SA-CASSCF and NEVPT2 levels of theory. Horizontal lines show the results for state-specific single reference calulations. **(b):** expected radius for the basis $3d_{z^2}$-orbital that compose the active and closed orbitals described in Equations (3) and (4) by filled squares and hollow circles, respectively. Dotted lines indicate the expected radius value for the optimal d-orbitals of the isolated $Cu^{2+}(3d^9)$, bottom one, and $Cu^+(3d^{10})$, top one.

18

$3d_{z^2}$-orbital to a mixture between this and ligand ones. As a consequence, the wave function for the specie S^{2+} changes from a $3d^9$-configuration to mixed $3d^9/3d^{10}$. Measuring the expected value for the radius of the $3d_{z^2}$ component in the molecular orbitals ϕ_s and ϕ_d, Equations (3) and (4), we can track how the usual SA-CASSCF converged orbitals relate with the more restricted ones used in Figure 5. Figure 6 b) shows the dependence of such orbitals with the weight used for the optimization. We notice that its size is similar to the ones for the isolated copper atom. The expected value for the radius of $Cu^+(3d^{10})$ and $Cu^{2+}(3d^9)$ optimal $3d$-orbitals are around 0.98 and 0.90 a_0, respectively, indicated by the dotted lines in Figure 6 b). Moreover, this molecular orbital quickly reaches the optimal size for the $3d^{10}$-occupancy, at 60% of the S^+ ground state weight, and is kept around this size for higher weights. This is somewhat expected, as the $3d$-orbital size has much more influence over the S^+ ground state total energy than over the S^{2+} one, as observed in Figure 2 a). For ϕ_d the radius is much larger than the one obtained for the isolated atom. This large size, together with its small coefficients, qualitatively depicted by c'_d in Equation (4), for weights smaller than 60%, indicates that those basis functions have only a polarization effect for small weights for S^+. On the other hand, for high weights their radii and coefficients are comparable with those for ϕ_s. Thus, the character that this $3d$ function has in the ϕ_d orbital changes from polarization to valence orbital.

As in the PFRS case, the usual approach has a weight where both CAS and NEVPT2 relative energies are the same. Therefore, we may suppose that this weight generates CASCI wave functions with similar accuracy, as they have the same amount of missing correlation energy. Such crossing occurs around 70%, as shown in the zoom of Figure 6 a). At this point the $3d_{z^2}$-orbital that composes the molecular orbitals is similar to the optimal one for the isolated $Cu^+(3d^{10})$, as shown in Figure 6 b). In terms of our proposed approach, this is similar to defining the metal weight as 100% of the d^{10}-occupancy and around 70% of the S^+ ground state for the ligand one. Comparing these values with the blue points in Figure 5, the leftmost point was 100% of the metal weight and 65% of the ligand with a relative energy of 8.89 eV. That is, our proposed protocol results in values close to the usual SA-CASSCF, with the advantage to control the bias for each group of orbitals. On the other hand, the usual SA-CASSCF has a bias cancellation, with the ligand orbitals biased towards the d^9-occupancy configurations and the d-orbitals biased towards d^{10}.

The molybdenum-copper model system

We observe a remarkable resemblance when comparing the electron affinity and wave function behavior of the copper subsystem with the charge transfer for the bi-centered molybdenum-copper model system. Such resemblance is neither unexpected nor accidental. It was hinted by the similarities in the experimental spectra of the multi-center complex and its sub parts, described by Herrera

19

et al.,[1] and this division has also been explored by Graaf et al. in the geometry optimization of the model system.[6]

