

Supporting Information for:

Extending Libraries of Extremely Localized Molecular Orbitals to Metal Organic Frameworks: a Preliminary Investigation

Erna K. Wieduwilt⁽¹⁾, Giovanni Macetti⁽¹⁾, Rebecca Scatena⁽²⁾,

Piero Macchi^{(3)*}, and Alessandro Genoni^{(1)*}

- (1) Université de Lorraine & CNRS, Laboratoire de Physique et Chimie Théoriques (LPCT), UMR CNRS 7019, 1 Boulevard Arago, F-57078 Metz, France.
- (2) Department of Physics, Clarendon Laboratory, University of Oxford, Parks Road, Oxford, OX1 3PU, United Kingdom.
- (3) Department of Chemistry, Materials and Chemical Engineering, Politecnico di Milano, via Mancinelli 7, 20131 Milano, Italy.

* Correspondence to:

- Alessandro Genoni, Université de Lorraine & CNRS, Laboratoire de Physique et Chimie Théoriques (LPCT), UMR CNRS 7019, 1 Boulevard Arago, F-57078 Metz, France. E-mail: alessandro.genoni@univ-lorraine.fr; Phone: +33 (0)3 72 74 91 70; Fax: +33 (0)3 72 74 91 87.
- Piero Macchi, Department of Chemistry, Materials and Chemical Engineering, Politecnico di Milano, via Mancinelli 7, 20131 Milano. E-mail: piero.macchi@polimi.it; Phone: +39 02-23993023.

S1. Basic theory of Extremely Localized Molecular Orbitals

ELMOs (Extremely Localized Molecular Orbitals) are molecular orbitals strictly localized on small molecular subunits. They can be calculated through the *a priori* definition of a localization scheme that enables the partition of the molecule under investigation into fragments.^{1,2} This subdivision automatically entails the definition of a local basis-set $\beta_i = \{|\chi_{i\mu}\rangle\}_{\mu=1}^{M_i}$ for each subunit. These basis-sets consist only of those basis functions corresponding to the atoms belonging to the fragments and are consequently used to expand the ELMOs of the different subunits. For example, we can express the generic ω -th ELMO for the i -th fragment as follows:

$$|\varphi_{i\omega}\rangle = \sum_{\mu=1}^{M_i} C_{i\mu,i\omega} |\chi_{i\mu}\rangle \quad (\text{S1})$$

Following Stoll,¹ the system under exam is described by means of a single Slater determinant constructed with the ELMOs defined by equation (S1). We obtain the following *ELMO wave function*:

$$|\Psi_{ELMO}\rangle = \frac{1}{\sqrt{(2N)! \det[\mathbf{S}]}} \hat{A} \left[\prod_{i=1}^f \prod_{\omega=1}^{n_i} \varphi_{i\omega} \bar{\varphi}_{i\omega} \right] \quad (\text{S2}).$$

\hat{A} is the usual antisymmetrizer, n_i is the number of occupied ELMOs for the i -th fragment, $\varphi_{i\omega}$ is a spin orbital with spatial part $\varphi_{i\omega}$ and spin part α and $\bar{\varphi}_{i\omega}$ is a spin orbital with spatial part $\varphi_{i\omega}$ and spin part β , while $\det[\mathbf{S}]$ is the determinant of the overlap matrix of the occupied ELMOs (which appears because of the ELMOs non-orthogonality).

ELMOs are calculated by minimizing the energy of the *ELMO wave function* with respect to the coefficients $\{C_{i\mu,i\omega}\}$ in expansion (S1). This leads to a set of modified Hartree-Fock equations (also known as Stoll equations) for each fragment:¹

$$\hat{F}_i |\varphi_{i\omega}\rangle = \varepsilon_{i\omega} |\varphi_{i\omega}\rangle \quad (\text{S3})$$

with \hat{F}_i as the modified Fock operator corresponding to the i -th subunit:

$$\hat{F}_i = (1 - \hat{\rho} + \hat{\rho}_i^\dagger) \hat{F} (1 - \hat{\rho} + \hat{\rho}_i) \quad (\text{S4}),$$

where \hat{F} is the traditional Fock operator, $\hat{\rho}$ is the global density operator depending on all the occupied ELMOs of the system, and $\hat{\rho}_i$ is the density operator for the i -th subunit depending only on the occupied ELMOs of the fragment.

