



**HAL**  
open science

## Investigating solvent effects on the magnetic properties of molybdate ions ( $\text{MoO}_4^{2-}$ ) with relativistic embedding

Loïc Halbert, Malgorzata Olejniczak, Valérie Vallet, André Severo Pereira Gomes

### ► To cite this version:

Loïc Halbert, Malgorzata Olejniczak, Valérie Vallet, André Severo Pereira Gomes. Investigating solvent effects on the magnetic properties of molybdate ions ( $\text{MoO}_4^{2-}$ ) with relativistic embedding. International Journal of Quantum Chemistry, 2020, pp.e26207. 10.1002/qua.26207 . hal-02413336

**HAL Id: hal-02413336**

**<https://hal.science/hal-02413336>**

Submitted on 6 Aug 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Investigating solvent effects on the magnetic properties of molybdate ions ( $\text{MoO}_4^{2-}$ ) with relativistic embedding

Loïc Halbert,<sup>1</sup> Małgorzata Olejniczak,<sup>2</sup> Valérie Vallet,<sup>1</sup> and André Severo Pereira Gomes<sup>1,\*</sup>

<sup>1</sup>*Univ. Lille, CNRS, UMR 8523 - PhLAM - Physique des Lasers Atomes et Molécules, F-59000 Lille, France*

<sup>2</sup>*Centre of New Technologies, University of Warsaw, S. Banacha 2c, 02-097 Warsaw, Poland*

(Dated: May 21, 2020)

We investigate the ability of mechanical and electronic density functional theory (DFT)-based embedding approaches to describe the solvent effects on nuclear magnetic resonance (NMR) shielding constants of the  $^{95}\text{Mo}$  nucleus in the molybdate ion in aqueous solution. From the description obtained from calculations with two- and four-component relativistic Hamiltonians, we find that for such systems spin-orbit coupling effects are clearly important for absolute shielding values, but for relative quantities a scalar relativistic treatment provides a sufficient estimation of the solvent effects. We find that the electronic contributions to the solvent effects are relatively modest yet decisive to provide a more accurate magnetic response of the system, when compared to reference supermolecular calculations. We analyze the errors in the embedding calculations by statistical methods as well as through a real-space representation of NMR shielding densities, which are shown to provide a clear picture of the physical processes at play.

**Keywords** NMR shielding, quantum embedding, relativistic effects, shielding density, molybdate, solvation.

## I. INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopy measures the interaction between the magnetic moments of nuclei and an applied external field, screened by the electrons of the system. NMR is extremely useful for characterizing molecules and materials since it provides detailed information of the local chemical environment around the responding nuclei, and does so in a non-destructive manner. In addition to that, it probes species in their electronic ground state and, in doing so, provides more direct information on the interactions in that state compared to other techniques.

The screening by the electron cloud of the magnetic interactions between a given atomic isotope  $K$  and the applied external magnetic field, is represented by the so-called NMR shielding constant,  $\sigma^K$ , which can be calculated as a second-order derivative of the energy of the molecular system with respect to the magnetic dipole moment of that nucleus ( $\vec{m}_K$ ) and an external magnetic field ( $\vec{B}$ ) at the zero-perturbation limit:

$$\sigma_{\alpha\beta}^K = \left. \frac{d^2 E}{dB_\alpha dm_{K;\beta}} \right|_{\vec{B}, \{\vec{m}_A\}=0}. \quad (1)$$

In most cases, what is measured in experiments is not  $\sigma^K$  but rather a signal relative to a chosen reference species, the chemical shift  $\delta^K$ , though in recent years experimental advances have revived the interest in determining absolute shielding scales [1–3].

NMR spectra can be quite difficult to interpret. The difficulties may come from the complexity due to the size

of composition of the system (species in solution, disordered materials, etc.), the broadening of signals due to quadrupole interactions (for nuclei with total angular momentum larger than  $I = 1/2$ ) or a combination of all of these [4]. Because of these difficulties, theoretical modeling has become indispensable for the interpretation of experiments [5–7]. Furthermore, for certain systems exhibiting a wide range of chemical shifts, or long acquisition times, theory is crucial to guide experiments by providing the approximate regions of the spectra in which to search for signals.

In this paper we are interested in exploring the use of quantum embedding [8–11] to describe the solvent effects on the NMR shielding constants of the  $^{95}\text{Mo}$  on the molybdate dianion in aqueous solution. The molybdate moiety is arguably the simplest experimentally relevant Mo-containing system that can be studied, and it is often used as an experimental reference system in the determination of  $^{95}\text{Mo}$  chemical shifts (in the form of an aqueous solution of sodium molybdate,  $\text{Na}_2\text{MoO}_4$  [12]. As such, it can be seen as ideal test system to evaluate theoretical approaches that can be applied to the modeling of other, more complicated species. Examples of more complex systems can be found in different classes of molybdenum oxides containing the  $\text{MoO}_3$  and  $\text{MoO}_4$  moieties. These oxides are found as component of catalysts [13–15], as fission products [16], and also make up materials that can be used for applications ranging from nonlinear optical materials [17] to glasses for vitrification of nuclear waste [18]. In most cases, the interest lies in understanding how the Mo atom modifies the properties (optical, mechanical etc) of the starting material, and that is where the characterization by NMR becomes interesting.

For materials and complex systems, often treated with band structure methods, the gauge-including projector augmented wave (GIPAW) [6, 7] approach, coupled to

\* [andre.gomes@univ-lille.fr](mailto:andre.gomes@univ-lille.fr)

density functional theory (DFT) calculations, is the de facto standard for NMR calculations useful to experimentalists. Quantum embedding, on the other hand, represent a class of theoretical methods, often used in combination to approaches developed for discrete systems (based on DFT but also on more accurate treatments based upon wavefunction theory, WFT). In it a given system is treated quantum mechanically but on the basis of a set of interacting subsystems, for which one can tune the accuracy of the description of each subsystem (by choosing for instance a DFT or WFT treatment), according to a desired compromise between accuracy and computational cost. As such, one can describe the interactions between the species of interest and its surroundings at a fraction of the cost of applying standard DFT or WFT approaches to the whole system.

Contrary to other embedding schemes [19–21], such as those in which only the changes in structure due to the environment are considered (mechanical embedding), or where the environment is represented classically (QM/MM) [22], quantum embedding is yet to be extensively used for describing magnetic properties in general (and NMR shieldings in particular), in spite of developments of response theory for magnetic perturbations for the frozen density embedding or subsystem DFT methods [23–27], as well as the generalizations of partition DFT to the spin-polarized case [?] and the subsystem formulation of time-dependent current density functional theory [?]. This is unfortunate since these approaches have the advantage of allowing, or providing a basis for, a seamless combination of different electronic structure approaches (DFT and WFT), but also the combination of different Hamiltonians, such as those taking into account relativistic effects.

Relativistic effects [28–31] (scalar relativistic effects and spin-orbit coupling) are now recognized as important (if not essential) to approach experimental accuracy in electronic structure calculations. While this is undisputed in the case of heavy elements ( $Z > 31$ ), it is increasingly the case that the effects of relativity are recognized and accounted for even for light systems (first and second-row atoms) when the molecular properties of interest involve a description of the electronic structure near the nucleus. This is the case of NMR properties [32, 33].

We expect that quantum embedding, if shown to be reliable, can provide a flexible framework and enable more sophisticated simulations, either by complementing DFT-only approaches such as GIPAW (by easily letting one explore the use of different relativistic Hamiltonians) or by allowing the use of more accurate methods, such as those based on the coupled cluster ansatz [34, 35], thus paving the way for a quantitative leap in the interpretation of NMR spectra. As the hydrated  $\text{MoO}_4^{2-}$  molybdate dianion appears to have well-structured hydrations shells [12], we are particularly interested in assessing how well quantum embedding can effectively reduce the size of the explicit quantum model, by accounting for hydration

effects.

The paper is organized as follows: in section II we outline the key theoretical aspects underpinning the work. This is followed in section III by a summary of the computational details of our calculations, with a particular emphasis in detailing which structural models are considered and introducing a shorthand notation to refer to these. In section IV we discuss the ability of the different embedding approaches to represent the molybdate ion in solution, the influence of the different relativistic Hamiltonians on the NMR shielding constants, and characterize (including in real space) the errors of the different embedding methods. We conclude and offer perspectives for future work in section V.

## II. THEORY

In what follows, we employ the standard notation for molecular orbital indices (with  $i, j, \dots$  corresponding to occupied,  $a, b, \dots$  - to virtual and  $p, q, \dots$  - to general molecular orbitals) and the summation convention over repeated indices. We use the SI-based atomic units ( $\hbar = m_e = e = 1/(4\pi\epsilon_0) = 1$ ) [36]. Employing the second-quantization formalism and an exponential parametrization of the unperturbed, closed-shell ground state wave function  $\psi(\boldsymbol{\kappa})$  (with the orbital-rotation operator,  $\hat{\kappa} = \sum_{pq} \kappa_{pq} p^\dagger q$ , acting on a trial wave function  $\psi_0$

$$\psi(\boldsymbol{\kappa}) = e^{\hat{\kappa}} \psi_0, \quad (2)$$

and a parametrization of the perturbed wave function in terms of first-order orbital-rotations ( $\hat{\kappa}^{(1)} = \sum_{pq} \kappa_{pq}^{(1)} p^\dagger q$ ) [27], the derivative in Eq. 1 can be further expressed as

$$\left. \frac{d^2 E}{dB_\alpha dm_{K;\beta}} \right|_{\epsilon=0} = \left. \frac{\partial^2 E}{\partial \kappa_{pq}^{(1)} \partial m_{K;\beta}} \frac{\partial \kappa_{pq}^{(1)}}{\partial B_\alpha} \right|_{\epsilon=0} + \left. \frac{\partial^2 E}{\partial B_\alpha \partial m_{K;\beta}} \right|_{\epsilon=0}, \quad (3)$$

where all perturbing fields are collected in  $\epsilon$ . Partial derivatives of the orbital rotation parameters ( $\{\kappa_{pq}^{(1)}\}$ ) with respect to the perturbing field are optimized in the linear response equations, which in the static-field regime can be compactly presented as:

$$E_0^{[2]} X_{\epsilon_1} = -E_{\epsilon_1}^{[1]}, \quad (4)$$

with  $E_0^{[2]}$ ,  $E_{\epsilon_1}^{[1]}$  and  $X_{\epsilon_1}$  corresponding to the electronic Hessian, the property gradient and the solution vector ( $X_{\epsilon_1} = (\partial \kappa_{pq}^{(1)} / \partial \epsilon_1)_{\epsilon=0}$ ), respectively [37]. The two former quantities require the formulation of various derivatives of Fock or Kohn-Sham matrices, therefore they are strictly connected to the formalism employed in the calculations.

In this work, we exploit the computational model based on relativistic Hamiltonians [38] and on the spin-density functional theory (SDFT) [39–41]. The closed-shell system is described by the electron density ( $\rho_0$ )

and the spin-density calculated in a non-collinear fashion as a norm of the spin-magnetization vector ( $s = |\rho_\mu|$ ,  $\mu \in \{x, y, z\}$ ). Following previous works [27, 42–44], we collect the electron density and the spin-magnetization vector into one variable - a general density component,  $\rho_k$  ( $k \in \{0, x, y, z\}$ ).

