Mayer's Effective Minimal Localized Molecular Orbitals

Á. Szabados*[1] and P. R. Surján†[1]

[1]Laboratory of Theoretical Chemistry, Institute of Chemistry, Faculty of Science, ELTE Eötvös Loránd University, Budapest, Hungary

*Email: agnes.szabados@ttk.elte.hu, †Email: peter.surjan@ttk.elte.hu

Abstract

Roughly thirty years ago Mayer presented a theory leading to the concept of an effective minimal atomic basis. A set of a posteriori defined, fragment-localized molecular orbitals (LMOs) was obtained as a byproduct. In this work, a new set of LMOs, termed Mayer's effective minimal LMOs, is derived from this byproduct; it can be viewed as arising from a compactification of atomic orbitals projected onto the occupied space. The procedure is fully self-contained, requiring no input beyond that of the underlying quantum chemical calculation. The resulting LMOs are orientation invariant, computationally economical, and can be partitioned into disjoint subsets assigned to atoms or multiatomic fragments.

Mayer's effective minimal LMOs span the occupied space, but their number exceeds its dimension; they therefore form a frame rather than a basis in the strict sense. The size of the set does not grow with the original AO basis but remains at a value prescribed by the user, a prudent choice being the minimal atomic orbital basis dimension.

Keywords

localized MOs, Magnasco-Perico criterion, projected atomic orbitals, orientation invariance, overcomplete set, frame theory

1 Introduction

In parallel with the widespread use of large, extended basis sets, containing multiple-zeta, polarization and diffuse functions and leading to high-accuracy quantum chemical calculations,

a demand of constructing orbital sets of reduced dimension is also present. The fundamental question of reproducing a large-basis calculation with smaller number of functions has been addressed by several studies. The goal has been often formulated as extracting a minimal set of atomic orbitals (AO), best representing the molecular self consistent field (SCF) solution. This idea led to the effective minimal basis of Mayer,[1,2] the quasiatomic minimal-basis-set orbitals of Ruedenberg and coworkers,[3,4] also developed by Szczepanik and Mrozek[5] or the intrinsic atomic orbitals proposed by Knizia.[6] The latter two were later found closely related.[7] Further studies along this line involve those of Adams,[8] Weinhold,[9,10] Ahlrichs and coworkers,[11,12] Cioslowski et al.,[13] Head-Gordon and coworkers.[14,15] These schemes result a reduced-dimensional set of one-electron orbitals, which are basis functions in the sense that they do not carry occupied/virtual categorization, at the same time they are single atom assigned even if expanded over AOs centered on multiple, neighbouring atoms. When occupied molecular orbitals (MO) are needed, the effective AOs are further transformed, generating e.g. intrinsic bond orbitals,[6] natural localized MOs,[16] etc. Localized molecular orbitals (LMO) derived from a reduced dimensional set of effective AOs proved extremely useful for chemical interpretation.[17–20]

In most applications however, localized MOs (LMO) are obtained without making recourse to a reduced set of effective AOs. From a practical aspect, localized orbitals are essential in linear scaling algorithms,[21–26] embedding[27,28] and fragmentation[29–31] techniques of molecular electronic structure. Apart from the overwhelming Boys,[32] Pipek-Mezey,[33] and Edmiston-Ruedenberg[34,35] schemes, a multitude of alternative ways exist for obtaining LMOs. Strict localization[36,37] and fragment localization[38–42] are important aspects for linear-scaling. Localization schemes are often designed for a particular embedding or divide and conquer approach[43–46] or for computing a particular quantity.[47] At difference with basis functions, localized MOs are either occupied or virtual and not single atom attributable, with the exception of cores and lone pairs. Interatomic/interfragment feature of bonding LMOs is vital and often warrants special adjustments in fragment based approaches.

Presently we introduce a crossbreed MO set, carrying features of the effective one-electron basis as well as those of LMOs. Revisiting the effective minimal AO construction scheme of Mayer,[1] we examine the union of occupied, atom-attributed LMOs obtained during the derivation. The outcome is a set of MOs, termed effective minimal Mayer-LMOs which are occupied as well as atom assigned. Dimension of the effective minimal Mayer-LMO set is in principle subject to user control. Following Mayer, it is fixed below as the dimension of the minimal AO basis set of the molecule. The union of atom localized LMOs receive some attention in Mayer's first paper[1] and appear without comment in the follow-up work.[2] His loss of interest might occurred due to the obvious linear dependence of the set. While

non-orthogonal character of localized orbitals has been often stressed to be beneficial,[36,48–54] linear dependence is mostly looked upon as an unwelcome feature, causing numerical burden. A handful of studies however report on deliberately exploiting linear dependence, in various scenarios.[55–61] In the present study, overcompleteness can be regarded as the price to pay for the dual nature of effective minimal Mayer-LMOs, i.e. occupied as well as atom assigned.

At the one-electron level, a linearly dependent set is easily obtained e.g. by projecting AOs into the occupied/virtual space. Application of the arising PS/QS-LMOs (with P/Q referring to the first-order particle/hole density matrix, and S standing for the AO overlap) mostly involve filtering to remove redundancy. We have recently pointed out the use of the unscreened, linearly dependent set of PS-LMOs in generating smoothly varying localized orbitals along reaction trajectories.[61] When propagating a system in time, multiple localization minima lead to unphysical jumps of traditional occupied orthogonal LMOs, while PS-related LMOs transition smoothly. At the same time, PS-LMOs are relatively expensive to deal with as their dimension is that of the AOs, in contrast with the occupied dimension of orthogonal LMOs. Moreover, PS-LMOs depend on the orientation of the molecule in the cartesian frame, when deduced from non s-type AOs.

Both unfavourable features of PS-LMOs are remedied by Mayer's effective minimal LMOs, as shown below. While they still form a redundant set, their number fixed at the minimal AO basis dimension of the molecule offers a good compromise when overcomplete LMOs are advantageous but numerical efficiency is also desired. Mayer's effective minimal LMOs can be viewed as a compactification of PS-LMOs on atomic or user-selected fragments, via a molecule-intrinsic criterion. The latter aspect ensures that Mayer's effective minimal LMOs are orientation independent.

Below Mayer's effective minimal LMOs are derived, extracted from a thirty-years-old paper.[2] Dealing with overcomplete orbital sets requires a generalization of the notion of bases, allowing for linear dependence. This foundation is provided by the mathematical theory of frames.[62–64] According to frame theory, redundant sets can be treated in many ways analogous to overlapping bases, using a generalized inverse of the overlap when necessary. We rely on the frame concept in the derivation and refer to our previous work on essentials of the underlying math.[61] The only background considered in some detail here is the overcomplete representation of the generalized eigenvalue problem, as this topic is missing from our previous report. Discussion of the formal features of effective minimal LMOs set is followed by illustrative numerical applications.

2 Theory

2.1 Mayer's effective LMOs

Suppose that spatial molecular orbitals (MO) φ_i arising as the self-consistent solution of the mean-field problem of a closed shell molecule are expanded on atomic orbitals (AO) χ_μ as

$$|\varphi_i\rangle = \sum_\mu |\chi_\mu\rangle C_{\mu i} \ , \tag{1}$$

allowing to express the projector of the occupied MO subspace as

$$\hat{P} = \sum_{\mu\nu} |\chi_\mu\rangle P_{\mu\nu} \langle\chi_\nu| \tag{2}$$

with elements of the AO based density matrix reading

$$P_{\mu\nu} = \sum_i^{\text{occ}} C_{\mu i} C^*_{\nu i} \ . \tag{3}$$

Mayer essentially reformulated[1] a Magnasco-Perico type localization[38,39] as a constrained extremum search. In this spirit, orbital ψ localized on a fragment (e.g. atom) A is obtained by taking the projector of AO-s assigned to fragment A

$$\hat{P}_A = \sum_{\mu\nu}^{\text{A}} |\chi_\mu\rangle \left(\mathbf{S}_A^{-1}\right)_{\mu\nu} \langle\chi_\nu| \tag{4}$$

and requiring that

$$M^A = \frac{\langle\psi|\hat{P}_A|\psi\rangle}{\langle\psi|\psi\rangle} \tag{5}$$

is stationary subject to the auxiliary condition that

$$\hat{P}|\psi\rangle = |\psi\rangle \ , \tag{6}$$

i.e. ψ should lie in the occupied subspace. In Eq.(4), $\mathbf{S}_A$ refers to the fragment block of the overlap matrix $\mathbf{S}$. For simplicity, atomic fragments are considered below. Extension to multiatomic fragments is straightforward.