Plotting the MMCT energy as explored by Graaf et al.,[8] we observe a similar, almost linear pattern between the usual SA-CASSCF values and the relative weight chosen among states. This is shown in Figure 7 a), for both the literature result[8] and our calculations. To reduce the computational cost we used a smaller basis set, but we can see from Figure 7 a) that the results change only slightly. In the same way as the copper subsystem, Figure 6 a), the energy variation with the weight is larger than 10 eV at both CASSCF and NEVPT2 levels of theory. To decide which weight is the appropriate, we can generalize the criterion used for the electron affinity of the copper subsystem, Equations (8) and (9). In this case, the relative energies calculated with the reference wave function and for the correlated calculation should be the same. These relative energies cross each other at the weight of 60% for the MMCT state, in an energy of 3.8 eV, for the smaller basis set. For the results from reference[8] with the larger basis set, extrapolation of the results suggests that CASSCF and NEVPT2 curves cross each other around 67%, in an energy of 3.1 eV. This energy is very close to the MMCT experimental transition for the original system (between 2.99 and 3.05 eV[1]).

It is relevant to point out that in these calculations the strongly and partially contracted versions of NEVPT2 lead to slightly different results. In the present work, the pc-NEVPT2 transition has a better linear correlation with the weight than the sc-NEVPT2 results (as shown in Figure S4). However, both contractions result in essentially the same relative energies for 20, 40 and 60%, therefore, their analysis give the same conclusion. On the other hand, pc-NEVPT2 results from Graaf et al.[8] (also shown in Figure S4) have larger relative energies when compared with their sc-NEVPT2 counterpart. Although it is not clear the reason for this large relative energy change, these pc-NEVPT2 values also have a smaller dependence with the weight. As a consequence, extrapolating these results the crossing point between pc-NEVPT2 and CASSCF is similar to the one using the sc-NEVPT2 values.

As in the subsystem case, the orbitals can be divided in groups, via the PFRS approach, to provide a better picture of the energies dependence with the convergence of the molecular orbitals. However, in this case there are three sets of orbitals, one for each metal center (copper and molybdenum) and one for the ligands. This increases the number of parameters, and thus increases the number of calculations needed to obtain a reasonable number of points. As the double d-shell effect is larger for first-row transition metals,[30] we choose to focus on the copper weight and use same weight for the molybdenum, that is, a single weight for both metals.[iii] Figure 7 b) shows how each

---

[iii]It is relevant to note that heavier transition metal atoms can have multi-d-occupancy character. However, their configurations' energies have smaller dependence with the optimal $4d$-orbital size. That way, the multi-d-occupancy can be reproduced with a larger range of d-orbitals. For instance, the $Mo^{4+} \rightarrow Mo^{5+}$ $4d$-orbitals size goes from 1.595 to 1.635 a_0, but the variation in the relative energy is of around 1 eV, at the CASSCF level of theory (see Figure S1).
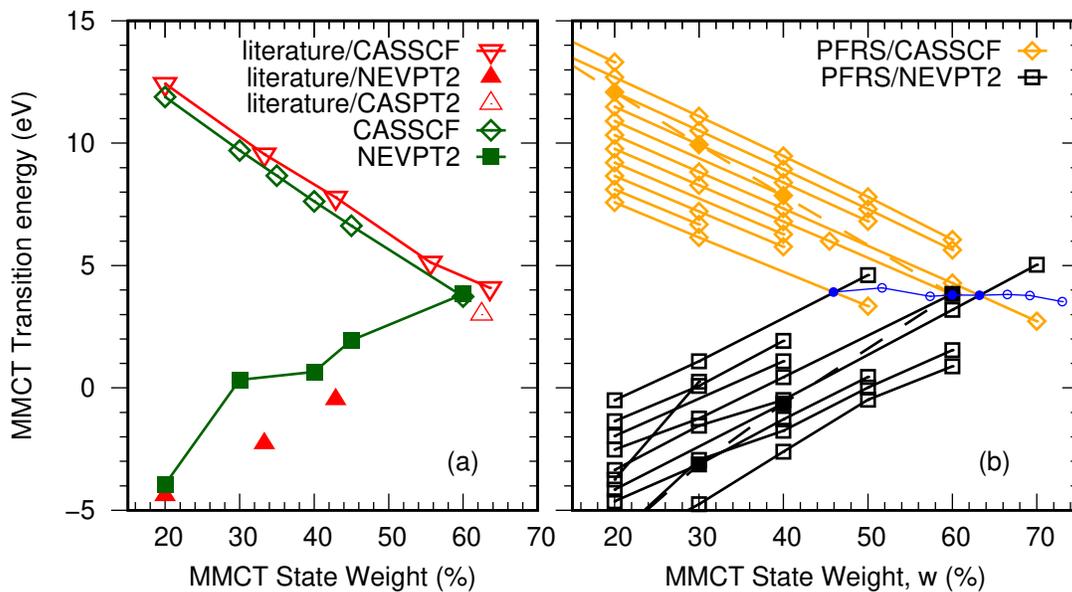
Figure 7: **(a):** Usual SA-CASSCF from reference,[8] represented by the "literature" label, and calculated in this work using a smaller basis set. **(b):** Using the proposed PFRS approach. In the x axis the variation of the w weight, full lines joining results for the same value of m and full points joined by dashed line for $m = w$. Blue circles indicate the weights where the theoretical MMCT transition at CASSCF and NEVPT2 level of theory. Full circles indicate energy "crossing" points obtained by interpolation and hollow ones by extrapolation.