This framework becomes slightly modified in the embedding situation. The DFT-subsystem-based approaches rely on the partitioning of the total system into (interacting) subsystems [8, 9, 45], which is realized by expressing the electron density of the full system as a sum of electron densities of subsystems [10, 46–48] (the same applies to the spin density if it is included in the formalism),

$$\rho_k^{\text{tot}}(\vec{r}) = \rho_k^I(\vec{r}) + \rho_k^{II}(\vec{r}) \quad (5)$$

Consequently, the total energy may be decomposed into contributions corresponding to energies of subsystems and the interaction energy term which depends on the (spin) densities of all subsystems [27],

$$E_{\text{tot}}[\rho_k^{\text{tot}}] = E_I[\rho_k^I] + E_{II}[\rho_k^{II}] + E_{\text{int}}[\rho_k^I, \rho_k^{II}]. \quad (6)$$

The analytical formula for the interaction energy term can be derived from the (S)DFT expression for the energy of the full system and reads as:

$$\begin{aligned} E_{\text{int}}[\rho_k^I, \rho_k^{II}] &= \int [\rho_0^I(\vec{r})v_{\text{nuc}}^{II}(\vec{r}) + \rho_0^{II}(\vec{r})v_{\text{nuc}}^I(\vec{r})] d\vec{r} \\ &+ \iint \frac{\rho_0^I(\vec{r}_1)\rho_0^{II}(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2 + E_{\text{nuc}}^{I,II} \\ &+ E_{\text{xc}}^{\text{nadd}} + T_s^{\text{nadd}}, \end{aligned} \quad (7)$$

where, in addition to terms describing the interaction of the electron density of a subsystem with the electron density and the nuclear potential of another subsystem and a term related to the nuclear repulsion energy between subsystems (first four terms in Eq. 7), the non-additive exchange-correlation ( $E_{\text{xc}}^{\text{nadd}}$ ) and kinetic energy ( $T_s^{\text{nadd}}$ ) contributions emerge. These non-additive contributions [49] are defined as:

$$X^{\text{nadd}} \equiv X^{\text{nadd}}[\rho_k^I, \rho_k^{II}] = X[\rho_k^{\text{tot}}] - X[\rho_k^I] - X[\rho_k^{II}] \quad (8)$$

and they depend on the general density components of all subsystems. In typical FDE calculations, one subsystem is chosen as *active*, while another subsystem constitute its *environment*. In this setting, the *active* density (e.g.  $\rho_k^I$ ) is optimized by solving the so-called Kohn-Sham (KS) equations for a constrained electron density (KSCED) [50], in which an effective KS potential known from a standard (S)DFT is augmented by the embedding potential responsible for describing the interaction of an *active* subsystem with the *environment* (here:  $\rho_k^{II}$ ):

$$\begin{aligned} v_{\text{emb};k}^I(\vec{r}) &= \frac{\delta E_{\text{int}}}{\delta \rho_k^I(\vec{r})} = \frac{\delta E_{\text{xc}}^{\text{nadd}}}{\delta \rho_k^I} + \frac{\delta T_s^{\text{nadd}}}{\delta \rho_k^I} + v_{\text{nuc}}^{II}(\vec{r}) \\ &+ \int \frac{\rho_0^{II}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}'. \end{aligned} \quad (9)$$

It is also possible to relax the density of the *environment* by solving the analogous KSCED equations set up for  $\rho_k^{II}$ . Repeating this procedure in an iterative manner, known as the *freeze-thaw* [25] (FnT) cycle allows to optimize the densities of all subsystems. Clearly, the embedding potential in Eq. 9 couples the systems in the ground state.

At this point it should be noted that although FDE is formally exact, its practical realization requires approximation to the non-additive exchange-correlation and kinetic energies. This is typically done with approximate density functionals, which - especially for the kinetic energy part - have limited accuracy [8, 51–54]. From the presence of the interaction energy term in the total energy expression (Eq. 5) and from its functional dependence on the densities of all subsystems, it becomes evident that the quantities which in the linear response theory are formulated as various energy derivatives ( $E_0^{[2]}$ ,  $E_{\varepsilon_1}^{[1]}$  in Eq. 4) now have to be augmented with analogous derivatives of  $E_{\text{int}}$ .

These additional contributions, derived by using the chain rule, involve the first- and second-order derivatives of  $E_{\text{int}}$  with respect to the densities of subsystems - the embedding potential (Eq. 9) and the embedding kernel (see [27] for details). In effect, the property gradient and the electronic Hessian (Eq. 4) are elegantly subdivided into contributions from isolated subsystems and terms related to the interaction input [49]. In practice, these derivatives can be coded analogously to the exchange-correlation terms in standard (S)DFT technique.

There is, however, an additional difficulty in the molecular property calculations, specific to the embedding setting. Namely, one needs to account for the fact that the property of interest arises as a response of the full system - comprising of all subsystems - to certain external perturbations. In particular, for the NMR shielding tensor of a nucleus K this can be symbolically expressed as

$$\sigma_{\alpha\beta}^K = \sigma_{\alpha\beta}^{K,I} + \sigma_{\alpha\beta}^{K,II}. \quad (10)$$

As evoked above and discussed in [27], the response of the interaction term will introduce terms in the electronic Hessian and property gradient that couple the response of the subsystems to the external perturbations, so each  $\sigma_{\alpha\beta}^{K,I}$  and  $\sigma_{\alpha\beta}^{K,II}$  formally contain intra-subsystem and inter-subsystem contributions. In this work we consider the same approximation as in prior works [23, 24] and disregard the inter-subsystem (coupling) contributions in the calculation of  $\sigma_{\alpha\beta}^{K,I}$  and  $\sigma_{\alpha\beta}^{K,II}$ .

We note that the term  $\sigma_{\alpha\beta}^{K,II}$ , assuming that the nucleus K belongs to subsystem I, can be interpreted as an effect that the perturbation due to  $B_\alpha$  in subsystem II has in the point of the center of nucleus K - an idea analogous to the concept of the nucleus independent chemical shift (NICS) [55]. A final complication in the formulation of the NMR shielding tensor arises as a consequence of the dependence of this property operator (more specifically - the Zeeman term) on the choice of the gauge origin,

which leads to unphysical results in incomplete basis set regime. The typical remedy is to use the London atomic orbitals (LAOs) [56, 57] instead of atomic orbitals, however with the price of the need to calculate additional terms in the property gradient [27, 44, 58].

The NMR shielding tensor (Eqs. 1, 3) can also be determined from numerical integration of the NMR shielding density [59, 60]. This property density can be calculated (analytically) in real space by using the relation to the magnetically-induced current density,  $j^B$  - also a second-order tensor quantity, which in the relativistic framework [61, 62] reads

$$j^B = \left. \frac{d\vec{j}}{d\vec{B}} \right|_{\vec{B}=0}; \quad \vec{j} = -e\psi_i^\dagger c\vec{\alpha}\psi_i. \quad (11)$$

The NMR shielding density is an integrand in the following expression:

$$\sigma_{\alpha\beta}^K = -\frac{1}{c^2} \int \frac{1}{r_K^3} \left( \vec{r}_K \times \vec{j}^{B\alpha} \right)_\beta d\vec{r}. \quad (12)$$

This reformulation of  $\sigma_{\alpha\beta}^K$  has multiple advantages. First, it allows to recalculate the NMR shielding values (therefore serves as a test of a fully analytical approach), secondly it opens the possibility to visualize the NMR shielding density on meshes, which has already proved to be an invaluable asset in the post-production analysis process [27]. Additional advantage of this formulation is that the NMR shielding density of a nucleus K can easily be calculated in any point in space, even if that point and the center of nucleus K belong to different subsystems, what permits to evaluate the second term of Eq. 10 [23].

### III. COMPUTATIONAL DETAILS

The solvated molybdate structures used in this work are taken from the Car-Parrinello molecular dynamics (CPMD [63]) calculations of Nguyen and coworkers [12]. We considered 517 snapshots, each containing the molybdate ion and 20 water molecules. For additional details, the reader is referred to the aforementioned paper. From these we defined four basic structural models:

- a supermolecular system, containing the molybdate ion and all water molecules;
- a subsystem embedding model, where the molybdate ion (and depending on the case, selected water molecules) is the active subsystem and the (remaining) water molecules make up the environment, but in which a given number of water molecules nearest to the active subsystem are relaxed through *freeze-thaw* cycles while the others remain frozen. We will refer to this model as FnT.

The electron densities and electrostatic potential for the environment are obtained either from calculations on individual water molecules (and we refer

to a “fragmented” environment) or by these forming a single subsystem (and we refer to a “grouped” environment);

- a frozen density embedding model, in which all water molecules in the environment are kept frozen. We will refer to this model as FDE, and the same notations for the environment (“fragmented” or “grouped”) as for FnT is used;
- an isolated molybdate ion. As the structures of the latter are not obtained in the gas phase but rather through the CPMD calculations, this model is equivalent to that of mechanical embedding, and as such we will refer to it as “mechanical embedding”.

Using the labels “A” for the molybdate ion and “B” for the water molecules, and indices **g** and **f** to denote the “grouped” or “fragmented” environment, we shall use the following shorthand notation to represent these four cases: [*Active* | *Relaxed* | *Frozen*]. This means, for example, that the supermolecular model is denoted by [*A* + 20*B*|0|0] (all species are in the active fragment), the mechanical embedding by [*A*|0|0], and a FnT model in which 20 molecules make up the environment but in which 11 are relaxed (and all are taken as individual fragments), is denoted by [*A*|11*B*<sub>f</sub>|9*B*<sub>f</sub>].

The execution of all calculations, including the preparation of the different fragments, has been handled with the PyADF [64] scripting framework. In all calculations, we employ Becke integration grids, and a gaussian nuclear charge distribution [65] to describe the finite nuclear volume [66], PBE [67, 68] exchange-correlation functional. For subsystem embedding and FDE embedding calculations (see below), we have employed the PW91k [69] and PBE [67, 68] functionals for calculating the non-additive kinetic and exchange-correlation contributions, respectively, and used a monomer expansion for the basis sets. In the case of subsystem embedding, we have performed 5 *freeze-thaw* cycles. In the NMR calculations we employ London orbitals.

#### A. One- and two-component calculations

The scalar relativistic (SR) and spin-orbit (SO) ZORA [70–73] calculations have been performed with the ADF [74] code, using the TZ2P basis sets [75]. The NMR calculations have been performed with the NMR program.

#### B. Four-component calculations

The calculations employing the Dirac-Coulomb (DC) Hamiltonian have been performed with the Dirac code [76], revision 54ab939. For the Mo atom the dyall.cv3z basis [77] is used, and for all lighter atoms

the aug-cc-pVTZ [78] basis is used. To aid in the convergence of the SCF procedure, in all calculations we have used the atomic start procedure, described in the DIRAC documentation.

For the subsystem embedding and FDE embedding calculations NMR calculations, we remain in the uncoupled response approximation, but take into account contributions from the spin density according to Ref. 27. We note that unless otherwise noted, the embedding calculations make use of results of embedded ZORA calculations with ADF (embedding potentials, electrostatic potentials and frozen electron densities and gradients of the density for the environment).