Here, we rewrite Mayer's theory, in particular method "C" of Ref.,[2] leading to the equations for LMOs ψ. The latter are considered as linear combination of the AOs projected to

the occupied subspace, the so-called PS-LMOs

$$|\chi_\mu^{PS}\rangle = \hat{P}|\chi_\mu\rangle = \sum_\nu |\chi_\nu\rangle(\mathbf{PS})_{\nu\mu} \ . \tag{7}$$

Writing ψ in the form

$$|\psi\rangle = \sum_\mu |\chi_\mu^{PS}\rangle \, d_\mu \tag{8}$$

ensures that $\delta\psi$ arising by arbitrary variation of $\mathbf{d}$

$$|\delta\psi\rangle = \sum_\mu |\chi_\mu^{PS}\rangle \, \delta d_\mu \ , \tag{9}$$

conforms Eq.(6).

Variation of the localization functional Eq.(5) upon variation of ψ results

$$\langle\delta\psi|\hat{P}_A|\psi\rangle - M^A\langle\delta\psi|\psi\rangle + c.c. = 0 \ . \tag{10}$$

Substituting Eq.(9) in Eq.(10) and assuming real orbitals, one is led to

$$\sum_{\mu\nu} \delta d_\mu \, \langle\chi_\mu^{PS}|\hat{P}_A|\chi_\nu^{PS}\rangle \, d_\nu = \sum_{\mu\nu} \delta d_\mu \, \langle\chi_\mu^{PS}|\chi_\nu^{PS}\rangle \, d_\nu \ M^A \ . \tag{11}$$

Since δd_μ are arbitrary, they can be eliminated, arriving at

$$\sum_\nu \langle\chi_\mu^{PS}|\hat{P}_A|\chi_\nu^{PS}\rangle \, d_\nu = \sum_\nu \langle\chi_\mu^{PS}|\chi_\nu^{PS}\rangle \, d_\nu \ M^A \ , \quad \mu = 1, \ldots, n_b \tag{12}$$

which is the generalized eigenvalue equation of $\hat{P}_A$ on the overcomplete set of PS-LMOs. In Eq.(12) indices μ, ν extend for the dimension of the AO basis, n_b. Introducing $\mathbf{P}_A$ for the AO representation of $\hat{P}_A$, matrix notation of Eq.(12) reads

$$\mathbf{SPP}_A\mathbf{PS}\,\mathbf{d}_i = \mathbf{SPS}\,\mathbf{d}_i \ M_i^A \ . \tag{13}$$

Either of the standard techniques is applicable to solve the symmetric, generalized eigenvalue problem of Eq.(13), with invoking a generalized, e.g. Moore-Penrose inverse for the singular overlap matrix of PS-LMOs, **SPS**, when needed. Section 2.2 outlines an approach which is analogous to Löwdin's symmetric orthogonalization. One has to bear in mind however, that components of $\mathbf{d}_i$ are non-unique for a given ψ_i, as a consequence of the redundancy

of PS-LMOs. At the same time, localized functions ψ_i, generated by $\mathbf{d}_i$ are unique and orthogonal, based on Eqs.(8) and (13) .

While the dimension of the problem in Eq.(13) is that of the AO basis, the number of nonzero eigenvalues, M_i^A is limited by two factors. On one hand it can not exceed the number of occupied orbitals, n_{occ}, as this gives the dimension of the space spanned by PS-LMOs. On the other hand the rank of operator $\hat{P}_A$ in Eq.(4) also sets an upper limit, which is n_A, the number of AOs centered on atom A. The rank of matrix formed with elements $\langle \chi_\mu^{PS} | \hat{P}_A | \chi_\nu^{PS} \rangle$ is therefore $\min\{n_{occ}, n_A\}$, giving the maximal number of nonzero M_i^A as well. In practice, the number of M_i^A significantly different from zero is typically even smaller than $\min\{n_{occ}, n_A\}$. This is in line with Eq.(5), reflecting that eigenvalue M_i^A can be interpreted as a measure of atomic character of ψ_i, at the extent $\hat{P}_A$ can be attributed to atom A. Numerical experience indicates, that the dimension of the traditional minimal basis often manifests in the number of significant M_i^A on atom A.[2]

As a consequence of the idempotent character of $\hat{P}_A$ and Eq.(5), the value of M_i^A is subject to the constraints

$$0 \leq M_i^A \leq 1 \ . \tag{14}$$

To show Eq.(14), consider the orthogonal and complementary projector to $\hat{P}_A$, denoted by $\hat{P}_{A\perp}$. Introducing $M_{i\perp}^A = \langle \psi_i | \hat{P}_{A\perp} | \psi_i \rangle / \langle \psi_i | \psi_i \rangle$, the sum fulfills $M_i^A + M_{i\perp}^A = 1$ and both M_i^A and $M_{i\perp}^A$ are non-negative. Occupation number terminology for M_i^A is justified by Eq.(14). Numerical studies of Mayer[2] indicate that M_i^A values reflect the extent of atom localization: larger occupation numbers correspond to core and lone pair orbitals and smaller values occur for valence bonding orbitals.

Regarding Eq.(13) from the computational point of view, it appears a waste of effort to solve a problem of the dimension of the AO basis, n_b, while the number of nontrivial eigenvalues is only $\min\{n_{occ}, n_A\}$. As Mayer pointed out,[1] it is in fact possible to compactify Eq.(13) exploiting the restricted summation of the fragment projector, Eq.(4). The key observation is that atomic projector, Eq.(4) is effective only in the occupied subspace in Eq.(12). In other words, $\hat{P}_A$ can be substituted for $\hat{P}\hat{P}_A$ on the left hand side of Eq.(12). Evaluating the effect of $\hat{P}\hat{P}_A$ on an arbitrary PS-LMO as

$$\hat{P}\hat{P}_A | \chi_\mu^{PS} \rangle = \sum_\nu^A | \chi_\nu^{PS} \rangle \left(\mathbf{S}_A^{-1} \mathbf{SPS} \right)_{\nu\mu} \ , \tag{15}$$

a linear combination of PS-LMOs centered on atom A is obtained. This means that the PS-LMOs centered on atom A represent an invariant subspace of $\hat{P}\hat{P}_A$. Consequently, nonzero

eigenvalues of Eq.(12) arise from the expansion of ψ restricted as

$$|\psi\rangle = \sum_{\mu}^{A} |\chi_{\mu}^{PS}\rangle \, d_{\mu} \ . \tag{16}$$

Imposing the same restriction on the variation of ψ in Eq.(9) and repeating the derivation one arrives at

$$\sum_{\nu}^{A} \left(\mathbf{SPP}_A\mathbf{PS}\right)_{\mu\nu} \, d_{\nu}^{i} = \sum_{\nu}^{A} \left(\mathbf{SPS}\right)_{\mu\nu} \, d_{\nu}^{i} \, M_i^A \ , \quad \mu \in A \ , \tag{17}$$

where upper index i of d_{ν}^{i} refers to distinct solutions $i = 1, \ldots, n_A$. As Eq.(17) indicates, the original eigenvalue problem of Eq.(13) can be restricted to the indices attributed to atom A. This does not represent any approximation to Eq.(13), it merely eliminates zero eigenvalues M_i^A and the corresponding eigenvectors $\mathbf{d}^i$. Since the atomic block of **SPS** is singular in general, Eq.(17) is to be treated with invoking the Moore-Penrose inverse, c.f. Section 2.2.

In the numerical applications of Section 3, Eq.(17) is the working equation generating fragment localized, orthogonal LMO sets. It is worth mentioning, that the overcomplete representation behind Eq.(17) can be sidestepped by e.g. expanding ψ on AOs as described in Ref.[2] This way, however, the benefit of Eq.(17) as compared to Eq.(13) is lost, since eigenvector expansions become n_b long.