weight affects the relative energy. Fitting the MMCT transition energy at the PFRS/CASSCF level of theory:

$$\Delta E_{\mathrm{MMCT}}^{\mathrm{CAS}}(m, w) \approx -0.055m - 0.157w + 16.31, \tag{10}$$

where m and w are, respectively, the metal and ligand weights in %, and the fitting residual is 0.13 eV. Therefore, the ligands weight has almost thrice the impact in the relative energy if compared with the metal ones. This suggests that, for a reliable description of this transition using the usual double d-shell active space approach, the inclusion of additional ligand orbitals would also be needed, besides the copper double d-shell. Otherwise, the lack of flexibility for the ligand orbitals would under-stabilize the d^{10}-occupancies configurations in relation to the d^9 ones, similarly to the discussed in Figure 2 b), leading to inaccurate relative energies and eventually loss of the multi-d-occupancy character in the reference wave function.

This necessity of having flexibility for ligand orbitals explains why the inclusion of the double d-shell active space (composed of 2 Mo $4d$-orbitals, two $3d$-shells of the copper and two π-orbitals of the CN bridge) by Graaf et al.[8] was not enough to obtain accurate values for the MMCT transition, or to include the double d-shell effect necessary to model it. On the other hand, when the weight for the MMCT state is increased, the metal and ligand orbitals are gradually changed from the

21

optimal for a specific d-occupancy into an intermediate type. This intermediate orbitals between the optimal for d^9 and d^{10}-occupancies is enough to describe the multi-d-occupancy character, and to include the double d-shell effect, according our proposed definition. Therefore, we argue that the results reported previously are much more accurate "for the right reason" than the authors originally thought, as expressed in their discussion (see reference[8]) "[...] we decided to stick to a minimal active space to have the MMCT state as low as possible in the list of roots. The price to pay is that we do not include the double d-shell effect and introduce a small uncertainty in the excitation energies.". Based on our proposed definition, that uses the multi-d-occupancy concept, their approach to optimize the orbitals is likely to have included more of the double d-shell effect than the usual double d-shell active space without the addition of extra ligand orbitals. That way, their good results are not a fortuity, but a (apparently unaware) systematic orbitals optimization to reach orbitals that describe both d-occupancy configurations with similar accuracy.