The NMR shielding densities used in the analysis of the embedding methods (see section IV C) have been obtained with the analysis module of the Dirac code, and are based on four-component calculations with the DC Hamiltonian and equation 12. In the case of embedded calculations, following equation 10, it was necessary to also perform calculations for the environment using the position of the Mo atom as the  $\vec{r}_K$  position.

### C. Plotting

All graphs have been prepared with Matplotlib [79] python library, with the exception of the volumetric density and NMR shielding density plots, for which the Mayavi [80] python library has been used.

## IV. RESULTS AND DISCUSSION

In what follows we are interested in the isotropic part of the NMR shielding tensor of the  $^{95}\text{Mo}$  nucleus,  $\sigma_{iso}$ , which is defined as:

$$\sigma_{iso} = 1/3(\sigma_{11} + \sigma_{22} + \sigma_{33}) \quad (13)$$

where the indices refer to the principal axes of the tensor according to the Mason's notation [81], with  $\sigma_{11} \leq \sigma_{22} \leq \sigma_{33}$ .

### A. Minimal structural model

An embedding-based computational strategy takes its most effective form, from a computational cost perspective, when the smallest possible molecular moiety can be considered as the active subsystem, and the rest of the system as the environment. In this work such a situation would correspond, as discussed above, to having only the  $\text{MoO}_4^{2-}$  ion in the active subsystem. In this subsection we explore this model. As our goal in this paper is to investigate the performance of embedding models, we use as reference for comparison the supermolecular model  $[A + 20B|0|0]$ , since that lets us directly trace back any discrepancies to the approximations introduced by the embedded models.

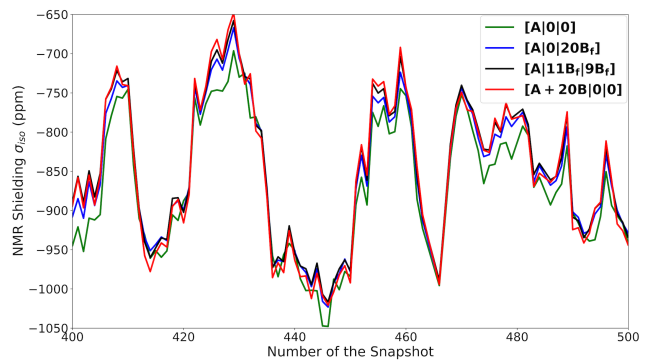


FIG. 1:  $\sigma_{iso}$  (in ppm) for a subset of the snapshots considered (snapshots 400 to 500), obtained from SR-ZORA calculations for the supermolecule ( $[A + 20B|0|0]$ ), electronic (FnT:  $[A|11B_f|9B_f]$  and FDE:  $[A|0|20B_f]$ ) and mechanical ( $[A|0|0]$ ) embedding.

### 1. Structural and electronic effects on NMR shieldings

We start the discussion with the analysis of the influence of the geometry of molecular complexes on  $\sigma_{iso}$  based on the SR-ZORA calculations. This is summarized in Figure 1, which shows the values of  $\sigma_{iso}$  obtained for a subset of 100 geometries out of the 517 considered MD simulations.

An interesting observation that emerges from this figure is that the description offered by different approximations (FnT, FDE, mechanical embedding) is not consistent for all the snapshots. While for certain geometries the electronic part of embedding is an essential component of the model in order to approach the supermolecular  $\sigma_{iso}$  values, for others these contributions are negligible and the mechanical embedding is sufficient to reproduce the supermolecular baseline.

Model	$\bar{\sigma}_{iso} \pm std$	$\Delta\bar{\sigma}_{iso}$ wrt. $[A + 20B 0 0]$
$[A + 20B 0 0]$	$-848.5 \pm 108.9$	-
$[A 11B_f 9B_f]$	$-841.5 \pm 103.9$	$7.0 \pm 8.5$
$[A 0 20B_f]$	$-848.0 \pm 101.8$	$3.5 \pm 13.6$
$[A 0 0]$	$-856.5 \pm 99.1$	$-7.9 \pm 24.5$

TABLE I: Mean ( $\bar{\sigma}_{iso}$ ) and Standard Deviations ( $std$ ) of  $\sigma_{iso}$  (in ppm) obtained from all snapshots for SR-ZORA calculations for various models: Supermolecule  $[A + 20B|0|0]$ , FnT  $[A|11B_f|9B_f]$ , FDE  $[A|0|20B_f]$  and isolated  $[A|0|0]$ .  $\Delta\bar{\sigma}_{iso}$  are the deviations with respect to  $[A + 20B|0|0]$ .

The fact that the values of  $\sigma_{iso}$  are often close to each other for the different snapshots as shown in Figure 1 is reflected on the similarity of the mean values of the  $^{95}\text{Mo}$  NMR shielding constants, featured as  $\bar{\sigma}_{iso}$  in Table I, calculated over all the snapshots for all the considered approximations (FnT, FDE, mechanical embedding). However, the choice of the computational model does affect the standard deviation of the  $\sigma_{iso}$  values distribution -

in this case the variance of the NMR shielding values over all the snapshots is significantly smaller if electronic embedding is included in the calculations. This signifies that the data from electronic embedding calculation is statistically less disperse, therefore in this case these can be considered as more accurate for the modeling of  $\sigma_{iso}$  and  $\bar{\sigma}_{iso}$ .

In order to understand these structural dependencies of  $\sigma_{iso}$  values, we decided to define a measure, the mean distance between the centers of Mo atom and O atoms in  $\text{MoO}_4^{2-}$ , hereafter referred to as  $\bar{r}[\text{Mo-O}]$ , and to plot the  $\sigma_{iso}$  vs  $\bar{r}[\text{Mo-O}]$  in Figure 2.

The first conclusion of this figure is that there is a significant correlation between the  $\sigma_{iso}$  and  $\bar{r}[\text{Mo-O}]$  for all approximations, as seen from the large values of  $R^2$ . The largest correlation between these two quantities was found for the mechanical coupling ( $R^2 = 0.992$ ) and the lowest for the supermolecular results ( $R^2 = 0.945$ ), which indicates the possibility that the effect of the interaction between two subsystem on  $\sigma_{iso}$  values cannot be simply attributed to the structural variations, or at least the ones represented by the  $\bar{r}[\text{Mo-O}]$  measure. The FDE and FnT methods recover some of the  $\bar{r}[\text{Mo-O}]$ -dependent contributions, as illustrated by the correlation coefficients ( $R^2 = 0.962$  and  $0.976$  for FnT and FDE, respectively). The slope of the linear regression also shows a better agreement between supermolecule and electronic embedding calculations than between supermolecule and mechanical embedding.

However, the linearity of the relationship between  $\sigma_{iso}$  and  $\bar{r}[\text{Mo-O}]$  is interesting since the NMR shielding tensor is considered to be a local property. Yet, as we shall discuss below, there might be other long-range effects that dampen the  $r^{-2}$  dependence of the corresponding property operator.

Since there is a non-negligible difference between the description of the supermolecular and embedded schemes, we have investigated the correlation between  $\bar{r}[\text{Mo-O}]$  and the values of the differences of  $\sigma_{iso}$  obtained from embedding calculations and the supermolecular ones, denoted by  $\Delta$  in Figure 3. From these results, we see that there is in fact no correlation between the two variables ( $R^2 \leq 0.214$ ). We also observe that there is a much larger dispersion of the data for the mechanical embedding than for the electronic embedding.

These  $\Delta$  values are a measure of the error of a given embedding approximation with respect to the supermolecular standard, which therefore cannot be attributed to the molecular geometry represented by the  $\bar{r}[\text{Mo-O}]$  descriptor. It remains to be tested whether this error can be associated with other structural descriptors (for instance, ones taking into account the asymmetry around the Mo nucleus in a given environment) or whether it should be attributed to more subtle effects, for instance due to spin-orbit coupling, which we discuss in the following.

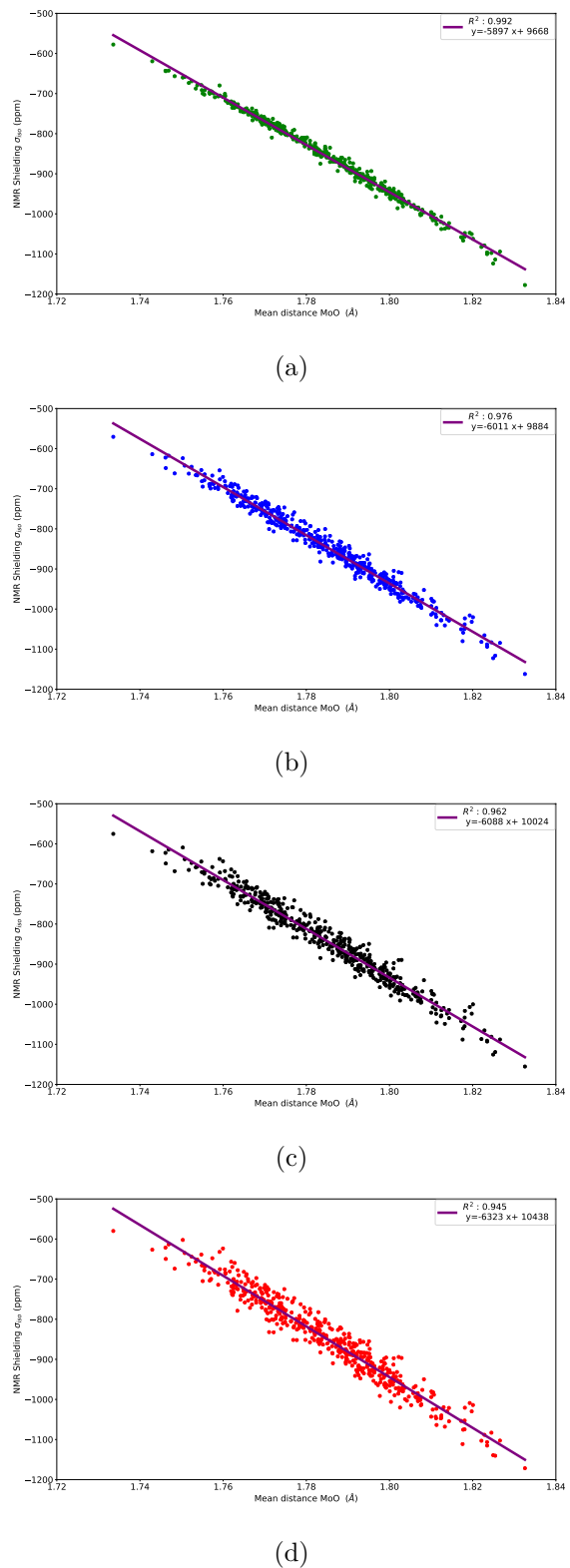


FIG. 2: Correlations between  $\sigma_{iso}$  (in ppm) and the mean distance between Mo and the oxygen atoms in the Molybdate ion ( $\bar{r}[\text{Mo-O}]$ , in Å), obtained from SR-ZORA calculations for various models: (a) isolated  $[A|0|0]$ ; (b) FDE  $[A|0|20B_f]$ ; (c) FnT  $[A|11B_f|9B_f]$ ; and (d) Supermolecule  $[A + 20B|0|0]$  for all snapshots.