Observing that occupation number M_i^A is characteristic not only of LMO ψ_i but also of its atomic "truncation" $\hat{P}_A\psi_i$, Mayer proceeded in the direction of generating effective minimal AO sets, determined by significant M_i^A's on a given atom A.[1] He also examined various atomic localization functionals and found that they result in similar occupation numbers,[2] moreover it was demonstrated that effective minimal AOs, exhibiting spherical harmonics characteristics are attainable even if completely eliminating the concept of atomic basis as input.[65] Studies along the same line[3,5,6] as well as extensive use of the resulting techniques[17–20] underline the significance of Mayer's result.

We presently take a step back from AOs and wish to work with LMOs ψ_i emerging from Eq.(17). Our treatment of occupation numbers M_i^A is also slightly different from Mayer's. Admitting that often, but not always does the number of significant M_i^A reflect the minimal basis dimension on atom A, we consider occupation numbers M_i^A merely as ordering parameters and prescribe the number of LMOs based on atom type, instead of the value of M_i^A. The number of ψ_i picked for atom A is the minimal basis dimension of the given atom, denoted l_A. E.g. $l_A = 1$ for H and He, $l_A = 5$ for Li till Ne and $l_A = 9$ for Na till Ar.

Having solved Eq.(17) for each atom of the molecule, the union of atom localized and selected orbitals ψ_i represent an overcomplete set of dimension $L = \sum_A l_A$, where L is the standard minimal basis dimension of the molecule.

To arrive at Mayer's effective minimal LMOs, consider the L-dimensional overlap built with

$$\langle \psi_i | \psi_j \rangle = T_{ij} \ , \qquad i,j = 1, \ldots, L \tag{18}$$

Matrix $\mathbf{T}$ features unit blocks of l_A dimension along the diagonal. Interatomic blocks of $\mathbf{T}$ are generally nonzero and exhibit a spatial decay with increasing distance between the fragments, resulting in a sparse matrix. Moreover, $\mathbf{T}$ is singular, since the set $\{\psi_i\}_{i=1}^{L}$ spans the occupied space and $n_{occ} < L$ holds in general. Eliminating redundancy is discarded as it would sacrifice the localized character of ψ_i. We advocate instead to generate an L-dimensional, generalized Löwdin-set of ψ_i according to

$$|\phi_i\rangle = \sum_{j=1}^{L} |\psi_j\rangle \left(\overline{\mathbf{T}}^{1/2}\right)_{ji} \ , \qquad i = 1, \ldots, L \tag{19}$$

where $\overline{\mathbf{T}}$ is the Moore-Penrose pseudo-inverse[66] of $\mathbf{T}$.

These are orbitals ϕ_i, that we call Mayer's effective minimal LMOs, exhibiting the following features

1. ϕ_i are occupied LMOs, their local character stemming from the atomic localization functional Eq.(5);

2. ϕ_i can be assigned to atomic centers based on $\hat{P}_A$, but their character is not necessarily mono-atomic, as seen in Section 3;

3. $\{\phi_i\}_{i=1}^{L}$ is the closest frame to $\{\psi_i\}_{i=1}^{L}$ in difference squared norm sense;[67]

4. ϕ_i do not depend on the orientation of the molecule (at difference with e.g. PS-LMOs);

5. though $\{\phi_i\}_{i=1}^{L}$ is an overcomplete set, its dimension is only that of the the minimal basis, which is typically much smaller than the dimension of the AO basis;

6. $\{\phi_i\}_{i=1}^{L}$ represent a so-called Parseval-frame, facilitating orbital-based expressions analogous to orthonormal orbitals.[61] In general, ϕ_i are not normalized. Squared norms

$$\langle \phi_i | \phi_i \rangle = n_i \tag{20}$$

can be interpreted as fractional occupation number of effective minimal Mayer-LMOs. Projector of the occupied space in terms of ϕ_i reads

$$\hat{P} = \sum_{i=1}^{L} |\phi_i\rangle\langle\phi_i| \ , \tag{21}$$

its trace fulfilling[61]

$$\mathrm{Tr}(\hat{P}) = \sum_{i=1}^{L} n_i \ . \tag{22}$$

Relating item 1, it deserves mention that alternatives to the localization functional "C" of Ref.[2] can be formulated, leading to distinct, L-dimensional effective minimal LMO sets. With fixing the atomic localization functional and the values for l_A, the effective minimal LMOs are well-defined.

An attractive feature of Mayer's effective minimal LMOs from the computational point of view is that the steepest scaling steps of the procedure do not depend on the size of the AO basis. Atomic sets $\{\psi_i\}_{i=1}^{l_A}$ arise from eigenproblems determined by the AO basis on atom A, c.f. Eq.(17). The Löwdin-transformation of Eq.(19), necessary to generate the Parseval-frame of $\{\phi_i\}_{i=1}^{L}$ is only of dimension L, the minimal AO basis dimension of the molecule, not the actual AO basis.

Item 6 indicates that overcompleteness can be treated without inducing overcounting. Non-unit norm of Parseval-frame LMOs is an important aspect in this matter. The number of effective minimal LMOs assigned to atom A, l_A deserves a remark at this point. The value of l_A should be equal or larger than twice the number of electrons of atom A, to ensure that $\{\phi_i\}_{i=1}^{L}$ spans the occupied space, i.e. Eq.(21) holds. An l_A value larger than minimal may have interpretational motives, e.g. studying hipervalency. This affects individual LMOs and has a toll on efficiency by increasing L. At the same time, the occupied projector $\hat{P}$ and quantities derived from it are the very same with a larger LMO set as with the smaller. From the point of view of $\hat{P}$, the M_i^A value where the selection of ψ_i cuts in the spectrum of Eq.(17) is of no concern, as long as l_A is not subminimal.

Note, that the above considerations relied on a restricted type SCF solution of a closed-shell molecule. Extension to the unrestricted case is trivial by separate treatment of the occupied alpha and beta orbital subspaces. It also deserves mention, that the above derivations can be performed for the virtual subspace, mutatis mutandis.

2.2 Frame treatment of the generalized eigenvalue problem

The generalized eigenvalue problems of Eqs.(13) and (17) of Section 2.1, are of the common form

$$\mathbf{Hc} = \mathbf{Gc}E \ . \tag{23}$$

where $\mathbf{H}$ is a Hermitian matrix and $\mathbf{G}$ is a positive semidefinite metric. The non-invertible character of $\mathbf{G}$ is manifest in zero eigenvalues in its spectral decomposition

$$\mathbf{G} = \mathbf{U} \begin{pmatrix} \mathbf{g}^{\mathrm{d}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U}^{\dagger} \ . \tag{24}$$

Entries along the diagonal of $\mathbf{g}^{\mathrm{d}}$ are considered nonzero. Partition $\mathbf{U}$ as

$$\mathbf{U} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \end{pmatrix} \tag{25}$$

allowing to express the diagonal form of $\mathbf{G}$ as

$$\begin{pmatrix} \mathbf{g}^{\mathrm{d}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \mathbf{U}^{\dagger}\mathbf{G}\mathbf{U} = \begin{pmatrix} \mathbf{X}^{\dagger}\mathbf{G}\mathbf{X} & \mathbf{X}^{\dagger}\mathbf{G}\mathbf{Y} \\ \mathbf{Y}^{\dagger}\mathbf{G}\mathbf{X} & \mathbf{Y}^{\dagger}\mathbf{G}\mathbf{Y} \end{pmatrix} \ . \tag{26}$$

The bottom right block of the above equation means, that the transformation of the original, overcomplete vector set (generating matrix representations $\mathbf{H}$ and $\mathbf{G}$) by $\mathbf{Y}$ produces vectors of zero norm, with nonzero components. Accordingly, the transformation of $\mathbf{H}$ as

$$\mathbf{U}^{\dagger}\mathbf{H}\mathbf{U} = \begin{pmatrix} \mathbf{X}^{\dagger}\mathbf{H}\mathbf{X} & \mathbf{X}^{\dagger}\mathbf{H}\mathbf{Y} \\ \mathbf{Y}^{\dagger}\mathbf{H}\mathbf{X} & \mathbf{Y}^{\dagger}\mathbf{H}\mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^{\dagger}\mathbf{H}\mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{27}$$

also exhibits just one nonzero block.