Although both the usual SA-CASSCF and our proposed step-wise optimization give us a systematic way to vary the relative stabilization between d-occupancies, there is no "end point" that leads to the "best" set of orbitals and nothing that prevents one to overshot the relative stabilization. As before, we have chosen to use the weights values for which the recovered correlation is the same for both states of interest, that is, the points where the relative energies calculated via the CASSCF and NEVPT2 methods are the same. However, this choice requires the calculation of points until the CASSCF and NEVPT energies cross each other and, as shown in Figure 7 b), it was not possible for all weights, due to root-flipping with a state at around 6 eV generated by an internal molybdenum d-d transition (this state energy can be found in Tables S6 and S7).[iv] The metal weights for which the convergence after the crossing points of CASSCF and NEVPT2 energies were possible are shown in full blue circles. For the others, where the convergence was not reached, we applied a simple linear extrapolation of the points. Such simple extrapolation resulted in "crossing" energies, marked by hollow circles, whose values are similar to the ones for which such extrapolation was not necessary. This suggests that the linear behavior of the energies with respect to the weight holds, and we will further explored it. For instance, the correlation between the weights w and m for the points where CASSCF and NEVPT2 energies are the same can be fitted as:

$$m \approx -2.877w + 232.2 \quad \text{and} \quad m \approx -2.937w + 235.0 \,, \tag{11}$$

where we considered only the genuine crossing points and those with the extrapolation, respectively, with residual values of 0.53% and 2.2%, respectively. In these equations the ratio of 1:3 between

---

[iv]It is possible to modify the active space to avoid such molybdenum internal transition by removing the second molybdenum $4d$-orbital. However, due to the copper internal transitions a CAS(11,6) would be necessary.[37]

weights is a consequence of the stability among the energy of the crossing points. Moreover, we notice that, like for the copper subsystem, the usual SA-CASSCF crossing point energy matches with region for the crossings using the PFRS approach.

Graaf et al. stumbled in the same "best orbitals" dilemma, and they used the CASPT2 energy dependence with the level shift as a measurement of orbital quality,[8] that is, the appropriate weight is the one that gives more stable results with respect to variations in the level shift (and thus with less intruder states problems). Although at first this may appear as a completely distinct approach, we understand it is related to our criterion. Considering that intruder states come from the reference wave function interaction with the first-order interacting space (FOIS),[54] a $3d^9$-occupancy with optimal orbitals may have minimal interaction with the FOIS, at the cost of larger interaction for states dominated by the $3d^{10}$ configurations. With the orbitals becoming closer to an intermediate between optimal for d^9 and d^{10}, the effect over the d^{10} configurations decreases, whereas this effect increases over the d^9 configuration, until the total effect over both configurations is minimal. Furthermore, if the recovered correlation energies are almost the same for all calculated states, the states relative energies should not drastically vary with the perturbative approach. As a consequence, for this set of reference states the presence of intruder states should be small, despite of the level shift. Therefore, both criteria to measure the orbitals quality give similar results. Numerically, Graaf et al. estimate a MMCT transition with 3.0 eV, while our PFRS result using the same basis set is 3.01 eV (for $m = w = 60\%$), and 2.99 eV using a larger basis set (def2-QZVPP for Cu, Mo and first coordination sphere atoms, and def2-SVP for all others[43]).

An additional advantage of the proposed step-wise orbitals optimization, the PFRS, is the decrease of computational cost, that comes with the reduction of the number of occupied-active-virtual rotations when performing orbital optimization. This leaded to a reduction of around three times in the calculation time, when our approach is compared to the simultaneous optimization of all orbitals in an usual SA-CASSCF for this system. The actual reduction in time will depend on the number of ligand orbitals and basis set size, but for the present calculations, with our computational resource, the average real time for the usual CASSCF was of 40 hour, while our proposed two-step procedure took 12 hour in average.

Conclusion

In this work we propose an alternative definition for the double d-shell effect: the interaction among distinct d-occupancies or, in analogy to the multireference character, the multi-d-occupancy character of the wave function. Therefore, a qualitative description of the wave function can only be obtained if all configurations with distinct d-occupancies are described with similar accuracy.