S2. Transferability and rotation of ELMOs

As a consequence of their strict localization, ELMOs are orbitals that can be easily transferred from a molecule to another.^{3,4} In particular, they can be transferred from a model system (which is usually a small molecule on which the orbital is originally determined) to the target system that one wants to study. This can be done by following a strategy originally proposed by Philipp and Friesner,⁵ a strategy that allows the definition of a matrix \mathbf{P} that transforms the ELMO coefficients (see equation (S1)) obtained on the geometry of the model molecule to the ELMO coefficients on the geometry of the target system. To obtain this rotation matrix, it is necessary to define i) a reference frame (\mathbf{a} , \mathbf{c} , \mathbf{d}) in the model molecule and ii) a reference frame (\mathbf{a}' , \mathbf{c}' , \mathbf{d}') in the target system (see Figure S2).

These two reference frames depend on the choice of two atomic triads (one for the model system and one for the target molecule). If the triads for the model and target systems are indicated as (A_1, A_2, A_3) and (A_1', A_2', A_3') , respectively, the vectors defining the two reference frames are: \mathbf{a} (\mathbf{a}'), which is the position vector of A_2 (A_2') relative to A_1 (A_1') (see Figure S2), while \mathbf{c} (\mathbf{c}') and \mathbf{d} (\mathbf{d}') are given by the following vector products:

$$\begin{cases} \mathbf{c} = \mathbf{a} \times \mathbf{b} & (\mathbf{c}' = \mathbf{a}' \times \mathbf{b}') \\ \mathbf{d} = \mathbf{c} \times \mathbf{a} & (\mathbf{d}' = \mathbf{c}' \times \mathbf{a}') \end{cases} \quad (\text{S5})$$

where \mathbf{b} (\mathbf{b}') is the position vector of A_3 (A_3') with respect to A_1 (A_1') (see Figure S2).

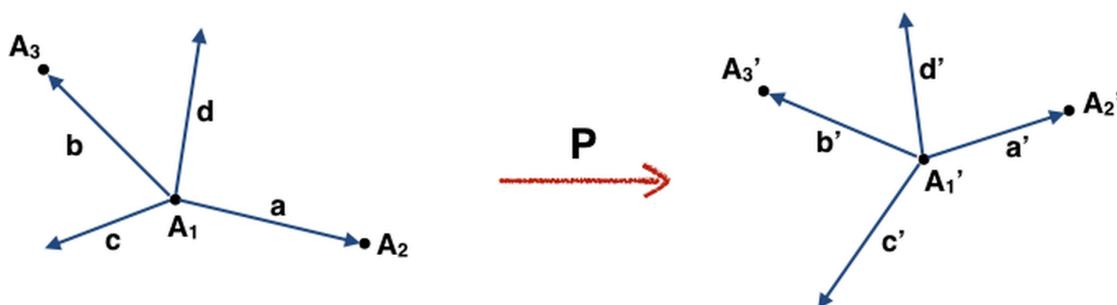


Figure S2. Schematic representation of the reference frames and of the atomic triads that are necessary to define rotation matrix \mathbf{P} allowing the transformation of the ELMOs from the geometry of the model system (left) to the geometry of the target molecule (right).

For ELMOs localized only on one atom (i.e., ELMOs corresponding to core or lone-pair electrons) the atomic triads simply consist of the atom on which the ELMO is localized and usually of two other bonded atoms. For ELMOs localized on two-center bonds, the triads correspond to the atoms forming the bond plus an atom describing the local dissymmetry of the bond under examination.