Löwdin's canonical orthogonalization[68] implies omitting the vectors generated by $\mathbf{Y}$ (redundancy filtering) and solving the reduced dimensional eigenvalue problem of $\mathbf{X}^{\dagger}\mathbf{H}\mathbf{X}$ in metric $\mathbf{g}^{\mathrm{d}}$. An alternative approach is to keep to the original sized problem. Introduce for this end

$$\mathbf{V} = \mathbf{U} \begin{pmatrix} \left(\mathbf{g}^{\mathrm{d}}\right)^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{28}$$

and its Moore-Penrose inverse[66]

$$\overline{\mathbf{V}} = \begin{pmatrix} \left(\mathbf{g}^{\mathrm{d}}\right)^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U}^{\dagger} \tag{29}$$

and write the metric on the rhs as $\mathbf{G} = \overline{\mathbf{V}}^{\dagger}\overline{\mathbf{V}}$. Insert also the idempotent matrix $\mathbf{V}\overline{\mathbf{V}}$ in between $\mathbf{H}$ and $\mathbf{c}$ on the lhs of Eq.(23), leading to

$$\mathbf{H}\mathbf{V}\overline{\mathbf{V}}\mathbf{c} = \overline{\mathbf{V}}^{\dagger}\overline{\mathbf{V}}\mathbf{c}E \ . \tag{30}$$

Introducing $\mathbf{V}\overline{\mathbf{V}}$ on the lhs is allowed in Eq.(30) because the $\mathbf{Y}$ block of $\mathbf{U}$ implies a transformation to the nullspace of the original set. To check the validity of Eq.(30), multiply by $\mathbf{U}^{\dagger}$ from the left. This way the reduced dimensional eigenproblem is recovered in the top block while identical zeroes are produced in the bottom block on the two sides.

The above observation also allows a Löwdin's symmetrical orthogonalization[69] type of treatment of Eq.(23), via the square-root of the Moore-Penrose inverse of $\mathbf{G}$ as

$$\overline{\mathbf{G}}^{1/2} = \mathbf{U} \begin{pmatrix} \left(\mathbf{g}^{\mathrm{d}}\right)^{-1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U}^{\dagger} \ . \tag{31}$$

Note, that $\overline{\mathbf{G}}^{1/2}\mathbf{G}^{1/2}$ produces the same idempotent matrix as $\mathbf{V}\overline{\mathbf{V}}$. Rewrite Eq.(23) as

$$\mathbf{H}\overline{\mathbf{G}}^{1/2}\mathbf{G}^{1/2}\mathbf{c} = \mathbf{G}^{1/2}\mathbf{G}^{1/2}\mathbf{c}E \tag{32}$$

and multiply by $\overline{\mathbf{G}}^{1/2}$ from the left to get

$$\overline{\mathbf{G}}^{1/2}\mathbf{H}\overline{\mathbf{G}}^{1/2}\ \mathbf{G}^{1/2}\mathbf{c} = \overline{\mathbf{G}}^{1/2}\mathbf{G}^{1/2}\mathbf{G}^{1/2}\mathbf{c}E \ . \tag{33}$$

Realizing, that the effect of $\overline{\mathbf{G}}^{1/2}\mathbf{G}^{1/2}$ is immaterial on the rhs, Eq.(33) reads as the ordinary eigenvalue problem of matrix $\mathbf{H}_{\mathrm{L}} = \overline{\mathbf{G}}^{1/2}\mathbf{H}\overline{\mathbf{G}}^{1/2}$, with the eigenvector obtained as $\mathbf{b} = \mathbf{G}^{1/2}\mathbf{c}$, according to

$$\mathbf{H}_{\mathrm{L}}\,\mathbf{b} = \mathbf{b}\,E \ . \tag{34}$$

Eigenvector $\mathbf{c}$, corresponding to the original, untransformed vector set is recovered by $\mathbf{c} = \overline{\mathbf{G}}^{1/2}\mathbf{b}$.

Handling Eq.(23) by conserving the original dimension of the problem, as described above, implies a frame treatment[61,62] of the overcomplete vector set generating matrix representa-

tions **H** and **G**. A frame based biorthogonal approach, operating with the Moore-Penrose inverse of **G** can be also easily cast in analogy with the linearly independent case.

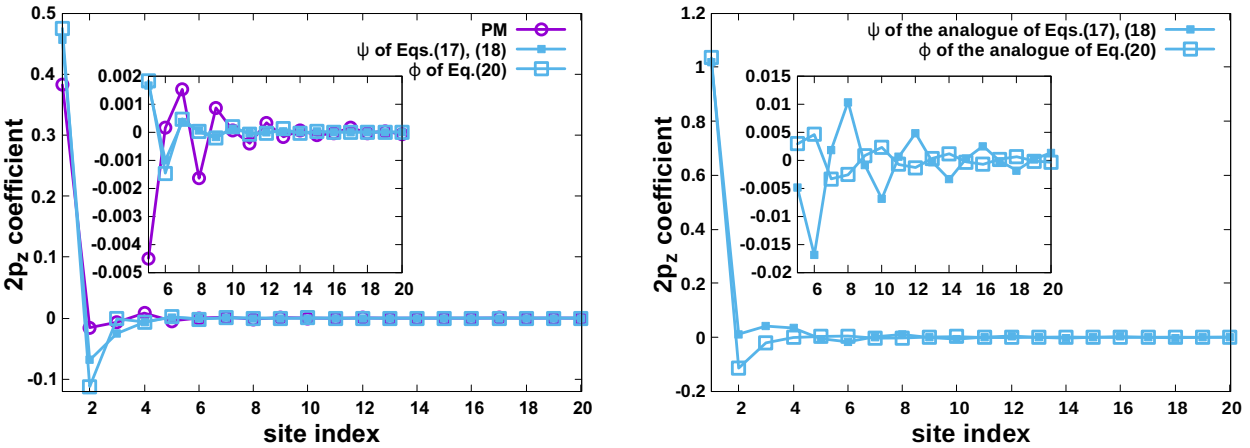
3 Numerical illustration

To generate Mayer's effective minimal LMOs, Eq.(17) was solved for each atom, applying the transformation of Eq.(34). Orbital subsets $\{\psi_i\}_{i=1}^{l_A}$ were then selected based on the largest M_i^A values and adjusting l_A by the element, i.e. five orbitals for carbon and one for hydrogen. Finally, the Löwdin-transformation of Eq.(19) was applied to the union of the atom-selected subsets.

Effective minimal Mayer-LMOs, produced this way are visualized either by contour plots, or by showing their AO expansion coefficients. In comparison, Boys[32] or Pipek-Mezey[33] (PM) orbitals are displayed with the same technique.

Our test cases serve to illustrate the localized character and orientational invariance of Mayer's effective minimal LMOs. We also demonstrate the smooth transition of ϕ_i for two processes where Boys and PM orbitals undergo a non-continuous switch along the reaction coordinate.

3.1 Localized character



(a) (b)

Figure 1: LMO expansion coefficients of the tightest $2p_z$ AOs of carbon atoms a $C_{20}H_2$ linear chain of metallic-type, equidistant geometry, computed in Dunning's DZ basis. The carbon–carbon distances are uniformly 1.3 Å , while the terminal C–H bonds are 1.1 Å long. Numbering of carbon atoms is indicated as site index. Panel (a) occupied LMOs; panel (b) virtual LMOs.

A $C_{20}H_2$ linear chain serves an example to compare localized nature of LMOs. For illustration π-LMOs centered on one of the penultimate atoms of the chain (site 1 in Fig. 1) are picked. Panel (a) of Fig. 1 depicts occupied orbitals, juxtaposing the PM-LMO with Mayer's effective minimal LMOs, the latter labeled ϕ. The outcome of the extremum problem of Eq.(17) is also displayed, labeled ψ. This latter can be considered the mid-step of generating the effective minimal Mayer-LMOs, and allows to grab the effect of the Löwdin-transformation of Eq.(19). Note, that ϕ are not normalized in general, at difference with ψ. For the sake of comparison, coefficients of the normalized frame orbital ϕ are plotted in Fig. 1.