23

In the usual double d-shell active space approach, this mixing of configurations with distinct d-occupancies is achieved with an active space that is capable to recover the optimal $3d$-orbital size of multiple d-occupancies, together with a large amount of the d-d dynamic correlation. However, we have shown, previously for small diatomic systems [29,31] and in the present work for a molybdenum-copper model, that a qualitatively correct wave function can be achieved using an intermediate set of orbitals. Moreover, as the mixing of configurations in a given wave function depends on the relative energies among them, our proposed definition also contemplates the reason why the qualitative description of excited states looks to suffer from the double d-shell effect. In this case, even if the ground state has no d-occupancy mixing, that is, does not have a substantial double d-shell effect, some of the excited states eventually will have, what can cause large errors in relative energies. For instance, this has been shown to be the problem for the definition of the triplet/quintet gap in the cobalt monofluoride diatomic molecule [31] and it is likely to be the reason for the CCSD(T)/MRCI discrepancy for the NiF ground state. [55]

In addition to the proposed definition for the double d-shell effect, we have shown that the stabilization of configurations with distinct d-occupancies is dependent on the orbitals spatial location. That is, the optimization of a large active space of local orbitals is not necessarily enough to describe multiple configurations simultaneously. Here we used the electronic affinity of the copper subsystem as example, where both the metal and ligands orbitals have similar relevance for the stabilization d^{10}-occupancy configurations. For the d^9 configurations, on the other hand, variations in the metal orbitals have almost no impact in the energy, but the ligand orbitals are twice more relevant than for the d^{10} configuration. That way, a minimum active space for the ligands will be optimized toward the d^9-occupancy optimal orbitals, as it leads to the larger energy reduction. As a consequence, a double d-shell active space that does not include enough ligand orbitals will likely not be enough to recover the multi-d-character of the wave function, because it only partially stabilizes the d^{10}-occupancy configurations. Such a conclusion, based on the multi-d-occupancy definition of the double d-shell effect, may explains why, in some cases, [15,56] the inclusion of a double d-shell active space is not enough to obtain accurate results without having the active space increased with ligand orbitals as well, often up to infeasible sizes.

As a consequence of the distinct tendency of each group of orbitals toward a given d-occupancy, when a balanced set of weights is used the usual SA-CASSCF is prone to converge each set of the orbitals biased towards one occupancy, what ends up canceling out the bias. To exemplify this, we have shown that varying the SA-CASSCF weights for the copper subsystems tends to converge $3d$-orbitals suitable to the d^{10}-occupancy and the ligands orbitals to the d^9-occupancy. For a better control of the orbital optimization we proposed to partially fix the reference space and optimize the orbitals in steps. Using this procedure we are able to optimize the ligand and metal orbitals in a way that both are intermediate for both occupancies and not an orbital bias cancellation error.

24

We point out that, although we have used this protocol to optimize an inner-valence type active space, this protocol may be used for a larger set of metal active orbitals. Moreover, an additional advantage of this protocol is that the step-wise optimization reduces the number of possible orbital rotations in each of the optimization step, reducing the total optimization time.

We expect that our proposed definition of the double d-shell effect, together with the orbitals division, can help bringing the construction of reference wave functions for transition metal containing compounds to the same level as in other systems. Ideally, a reference wave function should qualitatively describe the ground and low-lying excited states with the smallest as possible recovering of dynamic correlation and with the smallest active space necessary for this. That way, the inclusion of dynamic correlation remains as an exclusive attribution of post-CASCI methods. The present work is a step in this direction, helping the qualitative study of larger systems and improving of accuracy of high-level accurate calculations by reducing the reference size, allowing the use of larger basis sets. Therefore, the presently proposed procedure and definitions should clarify the wave function analysis and, thus, guide further developments in the electronic structure of transition metal containing compounds.

CRediT authorship contribution statement

Matheus Morato F. de Moraes: Conceptualization, Methodology, Software, Writing – original draft. Yuri Alexandre Aoto: Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

(1) Herrera, J. M.; Marvaud, V.; Verdaguer, M.; Marrot, J.; Kalisz, M.; Mathonière, C. Reversible Photoinduced Magnetic Properties in the Heptanuclear Complex $[Mo^{IV}(CN)_2(CN\text{-}CuL)_6]^{8+}$: A Photomagnetic High-Spin Molecule. *Angew. Chem. Int. Ed.* **2004**, *43*, 5468–5471.