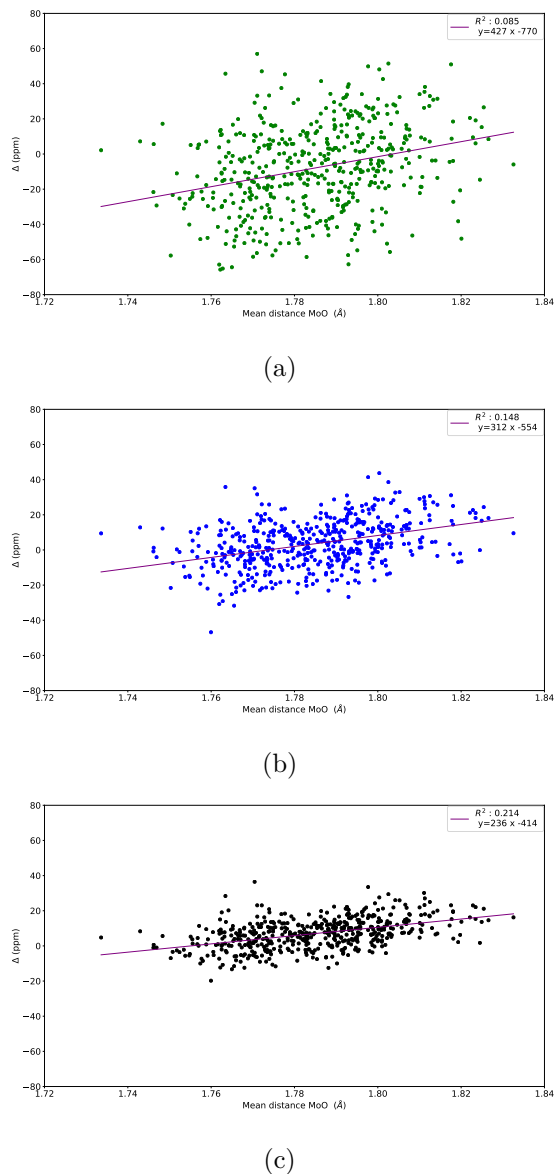


FIG. 3: Difference between the embedded models (a) isolated  $[A|0|0]$ ; (b) FDE  $[A|0|20B_f]$ ; and (c) FnT  $[A|11B_f|9B_f]$ , and the supermolecule SR-ZORA results ( $\Delta$ , in ppm) as a function of the mean distance between Mo and the oxygen atoms in the Molybdate ion ( $\bar{r}[\text{Mo-O}]$ ).

## 2. Assessment of spin-orbit Hamiltonians

In order to assess the importance of spin-orbit interactions on the the solvent effects, we begin by considering the SO-ZORA Hamiltonian. The  $\bar{\sigma}_{iso}$  values and standard deviations for supermolecular and embedded calculations are shown in Table II.

From these we observe that there are no major differences with respect to the SR-ZORA results (see Table I) if it comes to the magnitude of the solvent effects for all

Model	$\bar{\sigma}_{iso} \pm std$	$\Delta\bar{\sigma}_{iso}$ wrt. $[A + 20B 0 0]$
$[A + 20B 0 0]$	$-585.0 \pm 109.3$	-
$[A 11B_f 9B_f]$	$-581.9 \pm 104.2$	$3.1 \pm 8.8$
$[A 0 20B_f]$	$-585.8 \pm 102.1$	$-0.8 \pm 13.7$
$[A 0 0]$	$-600.3 \pm 99.1$	$-15.3 \pm 24.7$

TABLE II: Mean values for the isotropic NMR shielding constant ( $\bar{\sigma}_{iso}$ , in ppm) for the different computational models, and the respective standard deviations ( $std$ ), obtained from the complete set of CPMD snapshots, employing the SO-ZORA Hamiltonian.  $\Delta\bar{\sigma}_{iso}$  are the deviations with respect to  $[A + 20B|0|0]$ .

approaches. We do observe some differences in the error patterns (more readily grasped by inspecting the values of  $\Delta\bar{\sigma}_{iso}$ ), with both electronic embedded models now being closer to the supermolecular  $\bar{\sigma}_{iso}$  values than the mechanical embedding case, though once more the FDE  $\bar{\sigma}_{iso}$  is the closest to the supermolecular one. The standard deviations in each embedded model show the same pattern as in the SR-ZORA, with FnT yielding the smallest and mechanical embedding the largest  $std$  values.

Comparing the results in Tables I and II we also observe significant changes on the absolute  $\bar{\sigma}_{iso}$  values, which for SO-ZORA are upshifted by around 260 ppm from the SR-ZORA values. Such a spin-orbit effect on the absolute NMR shielding values is well-known and extensively discussed in the literature [82, 83], and therefore we refer the readers to the review [83] and recent examples [82, 84–87].

The nearly identical behavior of embedded SR-ZORA and SO-ZORA can be seen, first, as a manifestation of the fact that, since there are no significant spin-orbit coupling effects on the electronic structure of the solvent, the environment electron densities and electrostatic potentials will be rather similar in the two sets of calculations (SR and SO), as we can see from the rather similar standard deviations for all models considered. Second, since the Mo atom is not very heavy, spin-orbit coupling is not expected to bring about qualitative differences on the electron density of the  $\text{MoO}_4^{2-}$  ion.

Model	$\bar{\sigma}_{iso} \pm std$
$[A + 20B 0 0]^{216}$	$-845.7 \pm 45.4$
$[A 11B_f 9B_f]^{216}$	$-839.1 \pm 45.2$
$[A 11B_f 9B_f]_{Dirac}^{216}$	$-428.3 \pm 45.5$

TABLE III: Mean values for the isotropic NMR shielding constant ( $\bar{\sigma}_{iso}$ , in ppm) for the different computational models, and the respective standard deviations ( $std$ ), obtained from a subset of 216 CPMD snapshots chosen following the analysis of the normal probability plot (see supplemental information for details). The 4-component embedding calculations ( $[A|11B_f|9B_f]_{Dirac}^{216}$ ) are performed using embedding potentials and frozen densities obtained from the SR-ZORA  $[A|11B_f|9B_f]^{216}$  model.



An appealing feature of the SR-ZORA and SO-ZORA calculations is their relatively modest computational cost (respectively, around 3 and 8 hours CPU time per snapshot to calculate all the three embedded models and the supermolecular system, with the latter taking about one order of magnitude longer than each of the embedding calculations) which has allowed us to consider a rather large number of snapshots. This makes them ideal tools for exploring the importance of relativistic effects across the periodic table while considering a great number of configuration from molecular dynamics or the effect of increasing the size of the active space. As discussed above, this allowed us to verify that SR-ZORA calculations already capture the solvent effects for the molybdate case.

The exploration of a large number of snapshots also provided data with which to guide us in reducing the number of CPMD snapshots used for the more expensive calculations employing the DC Hamiltonian (about forty times more expensive than the ZORA-SO calculations in the current example). With the help of normal probability plots (see supplemental information), we have been able to use the SR-ZORA data to reduce the number of snapshots to 216, roughly half of the original sample space, while at the same time assuring to keep a normal distribution of  $\sigma_{iso}$  values for the SR-ZORA (supermolecule and embedded) calculations. We subsequently used this reduced sample to perform the DC calculations, for which results are shown in Table III.

While we cannot directly compare the standard deviations from the DC results to those of Tables I and II we see from the table that the DC standard deviation values are close to the SR-ZORA ones, calculated with the same reduced sample of snapshots, and therefore we expect to see a similar error behavior for DC as for ZORA, had we used the original sample size. Furthermore, as was the case when improving the ZORA Hamiltonian with the addition of spin-orbit coupling, the more accurate treatment of relativistic effects afforded by the DC Hamiltonian introduces a further upshift of about 150 ppm with respect to SO-ZORA mean difference:  $\bar{\sigma}_{iso}$ , so that the difference of the means between the SR-ZORA and DC results is now a little over 410 ppm.

This is a significant difference, given the growing interest in recent years in determining absolute shielding scales. Taking into account the steep computational cost of performing 4-component supermolecular calculations over hundreds of structures from MD simulations, we consider electronic embedding to be a promising way to calculate absolute NMR shielding constants in complex environments.

### B. Larger active subsystems

The results above suggest that the minimal structural model for solvated molybdate (comprising only the anion) is a rather good one for obtaining NMR shieldings in solution. In this section, we explore the effect of ex-

plicitly including nearest neighbor water molecules in the calculation.

As we have found little quantitative differences in the behavior of the embedding approaches for the different Hamiltonians, and that it is computationally very expensive to perform a systematic expansion of the active subsystem while averaging over hundreds of snapshot, here we have decided to focus on selected SR-ZORA calculations.

We therefore take two snapshots (numbered 10 and 85), representing respectively structures yielding NMR shieldings close to the  $\bar{\sigma}_{iso}$  value and towards the tails of the distribution. Figure 4, in which we compare embedding models with increasing number of water molecules in the active subsystem, summarizes our results. We only show results for up to 11 water molecules added to the active subsystem, due to the fact that no significant qualitative changes occur beyond this number.

Starting with now waters in the active subsystem ( $p = 0$ ), for snapshot 10 ( $\sigma_{iso}$  closer to  $\bar{\sigma}_{iso}$ ) we observe significant differences between embedding methods, with the subsystem embedding model ( $[A+pB|(11-p)B_f|9B_f]$ ) yielding results which are only a few ppm below the reference supermolecular calculations. The FDE model ( $[A+pB|0|(20-p)B_f]$ ) introduces more significant errors (over 10 ppm underestimation with respect to supermolecule), which from prior findings we can attribute to the importance of relaxing the environment species around a highly charged active subsystem. The mechanical embedding model fares the worst (over 40 ppm underestimation to supermolecule).

Adding one water molecule to the active subsystem ( $p = 1$ ) greatly improves the results for all embedding models, and gets the electronic embedding ones in very close agreement to the reference. From two ( $p = 2$ ) to four ( $p = 4$ ) explicit waters, however, we see a degradation of the electronic embedding results, which now overestimate  $\sigma_{iso}$ . At around six explicit water molecules ( $p = 6$ ), which roughly corresponds to the first solvation shell around the molybdate ion [12], the electronic embedding methods have converged to the reference result, and show no further significant variations. The mechanical embedding model, on the other hand, shows small errors but does not show converged results even after 11 explicit water molecules ( $p = 11$ ).

We note that a similar analysis has been carried out for 4-component calculations for snapshot 10, but restricting the reference calculations to 6 water molecules due to constraints in our computational resources (see computational details and supplemental information). As we observe the same trends in the DC calculations as in the SR-ZORA ones, we do not discuss the former explicitly.