Panel (b) of Fig. 1 displays the virtual analogues of ϕ and ψ obtained by substituting the occupied projector for the virtual in Eq.(2) and subsequently.

Localization on the terminal carbon atom is apparent for all LMOs shown in Fig. 1. The PM orbital in panel (a) is somewhat more spread out, than Mayer's effective minimal LMO, based on the inset. Coefficients of orbitals ϕ and ψ are closer in Fig. 1(a) and deviate somewhat more in Fig. 1(b). Both panels (a) and (b) of Fig. 1 indicate, that the Löwdin-transformation does not act against locality of the orbitals stemming from the extremum principle. This is in line with the smallest distortion theorem mentioned in item 3 of Section 2.1.

3.2 Orientation invariance

Rotation of a rigid H_2O molecule in the Cartesian frame serves to showcase orientation dependence of PS-LMOs at difference with Mayer's effective minimal LMOs. Instead of contour plots, Fig. 2 displays occupation numbers obtained by Eq.(20) and multiplied by two for spin summation. Occupation number of PS-LMOs is obtained by projecting all AOs to the occupied space and treating the full set by Eq.(19). The rotation angle being parallel with the y-axis of the Cartesian system, the $2p_y$-related PS-LMO is insensitive to the transformation. The $2p_x$ and $2p_z$ AOs are however transforming and so do the LMOs projected from them, as illustrated by the change in their occupation number in Fig. 2. Effective minimal Mayer LMOs, exhibiting distinct M_i^A values are invariant, as indicated by the straight horizontal lines in Fig. 2. Occupation numbers of effective minimal Mayer LMOs are larger than those of PS-LMOs since the former is a smaller set than the latter and both fulfill Eq.(22).

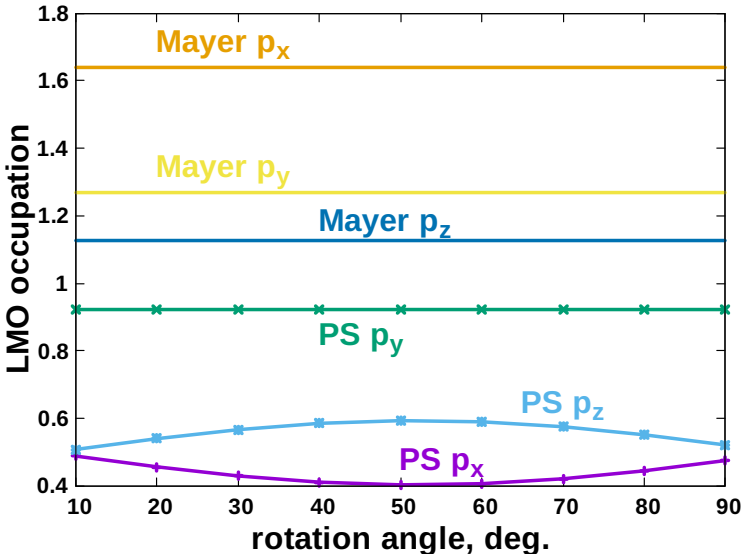


Figure 2: Occupation number of LMOs originating from the $2p$ AOs of atom O in the H_2O molecule, computed in Huzinaga's MIDI basis. The molecule lies in the (xz) plane, OH bond distances are 1.09 Å , HOH bond angle is 104 ^{o}. Rotation performed around an axis containing atom O and perpendicular to the molecular plane is measured on the horizontal axis of the plot. Mayer's effective minimal LMOs are compared to PS-LMOs.

3.3 Bond rearrangement

Rearrangement of double bonds in unsaturated molecules induced by geometrical distortion is a typical example where multiple minima can occur with traditional localization procedures at high-symmetry points of the reaction trajectory. Benzene serves as a simple example of this behavior, semibullvalene offers a more complex case.

3.3.1 Benzene

Kekulé-structures of benzene belong to the D_{3h} point group. We examine the gradual transition between the two Kekulé arrangements, through a higher-symmetry, D_{6h} geometry point. Carbon atoms are placed on a circle of 1.4096 Å radius. Central angle α of one CC bond serves as reaction coordinate. CH bonds point radially out, their length fixed at 1.0968 Å .

Since two completely equivalent sets of Boys LMOs are present at the sixfold symmetric arrangement, it is by chance, that Boys LMOs at $\alpha = 60^o$ are in correspondence with those at $\alpha = 54^o$ in Fig. 3. They could equally correspond rather to the LMOs at $\alpha = 66^o$. Either way, there occurs a switch in the LMOs along the reaction coordinate, i.e. progressing in the rows of Fig. 3. As discussed previously,[61] non-smooth orbital transition in this process is a feature of orthogonal, localized LMOs.

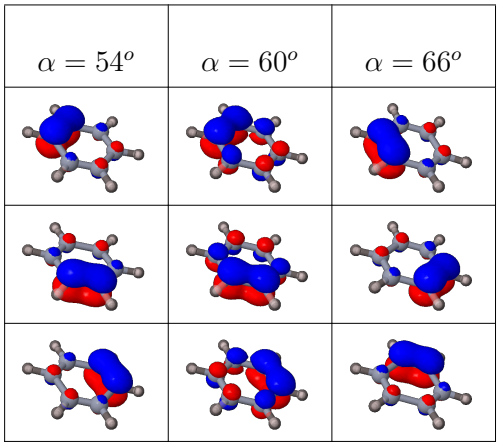


Figure 3: Boys-localized, occupied π-LMOs of the benzene molecule in cc-pVDZ basis. The central angle of one CC bond as seen from the center, parameter α is indicated in column header. Point group is D_{6h} at $\alpha = 60^o$. Kekulé-structures of D_{3h} symmetry correspond to $\alpha \neq 60^o$.

The overcomplete set of effective minimal Mayer-LMOs show a different picture in Fig. 4. The dimension of the set is six, as each carbon atom contributes one occupied π-LMO. When α is significantly far from 60^o, Mayer's effective minimal π-LMOs reflect the Kekulé-structure, with LMOs centered on the shorter CC bonds. Essentially the same LMO is obtained twice, by the contribution of the bonding carbon atoms. Getting closer to $\alpha = 60^o$, the pairwise roughly equivalent, Kekulé-type π-LMOs deviate more and more, developing at $\alpha = 60^o$ into six LMOs related by a single or multiple C_6 rotations and extending over three neighbouring carbon atoms. It is important to realize that the six effective minimal Mayer-LMOs represent the D_{6h} point group at $\alpha = 60^o$, while this is not true for the three Boys π-LMOs, either at the localization minima shown in Fig. 4 or the other one. This is the main reason behind Mayer's effective minimal LMOs transitioning smoothly in the rows of Fig. 4, gradually transforming from a CC π-bond of one Kekulé-structure to that of the other.

Mayer's effective minimal LMO set is not the only one varying smoothly with the reaction coordinate. The previously analyzed, PS-related LMOs also show this behavior. In the approach of Ref.,[61] the tightest $2p_z$ AOs would be projected in the occupied space and the resulting, six-dimensional Parseval frame shows essentially the same picture as in Fig. 4. While the outcome is similar on this example, an important difference in the procedure leading to effective minimal Mayer's and PS-related LMOs is the origin of dimension reduction form the AO basis to six. With PS-LMOs this is achieved by a user selection of AOs to be projected. With effective minimal Mayer-LMOs, Eq.(17) determines an optimal combination