(2) Stefańczyk, O.; Ohkoshi, S. Photoswitchable high-dimensional Co^{II}-$[W^V(CN)_8]$ networks: Past, present, and future. *J. Appl. Phys.* **2021**, *129*, 110901.

(3) Ohkoshi, S.; Shiraishi, K.; Nakagawa, K.; Ikeda, Y.; Stefanczyk, O.; Tokoro, H.; Namai, A. Reversible photoswitchable ferromagnetic thin film based on a cyanido-bridged RbCuMo complex. *J. Mater. Chem. C* **2021**, *9*, 3081–3087.

(4) Stefanczyk, O.; Kumar, K.; Pai, T.; Li, G.; ichi Ohkoshi, S. Integration of Trinuclear Triangle Copper(II) Secondary Building Units in Octacyanidometallates(IV)-Based Frameworks. *Inorg. Chem.* **2022**, *61*, 8930–8939.

(5) Pai, T.; Stefanczyk, O.; Kumar, K.; Mathonière, C.; Sieklucka, B.; Ohkoshi, S. Experimental and theoretical insights into the photomagnetic effects in trinuclear and ionic Cu(ii)–Mo(iv) systems. *Inorg. Chem. Front.* **2022**, *9*, 771–783.

(6) Carvajal, M.-A.; Reguero, M.; de Graaf, C. On the mechanism of the photoinduced magnetism in copper octacyanomolybdates. *Chem. Commun.* **2010**, *46*, 5737–5739.

(7) Carvajal, M.-A.; Caballol, R.; de Graaf, C. Insights on the photomagnetism in copper octacyanomolybdates. *Dalton Trans.* **2011**, *40*, 7295–7303.

(8) Domingo, A.; Carvajal, M. A.; de Graaf, C.; Sivalingam, K.; Neese, F.; Angeli, C. Metal-to-metal charge-transfer transitions: reliable excitation energies from ab initio calculations. *Theor. Chem. Acc.* **2012**, *131*, 1264.

(9) Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *J. Chem. Phys.* **2009**, *130*, 114108.

(10) Lischka, H.; Nachtigallová, D.; Aquino, A. J. A.; Szalay, P. G.; Plasser, F.; Machado, F. B. C.; Barbatti, M. Multireference Approaches for Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7293–7361.

(11) Magott, M.; Stefańczyk, O.; Sieklucka, B.; Pinkowicz, D. Octacyanidotungstate(IV) Coordination Chains Demonstrate a Light-Induced Excited Spin State Trapping Behavior and Magnetic Exchange Photoswitching. *Angew. Chem. Int. Ed.* **2017**, *56*, 13283–13287.

(12) Stefańczyk, O.; Pełka, R.; Majcher, A. M.; Mathonière, C.; Sieklucka, B. Irradiation Temperature Dependence of the Photomagnetic Mechanisms in a Cyanido-Bridged $Cu_2^{II}Mo^{IV}$ Trinuclear Molecule. *Inorg. Chem.* **2018**, *57*, 8137–8145.

(13) Manni, G. L.; Smart, S. D.; Alavi, A. Combining the Complete Active Space Self-Consistent Field Method and the Full Configuration Interaction Quantum Monte Carlo within a Super-CI Framework, with Application to Challenging Metal-Porphyrins. *J. Chem. Theory Comput.* **2016**, *12*, 1245–1258.

(14) Manni, G. L.; Alavi, A. Understanding the Mechanism Stabilizing Intermediate Spin States in Fe(II)-Porphyrin. *J. Phys. Chem. A* **2018**, *122*, 4935–4947.

(15) Manni, G. L.; Kats, D.; Tew, D. P.; Alavi, A. Role of Valence and Semicore Electron Correlation on Spin Gaps in Fe(II)-Porphyrins. *J. Chem. Theory Comput.* **2019**, *15*, 1492–1497.