For snapshot 85 ( $\sigma_{iso}$  away from  $\bar{\sigma}_{iso}$ ) we also observe the embedding approaches are pretty much converged to the reference after the first solvation shell is explicitly included. However, the behavior is quite different from snapshot 10 for the smaller number of explicitly included waters. Without any water molecule in the active sub-

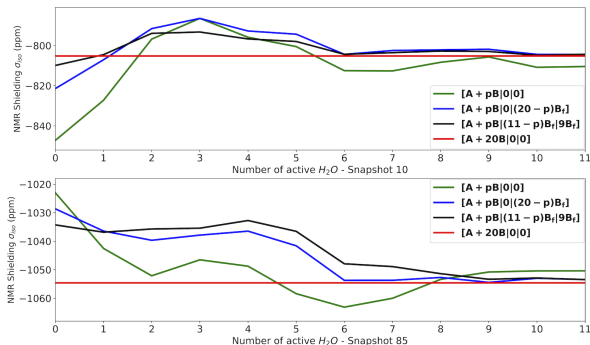


FIG. 4: Variations of  $\sigma_{iso}$  (in ppm) with increasing number of water molecules in the active subsystem for selected structures (top: snapshot 10; bottom: snapshot 85). All values are computed with the SR-ZORA Hamiltonian.

system, all methods underestimate  $\sigma_{iso}$ , and the mechanical embedding model is the worst among the three, but as the number of explicitly included water molecules increases, the electronic embedding methods do not show any improvement until that number reaches five.

The analysis of the snapshots in Figure 4 provides a first, but somewhat indirect, indication that electronic embedding approaches are more reliable than mechanical embedding for two key structural models: the embedded  $\text{MoO}_4^{2-}$ , and the  $[\text{MoO}_4(\text{H}_2\text{O})_6]^{2-}$  species. For a partial first solvation shell, it is nevertheless difficult to comprehend what is actually taking place from just following the  $\sigma_{iso}$  values.

### C. Real-space analysis of embedding

One can try to identify the underlying differences between embedding and supermolecular calculations by investigating the systems in real space. One option is to follow the differences in the electron density for the two treatments, as often done for other properties and previously done by Bulo and coworkers [24] for NMR shieldings. For convenience in the analysis, we decided to deviate slightly from the models discussed above and collected a number of water molecules closest to the molybdate ion into a single subsystem, while any molecules further away are still considered as individual subsystems. While such a modification appears to have little effect on the calculated shieldings (see Table S1 in the supplemental information), we discuss below its effect on density.

Figure 5 presents the difference in SR-ZORA density between the reference and the electronic embedding models ( $[A|11B_f|9B_f]$  and  $[A|0|20B_f]$ ). It is important to note here that this figure does not present density isosurfaces, but rather a volumetric description (i.e. the accumulated values ranging from lower (0.00) to the upper (0.01) bound considered) of the electron density. This

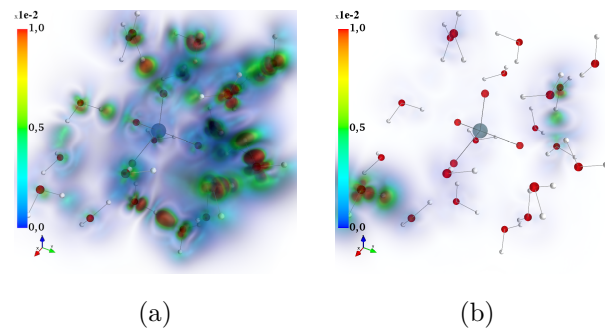


FIG. 5: Volumetric plot of the difference of SR-ZORA electron densities between (a)  $[A|11B_g|9B_f]$  and  $[A + 20B|0|0]$ ; and (b)  $[A|11B_f|9B_f]$  and  $[A|11B_g|9B_f]$  for snapshot 10. Displayed are difference density values at each grid point between 0.00 and 0.01 a.u.

representation has the advantage of enabling the visualization of high and low density regions in the same plot.

From Figure 5 we clearly see that inaccuracies in the embedding calculations, due to (for instance) the approximate treatment of the non-additive kinetic energy term, introduce small errors in the densities throughout the whole space. These errors tend to build up (green to red colors) in the regions away from the core of the  $\text{MoO}_4^{2-}$  unit, and for the  $[A|11B_f|9B_f]$  model they tend to be more significant around the second solvation shell, given that first solvation shell is fully relaxed in the presence of  $\text{MoO}_4^{2-}$  (Figure 5-a). We also see from that figure that there is a small but non-negligible effect on the densities of the environment when we consider the molecules as a single subsystem or as a collection of fragments—as (Figure 5-b), in the latter, we create additional frontiers between subsystems, due to the monomer expansion and the eventual buildup of errors due to the approximate kinetic energy functionals.

Whatever the case, the analysis of the electron density is at odds with our findings for the mean values of shielding, since there appears to be nothing that points to significant errors in the region of the Mo-O bonds. This suggests that trying to understand the behavior  $\sigma_{iso}$  values from changes in the electron density is not a suitable strategy.

An alternative to the visualization of the electron density is the analysis of the property densities. We have provided a first example for such an analysis for the analysis of embedding in the description of our 4-component implementation of subsystem embedding theory for magnetic properties, but for rather small systems [27]. Here we provide the analysis of a more complex example, while at the same time forsaking the use of isosurfaces in favor of volumetric plots.

Because the analysis of shielding densities is only implemented in the DIRAC code, we have restricted ourselves to the  $[\text{MoO}_4(\text{H}_2\text{O})_p]^{2-}$  species ( $p = 1 - 6$ ). These species are then considered as our references, and both mechanical and electronic embedding calculations con-

sidering only the molybdate ion in the active subsystem have been performed for each value of  $p$ . All fragments have therefore been treated with the DC Hamiltonian, with the embedding potentials being determined with *freeze-thaw* calculations using DIRAC. The results of these calculations are found in Figure 6, where we present the shielding density for the  $^{95}\text{Mo}$  atom on the references, along with the differences in shielding density between the references and the electronic and mechanical embedding results. For the embedding calculations, contributions from the solvent water molecules at the location of the Mo atom are obtained with the NICS procedure.

The first striking feature of the embedding results is that, unlike for the electron densities, the error in the embedding calculations is indeed localized within the  $\text{MoO}_4^{2-}$  species, in particular around the Mo atom. In addition to that, there are smaller errors also at the positions of the molybdate oxygen atoms.

We believe this points to a physical process that, while involving a rather local operator (the hyperfine operator has an effective  $r^{-2}$  dependence), makes the resulting property  $\sigma_{iso}$  (described as the cross product of the hyperfine and the magnetically induced current density) have a much less localized nature, and with that non-negligible contributions rather far from the responding atom (in the vicinity of the oxygen atoms) arise.

This goes to explain why, in both types of embedding, the dependence of  $\sigma_{iso}$  on the number of water molecules is so significant: while the water molecules themselves do not contribute significantly to the shielding density, their absence (in the case of mechanical embedding) or the relatively inaccurate representation of the active and environment subsystems (in the case of electronic embedding) is enough to perturb the contribution to the shielding density around the molybdate oxygen atoms.

Furthermore, we see that as the first solvation shell is built up by including the nearest water molecules to the active subsystem (we show in Figure 6 only the even  $p$  values, see supplemental information for the odd values), the perturbation on the molybdate oxygen atoms is not accounted for in a systematic manner by the embedding methods, so that errors may build up even in regions in which water molecules had already been added. We do not yet possess the analysis tools to fully understand the interplay between these different effects, and are currently pursuing the development of new analysis approaches to address the issue.

Beyond these similarities, we nevertheless see that the errors for the mechanical embedding calculations, though of similar magnitude than those for the electronic embedding, extend much farther than for the latter, and explain why electronic embedding is more reliable.

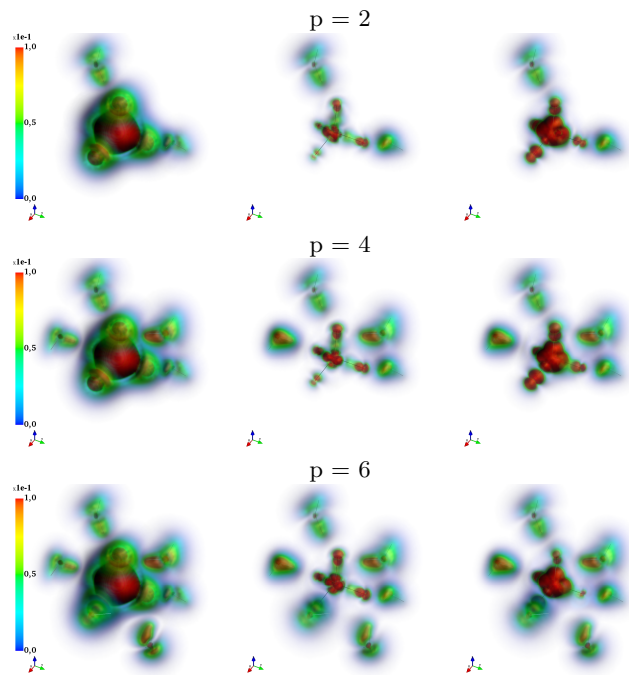


FIG. 6: Volumetric plots of shielding densities and differences in shielding densities between 0.0 and 0.1 ppm, computed with the DC Hamiltonian, for  $\text{MoO}_4(\text{H}_2\text{O})_p]^{2-}$  with increasing number of  $p$  water molecules. Left column: Supermolecule ( $[A + pB|0|0]$ ); Middle column: difference between supermolecular and electronic embedding ( $[A|6B_g|0]$ ) calculations; Right column: difference between supermolecular and mechanical embedding ( $[A|0|0] + [0|6B_g|0]$ ) calculations. For embedded models, water molecule contributions to  $^{95}\text{Mo}$  shielding densities are calculated with the NICS method.

## V. CONCLUSIONS AND PERSPECTIVES

In this manuscript we provide an assessment of approaches based upon the frozen density embedding (FDE) framework for the creation of computationally efficient models capable of capturing solvent effects on NMR shieldings of heavy elements. We have investigated how water influences the shielding of  $^{95}\text{Mo}$  in the molybdate dianion ( $\text{MoO}_4^{2-}$ ), which can be at times a reference for  $^{95}\text{Mo}$  NMR experiments, or a precursor for building Mo-containing materials.

A particular strength of FDE in this context is the ease with which different relativistic Hamiltonians can be used for different parts of the system. Here it allowed to compare the scalar relativistic (SR) and 2-component spin-orbit (SO) ZORA Hamiltonians, and the 4-component Dirac-Coulomb (DC) Hamiltonian to describe the active subsystem while using the SR ZORA Hamiltonian for the aqueous environment.

From our calculations on different structural mod-

els, employing a large set of structures obtained from CPMD trajectories from work previously described in the literature, we have established that for the  $^{95}\text{Mo}$  mean isotropic shielding value, there are weak but non-negligible solvent shifts, due to the modification of the response of the electronic wavefunction of solute by the solvent. We find that even an approach along the lines of mechanical embedding yields a mean isotropic shielding value in good agreement with reference results obtained with standard DFT calculations on the full system.

Such mechanical embedding calculations, however, show a much larger spread of shielding values (for the different CPMD snapshots) with respect to the reference calculations than either of the electronic embedding variants considered. This leads us to conclude that electronic embedding is in fact an essential component for computational models.

By investigating the dependence of the isotropic shielding value from embedding calculations on the size of the active subsystems for selected CPMD snapshots, we have observed that convergence to the reference calculations is not monotonic and large discrepancies may still occur until roughly a full first solvation shell is explicitly included in the active subsystem [22].