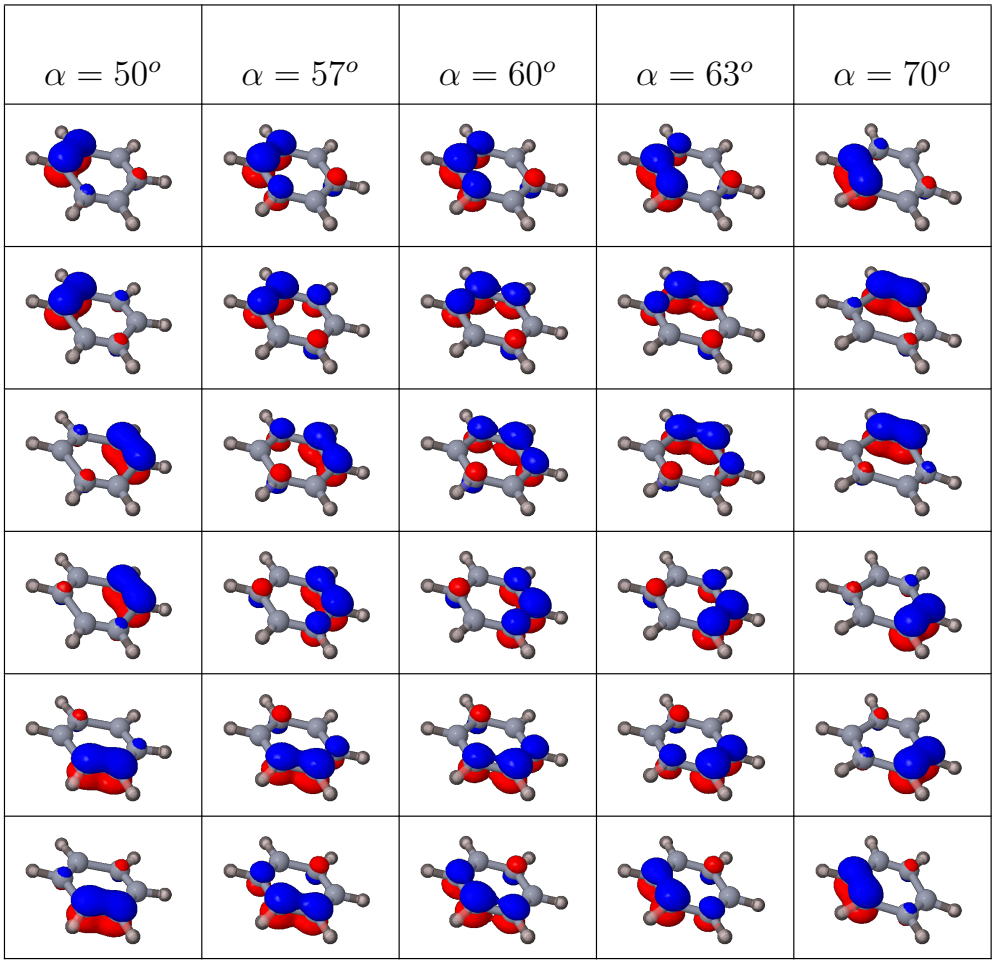


Figure 4: Mayer's effective minimal π-LMOs of the benzene molecule in cc-pVDZ basis. Parameter α is the central angle of one CC bond as seen from the center. Point group is D_{6h} at $\alpha = 60^o$. Kekulé-structures of D_{3h} symmetry correspond to $\alpha \neq 60^o$.

of the p_z carbon AOs, describing the occupied π-contribution of the atom in question.

In comparison with Boys-LMOs, two further comments are due. Separation of σ and π MOs is automatically ensured by Eq.(17) while rotations were restricted in the π-space to generate the orbitals shown in Fig. 3. Figs. 3 and 4 also reflect that local character of effective minimal Mayer and Boys π-LMOs is similar when bond length alternation is significant, i.e. α far from 60^o.

3.3.2 Semibullvalene

In the case of semibullvalene, σ-π separation does not apply to the MOs. For this reason, a minimal AO set for PS-treatment is less plausible to select, making semibullvalene tautomerization a more complicated case than benzene.

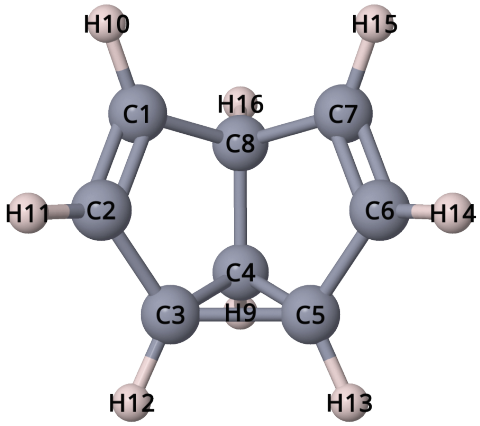


Figure 5: Molecular structure of semibullvalene at the initial point of the bond rearrangement process. Double bonds C1=C2, C6=C7 and single bond C3-C5 are cleaved during the reaction. Simultaneously, double bonds C2=C3, C5=C6 and single bond C1-C7 are formed.

We examine bond reorganization affecting the C1-C2 and C6-C7 double bonds and the C3-C5 single bond of the initial structure, shown in Fig. 5. Product of the reaction is a tautomer pair, with double bonds between C2-C3 and C5-C6 and a σ-bond between atoms C1-C7. Hartree-Fock computations at the discretized points of the reaction trajectory, taken from the Supplementary Information of Ref.[61] were performed in cc-pVDZ basis.

A common feature with the previous, benzene study case is the equivalence of two localization minima corresponding to the two Kekulé-structures, at the mid-point of the reaction trajectory. Variation of Boys or PM LMOs is accordingly characterized by a switch at the middle, C_{2v} symmetry point of the reaction path.

We have previously demonstrated that the overcomplete set of PS-related LMOs change smoothly in this process,[61] it was however difficult to match the orbital contour plots with the chemical picture of the breaking and forming σ and π bonds. When relying on PS-LMOs, the bond rearrangement effect gets smeared over all p-AO related LMOs of the carbon atoms. In parallel with orientation dependence getting rectified by Mayer's effective minimal LMOs, their character also becomes better agreeing with chemical notion. Effective minimal Mayer-LMOs are still atom attributable, contour plots are however rather di- or triatomic in Fig. 6. It is also relatively easy to make out simple combinations of diatomic bonds when the contour plots are rather tri- or multiatomic in Fig. 6. This is undoubtedly an advantage over PS-related LMOs examined in Ref.[61] The effect can be attributed to solving the extremum problem of Eq.(17). This step not only compactifies PS-LMOs, it also makes the bonds with contributions from atom A more expressed.

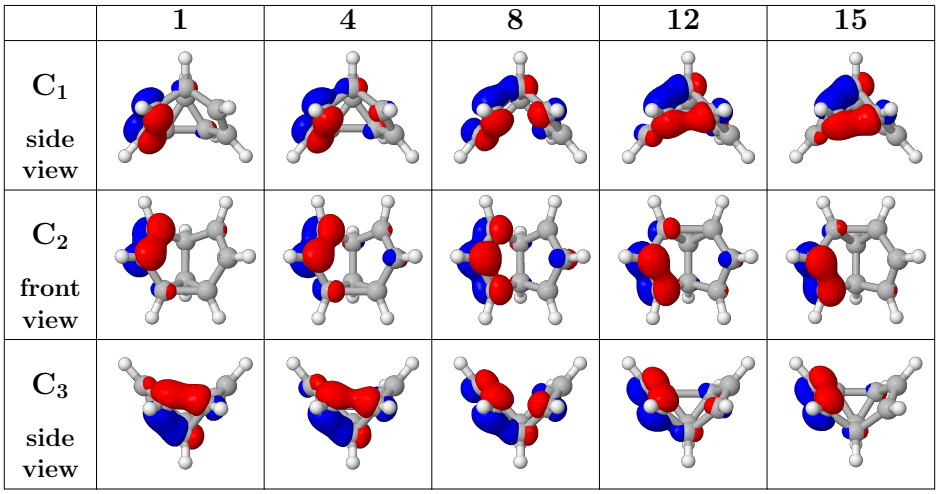


Figure 6: Contours of Mayer's effective minimal LMOs at selected steps of the tautomerization reaction of semibullvalene. Table header indicates step No. of the trajectory discretized for 15 points, taken from Ref.[61] Among the five LMOs originating from a carbon atom, the one mostly affected in the process is displayed. View of the molecule is adjusted to best show the LMO shape. See atom numbering in Fig. 5. Orbital plots of LMOs originating from C_5, C_6 and C_7 can be generated by C_s reflection of those shown.

Evolution of the C_2-originating effective minimal Mayer-LMO in Fig. 6 is reminiscent of the case of benzene. The C_1-related LMO in Fig. 6 clearly shows a π-bond transforming into a combination of σ-bonds while the C_3-related LMO undergoes the reverse process.