(16) Radoń, M.; Gassowaska, K.; Szklarzewicz, J.; Broclawik, E. Spin-State Energetics of Fe(III) and Ru(III) Aqua Complexes: Accurate ab Initio Calculations and Evidence for Huge Solvation Effects. *J. Chem. Theory Comput.* **2016**, *12*, 1592–1605.

(17) Radoń, M. Benchmarking quantum chemistry methods for spin-state energetics of iron complexes against quantitative experimental data. *Phys. Chem. Chem. Phys.* **2019**, *21*, 4854–4870.

(18) Drabik, G.; Szklarzewicz, J.; Radoń, M. Spin-state energetics of metallocenes: How do best wave function and density functional theory results compare with the experimental data? *Phys. Chem. Chem. Phys.* **2021**, *23*, 151–172.

(19) Drosuo, M.; Mitsopoulou, C. A.; Pantazis, D. A. Reconciling Local Coupled Cluster with Multireference Approaches for Transition Metal Spin-State Energetics. *J. Chem. Theory Comput.* **2022**, *18*, 3538–3548.

(20) Andersson, K.; Roos, B. O. Excitation energies in the nickel atom studied with the complete active space SCF method and second-order perturbation theory. *Chem. Phys. Lett.* **1992**, *191*, 507–514.

(21) Pierloot, K.; Persson, B. J.; Roos, B. O. Theoretical Study of the Chemical Conding in $Ni(C_2H_4)$ and Ferrocene. *J. Phys. Chem.* **1995**, *99*, 3465–3472.

(22) Fischer, C. F. Oscillator strengths for ^{2}S-^{2}P transitions in the copper sequence. *J. Phys. B: At. Mol. Phys.* **1977**, *10*, 1241–1251.

(23) Botch, B. H.; Dunning Jr, T. H.; Harrison, J. F. Valence correlation in the $s^2 d^n$, sd^{n+1}, and d^{n+2} states of the fist-row transition metal atoms. *J. Chem. Phys.* **1981**, *75*, 3466–3476.

(24) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-A.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. *Multiconfigurational Perturbation Theory: Applications in Electronic Spectroscopy*; John Wiley & Sons, Inc., 1996.

(25) Pierloot, K. Transition metals compounds: Outstanding challenges for multiconfigurational methods. *Int. J. Quantum Chem.* **2011**, *111*, 3291–3301.

(26) Li, C.; Evangelista, F. A. Spin-free formulation of the multireference driven similarity renormalization group: A benchmark study of first-row diatomic molecules and spin-crossover energetics. *J. Chem. Phys.* **2021**, *155*, 114111.

(27) Feldt, M.; Phung, Q. M. Ab Initio Methods in First-Row Transition Metal Chemistry. *Eur. J. Inorg. Chem.* **2022**, *2022*, e202200014.

(28) Balabanov, N. B.; Peterson, K. A. Basis set limit electronic excitation energies, ionization potentials, and electron affinities for the 3d transition metal atoms: Coupled cluster and multireference methods. *J. Chem. Phys.* **2006**, *125*, 074110–074120.

(29) de Moraes, M. M. F.; Aoto, Y. A. Reference spaces for multireference coupled-cluster theory: the challenge of the CoH molecule. *Theor. Chem. Acc.* **2020**, *139*, 71.

(30) Pierloot, K. In *Computational Organometallic Chemistry*; Cundari, T. R., Ed.; Marcel Dekker, Inc., 2001; Chapter 5.

(31) de Moraes, M. M. F.; Aoto, Y. A. Multi-reference and multi-occupancy character of the cobalt monofluoride. *J. Mol. Spectrosc.* **2022**, *385*, 111611.

(32) Chong, D. P.; Langhoff, S. R.; Bauschlicher Jr., C. W; Walch, S. P.; Partridge, H. Theoretical dipole moments for the firstrow transition metal hydrides. *J. Chem. Phys.* **1986**, *85*, 2850–2860.