We have found our results to be stable with respect to changes in the way the environment was described by FDE, that is, irrespective of whether we used embedding potentials obtained with the same (SR-ZORA or SO-ZORA) Hamiltonian throughout, or by using a combination of SR-ZORA based environment electron densities and electrostatic potentials together with a DC based calculation for the active subsystem.

This points to the possibility of devising efficient computational models for Mo-containing compounds, in which SR-ZORA DFT calculations are used to prepare embedding potentials for more sophisticated, 4-component based calculations on the Mo-containing regions of interest. The latter approach is particularly interesting to investigate the absolute shieldings for molecules in complex environments.

We have found a significant shift (around 260 ppm) when changing from the SR-ZORA to the SO-ZORA Hamiltonian, and another significant shift (around 150 ppm) when changing from the SO-ZORA to the DC Hamiltonians. However, given the rather systematic nature of the differences observed, for chemical shifts of Mo complexes it is likely that the SR-ZORA Hamiltonian provides sufficient accuracy.

Finally, we have explored the visualization of the electron and shielding densities, as a means to provide further insight on the mechanisms behind the solvent effects, in addition to the performance (and shortcomings) of the embedding approaches. Our results underscore the point

we have made previously, in that changes in shielding densities are better understood from the analysis of the differences of the corresponding shielding densities rather than from the changes in electron density.

Furthermore, we observe that the major source of errors in embedding calculations comes from the regions around the Mo nucleus, with less important discrepancies coming from all around the  $\text{MoO}_4^{2-}$  species. As the major difference between the FDE and reference calculations lies in the use of approximate kinetic energy functionals to describe the respective non-additive contributions, this suggests that for magnetic properties it would be of interest to try to improve the performance of such functionals.

## VI. ACKNOWLEDGEMENTS

The authors thank Dr. Jérôme Cuny (Laboratoire de Chimie et Physique Quantiques, UMR 5626, Toulouse) for kindly providing us with the snapshots from the CPMD simulations on the molybdate ion.

The members of the PhLAM laboratory acknowledge support from the CaPPA project (Chemical and Physical Properties of the Atmosphere), funded by the French National Research Agency (ANR) through the PIA (Programme d’Investissement d’Avenir) under contract “ANR-11-LABX-0005-01” and I-SITE ULNE project OVERSEE (ANR-16-IDEX-0004), as well as by the Ministry of Higher Education and Research, Hauts de France council and European Regional Development Fund (ERDF) through the Contrat de Projets Etat-Region (CPER) CLIMBIO (Changement climatique, dynamique de l’atmosphère, impacts sur la biodiversité et la santé humaine), the CNRS Institute of Physics (INP) via the PICS program (grant 6386), and computational time provided by the French national supercomputing facilities (grants DARI A0030801859, A0050801859).

MO acknowledges support from the Polish National Science Centre (NCN) (grant number 2016/23/D/ST4/03217).

## CONFLICT OF INTEREST

There are no conflicts to declare.

## SUPPORTING INFORMATION AVAILABLE:

The supporting information is available as appendix on page 14.

[1] G. A. Aucar and I. A. Aucar, in *Annual Reports on NMR Spectroscopy*, Vol. 96 (Academic Press, 2019) pp. 77–141.

[2] K. Jackowski and W. Makulski, *Magn. Reson. Chem.* **49**, 600 (2011).

- [3] M. Jaszuński, A. Antušek, P. Garbacz, K. Jackowski, W. Makulski, and M. Wilczek, *Progress in Nuclear Magnetic Resonance Spectroscopy* **67**, 49 (2012).
- [4] D. Grekov, T. Vancompernelle, M. Taoufik, L. Delevoye, and R. M. Gauvin, *Chem. Soc. Rev.* **47**, 2572 (2018).
- [5] J. Cuny, S. Cordier, C. Perrin, C. J. Pickard, L. Delevoye, J. Trébosc, Z. Gan, L. L. Pollès, and R. Gautier, *Inorganic Chemistry* **52**, 617 (2013), <https://doi.org/10.1021/ic301648s>.
- [6] T. Charpentier, *Solid State Nuclear Magnetic Resonance* **40**, 1 (2011).
- [7] G. J. O. Beran, J. D. Hartman, and Y. N. Heit, *Accounts of Chemical Research* **49**, 2501 (2016), <https://doi.org/10.1021/acs.accounts.6b00404>.
- [8] A. S. P. Gomes and C. R. Jacob, *Annual Reports Section "C" (Physical Chemistry)* **108**, 222 (2012).
- [9] C. R. Jacob and J. Neugebauer, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **4**, 325 (2014).
- [10] T. A. Wesolowski, S. Shedje, and X. Zhou, *Chem. Rev.* **115**, 5891 (2015).
- [11] Q. Sun and G. K.-L. Chan, *Acc. Chem. Res.* **49**, 2705 (2016).
- [12] T. T. Nguyen, J. Jung, X. Trivelli, J. Trébosc, S. Cordier, Y. Molard, L. L. Pollès, C. J. Pickard, J. Cuny, and R. Gautier, *Inorg. Chem.* **54**, 7673 (2015).
- [13] R. R. Schrock, *Angew. Chem., Int. Ed.* **47**, 5512 (2008).
- [14] Y. H. Jang and W. A. Goddard III, *Top. Catal.* **15**, 273 (2001).
- [15] N. Merle, F. Le Quémener, Y. Bouhoute, K. C. Szeto, A. De Mallmann, S. Barman, M. K. Samantaray, L. Delevoye, R. M. Gauvin, M. Taoufik, and J.-M. Bassett, *Journal of the American Chemical Society* **139**, 2144 (2017), <https://doi.org/10.1021/jacs.6b11220>.
- [16] G. Wallez, P. E. Raison, A. L. Smith, N. Clavier, and N. Dacheux, *Journal of Solid State Chemistry* **215**, 225 (2014).
- [17] K. Zhang, J. He, R. Shen, X. Sun, Q. Ouyang, Y. Chen, Y. Gao, and W. Ji, *Optical Materials* **99**, 109570 (2020).
- [18] R. J. Hand, R. J. Short, S. Morgan, N. C. Hyatt, G. Möbus, and W. E. Lee, *Glass Technology* **46**, 121 (2005).
- [19] H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.* **48**, 1198 (2009).
- [20] Q. Cui, , and M. Karplus, *J. Phys. Chem. B* **104**, 3721 (2000).
- [21] A. Frank, H. M. Möller, and T. E. Exner, *J. Chem. Theory Comput.* **8**, 1480 (2012).
- [22] J. Fukal, O. Páv, M. Buděšínský, I. Rosenberg, J. Šebera, and V. Sychrovský, *Physical Chemistry Chemical Physics* **21**, 9924 (2019).
- [23] C. R. Jacob and L. Visscher, *J. Chem. Phys.* **125**, 194104 (2006).
- [24] R. E. Bulo, C. R. Jacob, and L. Visscher, *J. Phys. Chem. A* **112**, 2640 (2008).
- [25] C. R. Jacob, J. Neugebauer, and L. Visscher, *J. Comput. Chem.* **29**, 1011 (2008).
- [26] A. W. Götz, J. Autschbach, and L. Visscher, *J. Chem. Phys.* **140**, 104107 (2014).
- [27] M. Olejniczak, R. Bast, and A. S. P. Gomes, *Phys. Chem. Chem. Phys.* **19**, 8400 (2017).
- [28] P. Pyykkö, *Chem. Rev.* **88**, 563 (1988).
- [29] K. G. Dyall and J. Knut Faegri, *Introduction to Relativistic Quantum Chemistry* (Oxford University Press, 2007).
- [30] M. Reiher and A. Wolf, *Relativistic Quantum Chemistry. The Fundamental Theory of Molecular Science* (Wiley, 2009).
- [31] J. Autschbach, *J. Chem. Phys.* **136**, 150902 (2012).
- [32] M. Kaupp, M. Bühl, and V. G. Malkin, *Calculation of NMR and EPR Parameters: Theory and Applications* (Wiley-VCH, 2004).
- [33] J. Autschbach, *Phil. Trans. R. Soc. A* **372**, 20120489 (2014).
- [34] R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).
- [35] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, and K. Ruud, *Chem. Rev.* **112**, 543 (2012).
- [36] D. H. Whiffen, *Pure Appl. Chem.* **50**, 75 (1978).
- [37] T. Saue and H. J. A. Jensen, *J. Chem. Phys.* **118**, 522 (2003).
- [38] T. Saue, *Chemphyschem : a European journal of chemical physics and physical chemistry* **12**, 3077 (2011).
- [39] U. von Barth and L. Hedin, *J. Phys. C Solid State Phys.* **5**, 1629 (1972).
- [40] A. K. Rajagopal and J. Callaway, *Phys. Rev. B* **7**, 1912 (1973).
- [41] C. R. Jacob and M. Reiher, *Int. J. Quant. Chem.* **112**, 3661 (2012).
- [42] R. Bast, H. J. A. Jensen, and T. Saue, *Int. J. Quant. Chem.* **109**, 2091 (2009).
- [43] S. Komorovský, M. Repiský, O. L. Malkina, V. G. Malkin, I. M. Ondík, and M. Kaupp, *J. Chem. Phys.* **128**, 104101 (2008).
- [44] M. Olejniczak, R. Bast, T. Saue, and M. Pecul, *J. Chem. Phys.* **136**, 014108 (2012), erratum: *ibid.* **136**, 239902, (2012).
- [45] A. Goez and J. Neugebauer, in *Frontiers of Quantum Chemistry*, edited by M. J. Wojcik, H. Nakatsuji, B. Kirtman, and Y. Ozaki (Springer Singapore, Singapore, 2018) Chap. Embedding, pp. 139–179.
- [46] T. A. Wesolowski and A. Warshel, *J. Phys. Chem.* **97**, 8050 (1993).
- [47] T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.* **248**, 71 (1996).
- [48] T. A. Wesolowski and J. Weber, "Recent advances in density functional methods," (World Scientific, 2002) Chap. 24, p. 371.
- [49] S. Höfener, A. S. P. Gomes, and L. Visscher, *J. Chem. Phys.* **136**, 044104 (2012).
- [50] T. A. Wesolowski and A. Warshel, *The Journal of Physical Chemistry* **97**, 8050 (1993).
- [51] C. R. Jacob and J. Neugebauer, *Wiley Interdiscip Rev Comput Mol Sci* **4**, 325 (2014).
- [52] D. Schlüns, K. Klahr, C. Mück-Lichtenfeld, L. Visscher, and J. Neugebauer, *Phys. Chem. Chem. Phys.* **17**, 14323 (2015).
- [53] S. Fux, C. R. Jacob, J. Neugebauer, L. Visscher, and M. Reiher, *J. Chem. Phys.* **132**, 164101 (2010), comment: T. A. Wesolowski, *ibid.* **135**, 027101, (2011). Response: S. Fux, C. R. Jacob, J. Neugebauer and M. Reiher, *ibid.* **135**, 027102, (2011).
- [54] D. G. Artiukhin, C. R. Jacob, and J. Neugebauer, *J. Chem. Phys.* **142**, 234101 (2015).
- [55] P. v. R. Schleyer, C. Maerker, A. Dransfeld, H. Jiao, and N. J. Van Eikema Hommes, *J. Am. Chem. Soc.* **118**, 6317 (1996).
- [56] F. London, *J. Phys. Radium* **8**, 397 (1937).