There are three effective minimal Mayer-LMOs, originating from C_5, C_6 and C_7, related by C_s reflection to those in Fig. 6. Altogether, these six LMOs represent the C_{2v} point group of the system at the mid-point of the trajectory, i.e. at column 8 of Fig. 6. Similar to the case of benzene, this symmetry argument is the main origin of overcomplete LMOs affected by the bond rearrangement being free from any discontinuity along the reaction coordinate, i.e. in the rows of the table.

4 Conclusion

Summarizing, Mayer's effective minimal LMOs are similar to projected AOs in providing disjoint (though non-orthogonal) subsets of occupied orbitals assigned to atoms or multiatomic fragments. While effective minimal Mayer-LMOs are atom related, they lie in the *molecular* occupied subspace. They therefore facilitate the distribution of occupied orbitals for atomic or larger fragments.

Superior feature of effective minimal Mayer-LMOs compared to projected AOs is on one

hand their reduced dimension. The number of Mayer's effective minimal LMOs is subject to user control. When dealing with the occupied subspace, the minimal AO basis set dimension appears a practical choice. Mayer's effective minimal LMOs also improve on the behavior of projected AOs with rotating the cartesian frame. Effective minimal Mayer-LMOs are invariant to changing the orientation of the coordinate system, while the same statement holds for the subspace of the full $p, d, \ldots$ shell derived PS-LMOs.

Mayer's effective minimal LMOs arise by a compactification of projected AOs, achieved by solving Eq.(17) and selecting the most important MOs based on the occupation numbers, M_i^A. While native fragment orbitals ψ, stemming from Eqs.(16) and (17) are orthogonal, their Löwdin-transformed counterparts, ϕ generated by Eq.(19), correspond to the entire molecule, and they are not only overlapping but also linearly dependent. Overcompleteness originates from essentially the same bonding LMOs or LMO subsets being generated by the participating atoms. Overcompleteness is a key element facilitating fragmentation of the occupied LMOs for atomic subspaces. Redundant feature of effective minimal Mayer-LMOs is also essential for providing a unique set when traditional localization methods produce multiple, equivalent sets of localized orbitals. It is important to stress that overcompleteness of Mayer's effective minimal LMOs does not generate overcounting. E.g., correct trace of the occupied subspace projector is obtained as a sum of fractional occupation numbers, n_i of Eq.(20), when expressed with Mayer's effective minimal LMOs.

Overcompleteness applies not only to effective minimal Mayer-LMOs but also to PS-LMOs used for representing the eigenvalue problem, resulting from the Magnasco-Perico type extremum condition. It is for this latter reason that the frame concept appears when solving Eq.(17).

Numerical illustrations indicate that valence effective minimal Mayer-LMOs, localized on multiatomic fragments, exhibit the participation of the selected atom in bonds with the neighbouring atoms. This result is a consequence of the extremum condition, yielding native fragment MOs, ψ. The important point here is that the Löwdin-transformation of Eq.(19) does not deteriorate the localized character of the MOs.

To our knowledge, effective minimal Mayer-LMOs are unique in showing the characteristics of atom/fragment attributed occupied orbitals, which are orientation invariant, of relatively low dimension and do not rely on external data, e.g. free-atom quantities or a predefined minimal basis as input. In addition, they are computationally economical as the most time consuming steps are performed in considerably smaller dimension than the molecular AO basis. These are potentially beneficial features for fragmentation and embedding methods. The present study focused on occupied orbitals but the analogous procedure can be formulated in the virtual space by trivial replacement of the occupied projector for the vir-

tual. Effective minimal virtual Mayer-LMOs may be useful for localization enhanced linear scaling techniques, as a compactified QS-LMO set. Machine learning based approaches that require rotationally invariant atomic descriptors[70,71] may also benefit from effective minimal Mayer-LMOs. Extension to correlated wavefunctions is also possible via partitioning the correlated one-particle density for atoms or fragments, based on e.g. scheme A of Ref.[2] or applying an analysis in the 3-dimensional space, c.f. schemes B and E of the same paper.

Acknowledgments

Computations have been performed with the Budapest version[72] of the MUNGAUSS program system.[73] Orbital contour plots were drawn with the cutoff value 0.05 in Jmol.[74] The work was supported by the National Research, Development and Innovation Office (NKFIH) under the Hungarian National Excellence Programme (NKKP) ADVANCED grant scheme, No. 150283.

Discussions with J. Daru (ELTE, Eötvös Loránd University, Budapest) are gratefully acknowledged.

References

[1] MAYER, I. Non-orthogonal localized orbitals and orthogonal atomic hybrids derived from Mulliken's population analysis. *Chemical Physics Letters* **1995**, *242*, 499–506.

[2] Mayer, I. Atomic orbitals from molecular wave functions: The effective minimal basis. *Journal of Physical Chemistry* **1996**, *100*, 6249–6257.

[3] Lu, W.; Wang, C.; Schmidt, M.; Bytauas, L.; Ho, K.; Ruedenberg, K. Molecule intrinsic minimal basis sets I. Exact resolution of "*ab initio*" optimized molecular orbitals in terms of deformed atomic minimal-basis orbitals. *The Journal of Chemical Physics* **2004**, *120*, 2629.

[4] West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. A comprehensive analysis of molecule-intrinsic quasi-atomic, bonding, and correlating orbitals. I. Hartree-Fock wave functions. *The Journal of Chemical Physics* **2013**, *139*, 234107.

[5] Szczepanik, D.; Mrozek, J. Minimal set of molecule-adapted atomic orbitals from maximum overlap criterion. *Journal of Mathematical Chemistry* **2013**, *51*, 2687–2698.

[6] Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *Journal of Chemical Theory and Computation* **2013**, *9*, 4834–4843.

[7] Janowski, T. Near Equivalence of Intrinsic Atomic Orbitals and Quasiatomic Orbitals. *Journal of Chemical Theory and Computation* **2014**, *10*, 3085–3091.

[8] Adams, W. H. Orbital Theories of Electronic Structure. *The Journal of Chemical Physics* **1962**, *37*, 2009–2018.

[9] J.P.Foster; F.Weinhold Natural hybrid orbitals. *Journal of the American Chemical Society* **1980**, *102*, 7211.

[10] A.E.Reed; L.A.Curtiss; F.Weinhold Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chemical Reviews* **1988**, *88*, 899.

[11] Heinzmann, R.; Ahlrichs, R. Population analysis based on occupation numbers of modified atomic orbitals (MAOs). *Theoretica chimica acta* **1976**, *42*, 33–45.

[12] Ehrhardt, C.; Ahlrichs, R. Population analysis based on occupation numbers II. Relationship between shared electron numbers and bond energies and characterization of hypervalent contributions. *Theoretica chimica acta* **1985**, *68*, 231–245.

[13] Cioslowski, J.; Liashenko, A. Atomic orbitals in molecules. *The Journal of Chemical Physics* **1998**, *108*, 4405–4412.

[14] Lee, M. S.; Head-Gordon, M. Polarized atomic orbitals for self-consistent field electronic structure calculations. *The Journal of Chemical Physics* **1997**, *107*, 9085–9095.

[15] Lee, M. S.; Head-Gordon, M. Extracting polarized atomic orbitals from molecular orbital calculations. *International Journal of Quantum Chemistry* **2000**, *76*, 169–184.

[16] Reed, A. E.; Weinhold, F. Natural localized molecular orbitals. *The Journal of Chemical Physics* **1985**, *83*, 1736–1740.

[17] Knizia, G.; Klein, J. E. M. N. Electron Flow in Reaction Mechanisms—Revealed from First Principles. *Angewandte Chemie International Edition* **2015**, *54*, 5518–5522.

[18] Ramos-Cordoba, E.; Postils, V.; Salvador, P. Oxidation States from Wave Function Analysis. *Journal of Chemical Theory and Computation* **2015**, *11*, 1501–1508.

[19] Dhar, D.; Yee, G. M.; Spaeth, A. D.; Boyce, D. W.; Zhang, H.; Tolman, W. B. Mechanisms for Hydrogen-Atom Abstraction by Mononuclear Copper–Oxygen Complexes. *Journal of the American Chemical Society* **2019**, *141*, 17236–17245.