For ELMOs localized on three centers (e.g., ELMOs that are used to deal with situations in which it is crucial to take into account the delocalized nature of the electronic structure, such as in peptide bonds or aromatic rings), the triads of atoms are automatically selected. For ELMOs localized on more than three atoms, the definition of triads (and of reference frames) that simultaneously take into account the orientation of all the atoms in the subunit is impossible. For this reason, in the current databanks, all the ELMOs are localized at the largest on three atoms (see the next subsection).

Rotation matrix \mathbf{P} that brings from reference frame $(\mathbf{a}, \mathbf{c}, \mathbf{d})$ to reference frame $(\mathbf{a}', \mathbf{c}', \mathbf{d}')$ is actually a matrix that enables to obtain all the rotation matrices for all the types of basis functions and associated ELMO coefficients. In fact, not considering the *s*-type basis functions (which are invariant to rotations because they have spherical symmetry), it is trivial to show that *p*-type atomic orbitals (and the related coefficients) can be rotated according to matrix \mathbf{P} ,³ while basis functions (and corresponding coefficients) with angular momentum greater than 1 can be transformed using matrices that can be constructed as a function of \mathbf{P} .

S3. Supplementary Table

Table S1. Electron density at the bond critical point (in e/bohr^3) and values of the delocalization index (DI) for the Cu-Cu and Cu-O interactions at different spin states as obtained from standard B3LYP and B3LYP/ELMO calculations with different sizes of the QM region on the model system extracted from the HKUST-1 experimental crystal structure with water chemisorption degree of 14.3(6)%.^(a)

Interaction	Spin State	B3LYP/ELMO (small QM region) ^(b)		B3LYP/ELMO (large QM region) ^(c)		Standard B3LYP	
		ρ_{bcp}	DI	ρ_{bcp}	DI	ρ_{bcp}	DI
Cu-Cu	Diamagnetic	0.029	1.031	0.036	0.477	0.036	0.426
	Anti-Ferromagnetic	0.029	0.100	0.035	0.138	0.035	0.146
	Ferromagnetic	0.029	0.100	0.035	0.137	0.035	0.144
Cu-O	Diamagnetic	0.066	0.227	0.094	0.428	0.095	0.429
	Anti-Ferromagnetic	0.065	0.233	0.093	0.463	0.093	0.466
	Ferromagnetic	0.065	0.233	0.093	0.463	0.093	0.466

^(a) All calculations performed with basis-set 6-311G(2d,2p); ^(b) QM region consisting of the Cu-Cu pair;

^(c) QM region comprising the Cu-Cu pair and the four surrounding carboxylate groups.

S3. References for the Supporting Information

- ¹ Stoll, H.; Wagenblast, G.; Preuss, H. On the Use of Local Basis Sets for Localized Molecular Orbitals. *Theor. Chim. Acta* **1980**, *57*, 169–178.
- ² Sironi, M.; Genoni, A.; Civera, M.; Pieraccini, S.; Ghitti, M. Extremely Localized Molecular Orbitals: Theory and Applications. *Theor. Chem. Acc.* **2007**, *117*, 685-698.
- ³ Meyer, B.; Guillot, B.; Ruiz-Lopez, M. F.; Genoni, A. Libraries of Extremely Localized Molecular Orbitals. 1. Model Molecules Approximation and Molecular Orbitals Transferability. *J. Chem. Theory. Comput.* **2016**, *12*, 1052-1067.
- ⁴ Meyer, B.; Guillot, B.; Ruiz-Lopez, M. F.; Jelsch, C.; Genoni, A. Libraries of Extremely Localized Molecular Orbitals. 2. Comparison with the Pseudoatoms Transferability. *J. Chem. Theory. Comput.* **2016**, *12*, 1068-1081.
- ⁵ Philipp, D. M.; Friesner, R. A. Mixed Ab Initio QM/MM Modeling Using Frozen Orbitals and Tests with Alanine Dipeptide and Tetrapeptide. *J. Comput. Chem.* **1999**, *20*, 1468-1494.