- [57] T. Helgaker, M. Jaszuński, and K. Ruud, *Chem. Rev.* **99**, 293 (1999).
- [58] M. Iliáš, T. Saue, T. Enevoldsen, and H. J. A. Jensen, *J. Chem. Phys.* **131**, 124119 (2009).
- [59] C. J. Jameson and A. D. Buckingham, *J. Chem. Phys.* **73**, 5684 (1980).
- [60] C. J. Jameson and A. Buckingham, *J. Phys. Chem.* **83**, 3366 (1979).
- [61] R. Bast, J. Jusélius, and T. Saue, *Chem. Phys.* **356**, 187 (2009).
- [62] D. Sulzer, M. Olejniczak, R. Bast, and T. Saue, *Phys. Chem. Chem. Phys.* **13**, 20682 (2011).
- [63] J. Hutter, *WIREs Comput. Mol. Sci.* **2**, 604 (2011).
- [64] C. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema, and L. Visscher, *J. Comput. Chem.* **32**, 2328 (2011).
- [65] L. Visscher and K. G. Dyall, *Atomic Data and Nuclear Data Tables* **67**, 207 (1997), see also <http://www.few.vu.nl/visscher/FiniteNuclei/FiniteNuclei.htm>.
- [66] D. Andrae, *Phys. Rep.* **336**, 413 (2000).
- [67] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [68] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **78**, 1396 (1997).
- [69] A. Lembarki and H. Chermette, *Phys. Rev. A* **50**, 5328 (1994).
- [70] E. Van Lenthe, J. G. Snijders, and E. J. Baerends, *J. Chem. Phys.* **105**, 6505 (1996).
- [71] E. Van Lenthe, E. J. Baerends, and J. G. Snijders, *J. Chem. Phys.* **101**, 9783 (1994).
- [72] E. Van Lenthe, R. Van Leeuwen, E. J. Baerends, and J. G. Snijders, *Int. J. Quant. Chem.* **57**, 281 (1996).
- [73] E. Van Lenthe, *J. Chem. Phys.* **110**, 8943 (1999).
- [74] “SCM ADF software for chemistry & materials,” <https://www.scm.com>.
- [75] E. Van Lenthe and E. J. Baerends, *J. Comput. Chem.* **24**, 1142 (2003).
- [76] DIRAC, a relativistic ab initio electronic structure program, Release DIRAC18 (2018), written by T. Saue, L. Visscher, H. J. Aa. Jensen, and R. Bast, with contributions from V. Bakken, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, E. Faßhauer, T. Fleig, O. Fossgaard, A. S. P. Gomes, E. D. Hedegård, T. Helgaker, J. Henriksson, M. Iliáš, Ch. R. Jacob, S. Knecht, S. Komorovský, O. Kullie, J. K. Lærdahl, C. V. Larsen, Y. S. Lee, H. S. Nataraj, M. K. Nayak, P. Norman, G. Olejniczak, J. Olsen, J. M. H. Olsen, Y. C. Park, J. K. Pedersen, M. Pernpointner, R. di Remigio, K. Ruud, P. Sałek, B. Schimmelpfennig, A. Shee, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, S. Villaume, O. Visser, T. Winther, and S. Yamamoto (available at <https://doi.org/10.5281/zenodo.2253986>, see also <http://www.diracprogram.org>).
- [77] K. G. Dyall, *Theor. Chem. Acc.* **117**, 483 (2007).
- [78] T. H. D. Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- [79] J. D. Hunter, *Computing In Science & Engineering* **9**, 90 (2007).
- [80] P. Ramachandran and G. Varoquaux, *Comput. Sci. Eng.* **13**, 40 (2011).
- [81] J. Mason, *Solid State Nucl. Mag.* **2**, 285 (1993).
- [82] P. Lantto, R. H. Romero, S. S. Gómez, G. A. Aucar, and J. Vaara, *Journal of Chemical Physics* **125** (2006), 10.1063/1.2378737.
- [83] M. Jaszuński, A. Rizzo, and K. Ruud, in *Handbook of Computational Chemistry* (Springer International Publishing, Cham, 2017) pp. 497–592.
- [84] F. Alkan and C. Dybowski, *Solid State Nucl. Mag.* **95**, 6 (2018).
- [85] Y. Ootani, H. Yamaguti, H. Maeda, and H. Fukui, *J. Chem. Phys.* **125** (2006), 10.1063/1.2361292.
- [86] V. A. Semenov, D. O. Samultsev, I. L. Rusakova, and L. B. Krivdin, *J. Phys. Chem. A* **123**, 4908 (2019).
- [87] S. T. Holmes and R. W. Schurko, *J. Chem. Theory Comput.* **15**, 1785 (2019).

## Appendix A: Supplemental information

### 1. Influence of the embedding model on $^{95}\text{Mo}$ shielding values

Model	snapshot 10		snapshot 85	
	$\sigma_{iso}$	$\Delta\sigma_{iso}^{ref}$	$\sigma_{iso}$	$\Delta\sigma_{iso}^{ref}$
$[A + 20B 0 0]$ [ <i>ref</i> ]	-805.1		-1054.6	
$[A 11B_f 9B_f]$	-808.9	-3.7	-1034.5	x7 20.1
$[A 11B_g 9B_f]$	-808.9	-3.7	-1035.9	18.7
$[A 0 20B_f]$	-822.9	-17.7	-1028.9	25.7
$[A 0 11B_g + 9B_f]$	-822.4	-17.3	-1032.4	22.2

TABLE S1: SR-ZORA values of  $\sigma_{iso}$  as a function of the fragmentation of the first solvation layer of  $\text{MoO}_4^{2-}$ , and  $\Delta\sigma_{iso}^{ref}$ , with respect to the supermolecule  $[A + 20B|0|0]$  reference, for snapshots 10 and 85.

The electron densities and electrostatic potential for the environment are obtained either from calculations on individual water molecules (and we refer to a “fragmented” environment, noted *f*) or by these forming a single subsystem (and we refer to a “grouped” environment, noted *g*). We have looked at the influence of these two ways of treating the water molecules for two snapshots of the CPMD trajectories, snapshot 10 yielding results close to the average and snapshot 85 standing in the tail of the result distribution.

Table S1 reveals that the  $\sigma_{iso}$  values obtained with FDE and FnT embeddings are marginally sensitive to the way the embedding density is calculated. For the sake of minimizing the computational cost, we will rely on “fragmented” water clusters for the generation of the FnT and FDE embedding potentials.

### 2. Influence of the relativistic Hamiltonian on $\sigma_{iso}$

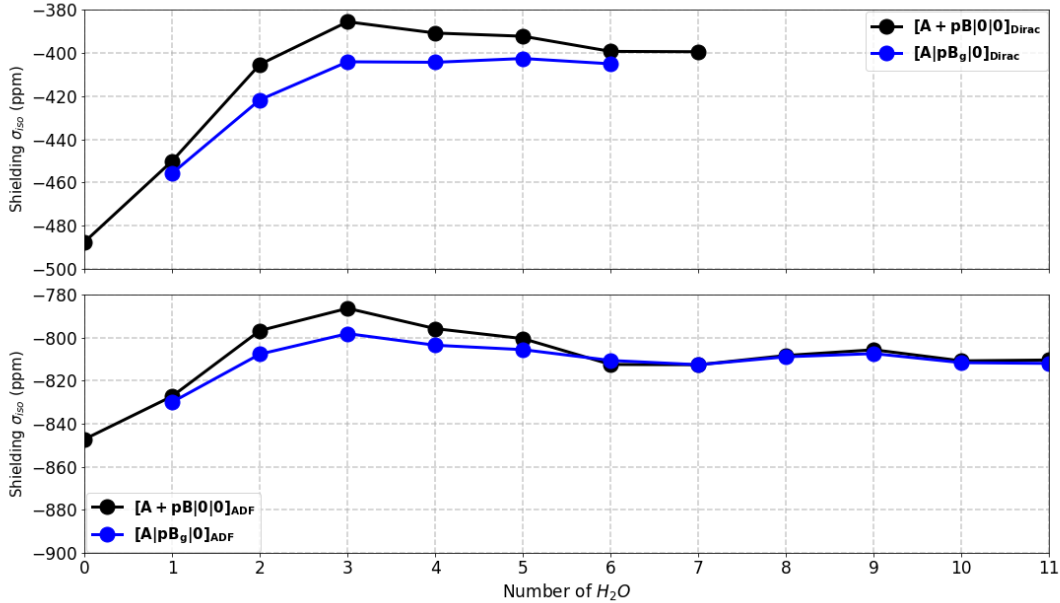


FIG. S1: Evolution of the isotropic shielding  $\sigma_{iso}$  (in ppm) as a function of the number of water molecules  $p$  in the active system ( $[A + pB|0|0]$ ) or in the FnT process ( $[A|pB_g|0]$ ), with Dirac (top) and with ADF (bottom), for snapshot 10.

In the manuscript, we have noted that the change of relativistic Hamiltonian from SR-ZORA (ADF) to DC (DIRAC) translates into a change 410 ppm of the absolute shielding constant for the  $[A|11B_f|9B_f]$  model (See Table III). To

verify whether this shift is independent of the chemical model, we have performed a systematic comparison of  $\sigma_{iso}$  values obtained with SR-ZORA and DC Hamiltonians for molybdate  $[\text{MoO}_4(\text{H}_2\text{O})_p]^{2-}$  hydrated by  $p = 1, 6$  water molecules either treated explicitly ( $[A + pB|0|0]_{ADF}$  and  $[A + pB|0|0]_{Dirac}$ ), or by an embedding potential treating the  $(\text{H}_2\text{O})_p$  water molecules as a “grouped” water cluster ( $[A|pB_g|0]_{ADF}$  and  $[A|pB_g|0]_{Dirac}$ ). The values drawn on Figure S1 reveal very similar trends between the SR-ZORA and DC values. For both Hamiltonians, the difference between FnT ( $[A|pB_g|0]$ ) values and the supermolecular  $[A + pB|0|0]$ , vanishes at  $p = 6$  water molecules.

### 3. Statistical analysis of $\sigma_{iso}$ values; selection of a snapshot subset

Figure S2 plots the distribution of the  $\sigma_{iso}$  values obtained from the 517 snapshots (red dotted line). With the aim of performing DC calculations, which are computationally more expensive than the ADF ones, it was necessary to reduce the number of snapshots, but keeping a normal distribution. The normal probability plots of Figures S3 and S4, reveal that the ADF values, both for supermolecular calculations (dashed red line) and FnT calculations (black line) obtained for a subset of 216 snapshots follow a normal law, though with a smaller variance. The DC values (red lines) have the same statistical distribution, shifted by 411 ppm.