(33) Bauschlicher Jr., C. W. Configuration Interaction Benchmark Calculations for TiH. *J. Phys. Chem.* **1988**, *92*, 3020–3023.

(34) Hay, P. J. Gaussian basis sets for molecular calculations. The representation of 3d orbitals in transition-metal atoms. *J. Chem. Phys.* **1977**, *66*, 4377–4384.

(35) Werner, H.-J.; Knowles, P. J. A second order multiconfiguration SCF procedure with optimum convergence. *J. Chem. Phys.* **1985**, *82*, 5053–5063.

(36) Knowles, P. J.; Werner, H.-J. An efficient second-order MC SCF method for long configuration expansions. *Chem. Phys. Lett.* **1985**, *115*, 259–267.

(37) de Moraes, M. M. F. The *d*-orbital size and occupancy effects over the theoretical description of transition metal containing compounds. Ph.D. thesis, Federal University of ABC, Santo André-SP, Brazil, 2023.

(38) Levine, B. G.; Durden, A. S.; Esch, M. P.; Liang, F.; Shu, Y. CAS without SCF-Why to use CASCI and where to get the orbitals. *J. Chem. Phys.* **2021**, *154*, 090902.

(39) Freidorf, M.; Marian, C. M.; Hess, B. A. Theoretical study of the electronic spectrum of the CoH molecule. *J. Chem. Phys.* **1993**, *99*, 1215–1223.

(40) MOLPRO, version 2021 a package of *ab initio* programs. H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, and others; available from `http://www.molpro.net`.

(41) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242–253.

(42) Werner, H.-J. et al. The Molpro quantum chemistry package. *J. Chem. Phys.* **2020**, *152*, 144107.

(43) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(44) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. N-electron valence state perturbation theory: a fast implementation of the strongly contracted variant. *Chem. Phys. Lett.* **2001**, *350*, 297–305.

(45) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.

(46) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. n-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.* **2002**, *117*, 9138–9153.

(47) Williams, T.; Kelley, C.; many others, Gnuplot 5.4: an interactive plotting program. `http://gnuplot.sourceforge.net/`, 2020.

(48) Hampel, C.; Peterson, K. A.; Werner, H.-J. A comparison of the efficiency and accuracy of the quadratic configuration-interaction (QCISC), coupled cluster (CCSD), and brueckner coupled cluster (BCCD) methods. *Chem. Phys. Lett.* **1992**, *190*, 1–12.

(49) Harrison, J. F. Electronic Structure of Diatomic Molecules Composed of a First-Row Transition Metal and Main-Group Element (H-F). *Chem. Rev.* **2000**, *100*, 679–716.

(50) Jiang, W.; DeYonker, N. J.; Wilson, A. K. Multirefernce Character for 3d Transition-Metal-Containing Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 460–468.

(51) Wiberg, K. B.; Rablen, P. R. Comparison of atomic charges derived via different procedures. *J. Comp. Chem.* **1993**, *14*, 1504–1518.

(52) Grier, D. L.; Streitwieser Jr., A. Electron Density Analysis of Sustituted Carbonyl Groups. *J. Am. Chem. Soc.* **1982**, *104*, 3556–3564.

(53) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.

(54) Roos, B. O.; Anderson, K. Multiconfigurational perturbation theory with level shift - the Cr_2 potential revisited. *Chem. Phys. Lett.* **1995**, *245*, 215–223.

(55) Koukounas, C.; Mavridis, A. Ab initio Study of the Diatomic Fluorides FeF, CoF, NiF, and CuF. *J. Phys. Chem. A* **2008**, *112*, 11235–11250.

(56) Roemelt, M.; Pantazis, D. A. Multireference Approaches to Spin-State Energetics of Transition Metal Complexes Utilizing the Density Matrix Renormalization Group. *Adv. Theory. Simul.* **2019**, *2*, 1800201.