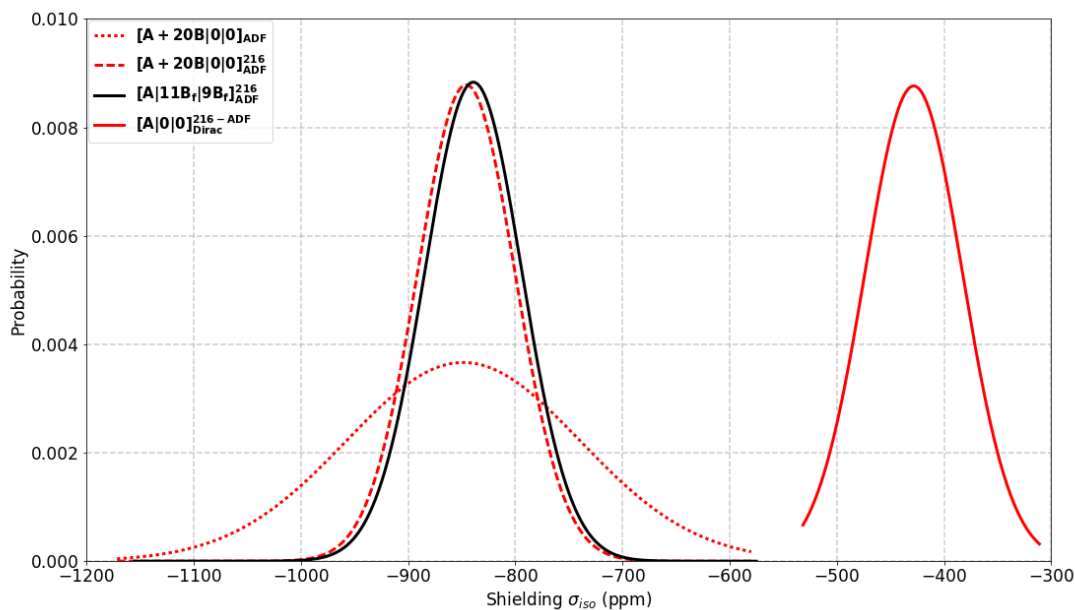


FIG. S2: Normalized Gaussians representing  $\sigma_{iso}$  (in ppm). The red dotted line refers to supermolecular ADF for all snapshots ( $[A + 20B|0|0]_{ADF}$ ), while the red dashed line refers to the subset of 216 selected snapshots ( $[A + 20B|0|0]_{ADF}^{216}$ ). The FnT ADF calculations ( $[A|11B_f|9B_f]_{ADF}^{216}$ ) for this 216 snapshot subset are drawn with a black line, while the DIRAC calculations using SR-ZORA (ADF) embedding potentials and electrostatic potentials are drawn in red.



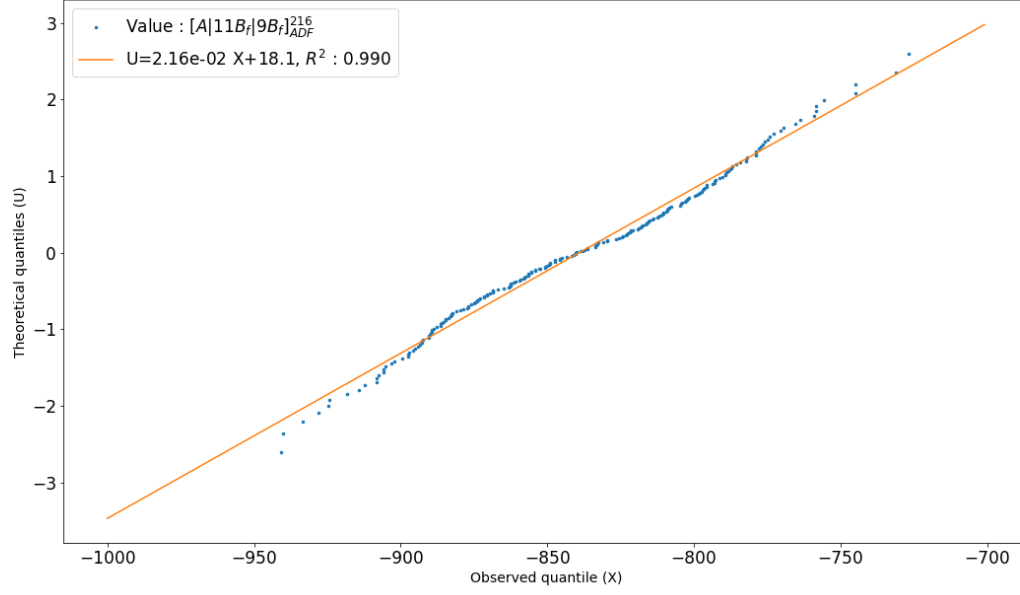


FIG. S3: Normal probability plot of  $\sigma_{iso}$  for the 216 selected snapshots with ADF (FnT) :  $[A|11B_f|9B_f]_{ADF}^{216}$ .

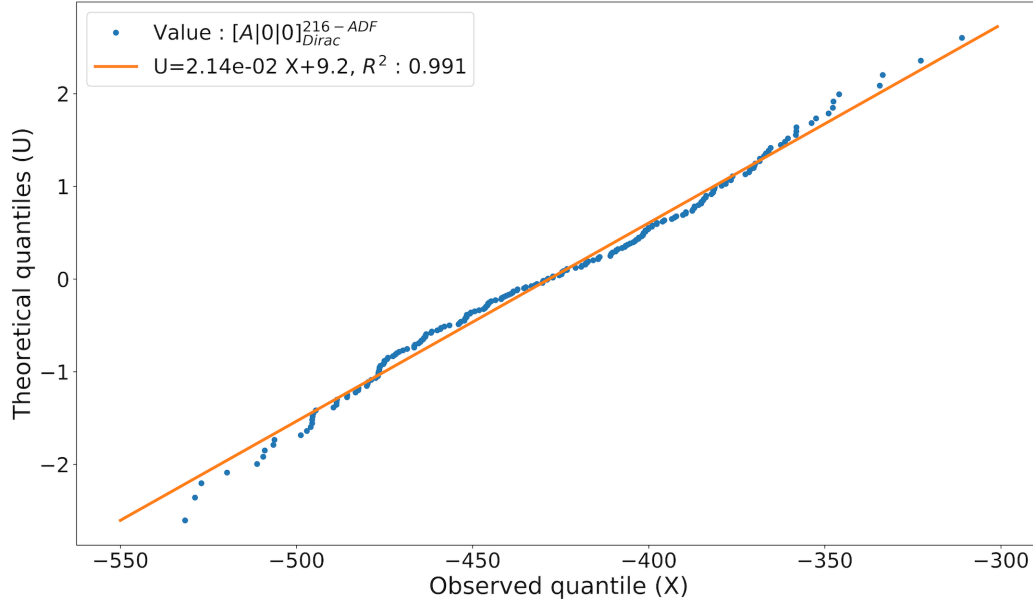


FIG. S4: Normal probability plot of  $\sigma_{iso}$  for the 216 selected snapshots with  $[A|11B_f|9B_f]_{Dirac}^{216} : [A|0|0]_{Dirac}^{216-ADF}$ .

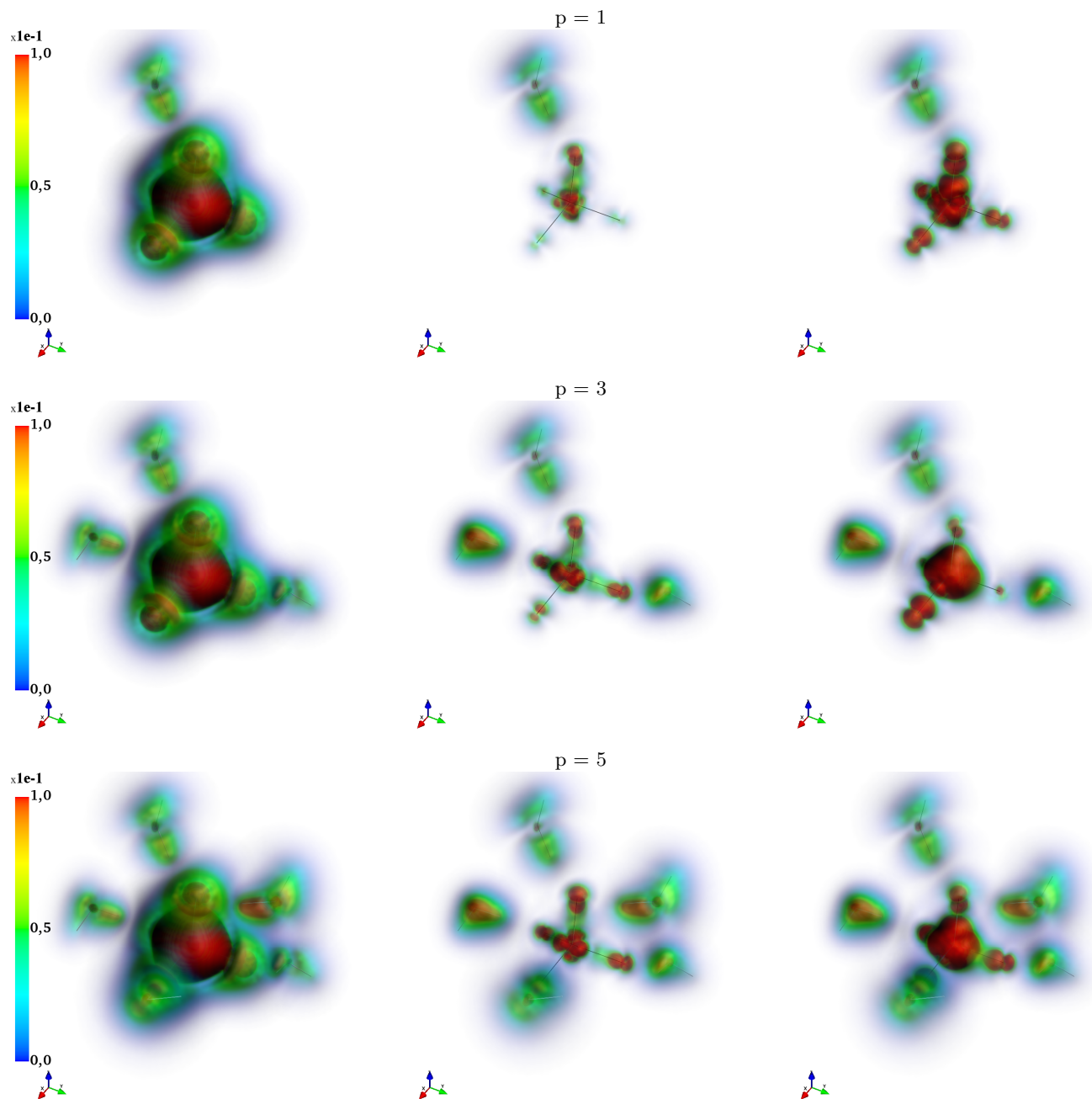


FIG. S5: Volumetric plots of shielding densities and differences in shielding densities between 0.0 and 0.1 ppm, computed with the DC Hamiltonian, for  $[\text{MoO}_4(\text{H}_2\text{O})_p]^{2-}$  with increasing number of  $p$  water molecules ( $p = 1, 3, 5$ ).

Left column: Supermolecule ( $[A + pB|0|0]$ ); Middle column: difference between supermolecular and electronic embedding ( $[A|6B_g|0]$ ) calculations; Right column: difference between supermolecular and mechanical embedding ( $[A|0|0] + [0|6B_g|0]$ ) calculations. For embedded models, water molecule contributions to  $^{95}\text{Mo}$  shielding densities are calculated with the NICS method.