[20] Zhang, P.; Zhao, Y.; Liu, F.; Houk, K. N. Involving Carbene or Not? Mechanism of Corey–Winter Reaction. *Journal of the American Chemical Society* **2025**, *147*, 13915–13922.

[21] Rubio, J.; Povill, A.; Malrieu, J.-P.; Reinhardt, P. Direct determination of localized Hartree–Fock orbitals as a step toward N scaling procedures. *The Journal of Chemical Physics* **1997**, *107*, 10044–10050.

[22] Maslen, P. E.; Ochsenfeld, C.; White, C. A.; Lee, M. S.; Head-Gordon, M. Locality and Sparsity of Ab Initio One-Particle Density Matrices and Localized Orbitals. *The Journal of Physical Chemistry A* **1998**, *102*, 2215–2222.

[23] Schütz, M.; Hetzer, G.; Werner, H.-J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *The Journal of Chemical Physics* **1999**, *111*, 5691–5705.

[24] Li, S.; Shen, J.; Li, W.; Jiang, Y. The cluster-in-molecule method for electron correlation of large systems. *The Journal of Chemical Physics* **2006**, *125*, 074109.

[25] Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *The Journal of Chemical Physics* **2013**, *138*, 034106.

[26] Nagy, P. R.; Samu, G.; Kállay, M. An integral-direct linear-scaling second-order Møller–Plesset approach. *Journal of Chemical Theory and Computation* **2016**, *12*, 4897.

[27] Sun, Q.; Chan, G. K.-L. Quantum Embedding Theories. *Accounts of Chemical Research* **2016**, *49*, 2705–2712.

[28] Szirmai, A. B.; Hégely, B.; Tajti, A.; Kállay, M.; Szalay, P. Projected Atomic Orbitals As Optimal Virtual Space for Excited State Projection-Based Embedding Calculations. *Journal of Chemical Theory and Computation* **2024**, *20*, 3420–3425.

[29] Richard, R. M.; Lao, K. U.; Herbert, J. M. Aiming for Benchmark Accuracy with the Many-Body Expansion. *Accounts of Chemical Research* **2014**, *47*, 2828–2836.

[30] Peng, L.; Peng, D.; Gu, F. L.; Yang, W. Regularized Localized Molecular Orbitals in a Divide-and-Conquer Approach for Linear Scaling Calculations. *Journal of Chemical Theory and Computation* **2022**, *18*, 2975–2982.

[31] Nakai, H.; Kobayashi, M.; Yoshikawa, T.; Seino, J.; Ikabata, Y.; Nishimura, Y. Divide-and-Conquer Linear-Scaling Quantum Chemical Computations. *The Journal of Physical Chemistry A* **2023**, *127*, 589–618.

[32] Foster, J.; Boys, S. The Quantum Theory of Molecular Orbitals. *Reviews of Modern Physics* **1960**, *32*, 300.

[33] J. Pipek and P. G. Mezey A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *The Journal of Chemical Physics* **1989**, *90*, 4916.

[34] Edmiston, C.; Ruedenberg, K. Localized Atomic and Molecular Orbitals. *Reviews of Modern Physics* **1963**, *35*, 457.

[35] Edmiston, C.; Ruedenberg, K. Localized Atomic and Molecular Orbitals. II. *The Journal of Chemical Physics* **1965**, *43*, S97.

[36] Famulari, A.; Sironi, M.; Raimondi, M.; Gianinetti, E. An orthogonal approach to determine extremely localised molecular orbitals. *Theoretical Chemistry Accounts* **2000**, *103*, 417–422.

[37] Ferenczy, G. G. In *Hungarian Quantum Chemistry: Part B - Contemporary Research*; Surján, P. R., Szabados, Á., Eds.; Advances in Quantum Chemistry; Academic Press: Amsterdam, 2026; Vol. 94; Chapter 3.

[38] Magnasco, V.; Perico, A. Uniform Localization of Atomic and Molecular Orbitals. I. *The Journal of Chemical Physics* **1967**, *47*, 971.

[39] Magnasco, V.; Perico, A. Uniform Localization of Atomic and Molecular Orbitals. II. *The Journal of Chemical Physics* **1968**, *48*, 800.

[40] Wu, F.; Liu, W.; Zhang, Y.; Li, Z. Linear-Scaling Time-Dependent Density Functional Theory Based on the Idea of "From Fragments to Molecule". *Journal of Chemical Theory and Computation* **2011**, *7*, 3643–3660.

[41] Vanfleteren, D.; Van Neck, D.; Bultinck, P.; Ayers, P. W.; Waroquier, M. Stockholder projector analysis: A Hilbert-space partitioning of the molecular one-electron density matrix with orthogonal projectors. *The Journal of Chemical Physics* **2012**, *136*, 014107.

[42] Náray-Szabó, G.; Mezey, P. G. A wavefunction model to chemical bonding. *International Journal of Quantum Chemistry* **2022**, *122*, e26686.

[43] Loos, P.-F.; Assfeld, X. Self-Consistent Strictly Localized Orbitals. *Journal of Chemical Theory and Computation* **2007**, *3*, 1047–1053.

[44] de Silva, P.; Giebułtowski, M.; Korchowiec, J. Fast orbital localization scheme in molecular fragments resolution. *Physical Chemistry Chemical Physics* **2012**, *14*, 546–552.

[45] Claudino, D.; Mayhall, N. J. Automatic Partition of Orbital Spaces Based on Singular Value Decomposition in the Context of Embedding Theories. *Journal of Chemical Theory and Computation* **2019**, *15*, 1053–1064.

[46] Giovannini, T.; Koch, H. Fragment Localized Molecular Orbitals. *Journal of Chemical Theory and Computation* **2022**, *18*, 4806–4813.

[47] Gimferrer, M.; Aldossary, A.; Salvador, P.; Head-Gordon, M. Oxidation State Localized Orbitals: A Method for Assignin Oxidation States Using Optimally Fragment-Localized Orbitals and Fragment Orbital Localization Index. *Journal of Chemical Theory and Computation* **2021**, *18*, 309–322.

[48] Meyer, W. PNO–CI Studies of electron correlation effects. I. Configuration expansion by means of nonorthogonal orbitals, and application to the ground state and ionized states of methane. *Journal of Chemical Physics* **1973**, *58*, 1017–1035.

[49] Mayer, I. Non-orthogonal localized orbitals to study delocalization effects. *Chemical Physics Letters* **1982**, *89*, 390–394.

[50] Feng, H.; Bian, J.; Li, L.; Yang, W. An efficient method for constructing nonorthogonal localized molecular orbitals. *The Journal of Chemical Physics* **2004**, *120*, 9458–9466.

[51] Usvyat, D.; Schütz, M. Nonorthogonal ultralocalized functions and fitted Wannier functions for local electron correlation methods for solids. *Theoretical Chemistry Accounts* **2005**, *114*, 276–282.

[52] Zoboki, T.; Mayer, I. Extremely localized nonorthogonal orbitals by the pairing theorem. *Journal of Computational Chemistry* **2011**, *32*, 689–695.

[53] Høyvik, I.-M.; Olsen, J.; Jørgensen, P. Generalising localisation schemes of orthogonal orbitals to the localisation of non-orthogonal orbitals. *Molecular Physics* **2017**, *115*, 16 – 25.

[54] Luo, Z.; Khaliullin, R. Z. Variable-Metric Localization of Occupied and Virtual Orbitals. *Journal of Chemical Theory and Computation* **2021**, *17*, 5568–5581.

[55] Christiansen, O.; Manninen, P.; Jørgensen, P.; Olsen, J. Coupled-cluster theory in a projected atomic orbital basis. *The Journal of Chemical Physics* **2006**, *124*, 084103.

[56] Weijo, V.; Manninen, P.; Jørgensen, P.; Christiansen, O.; Olsen, J. General biorthogonal projected bases as applied to second-order Møller-Plesset perturbation theory. *The Journal of Chemical Physics* **2007**, *127*, 074106.

[57] Margócsy, Á.; Szabados, Á. Ring coupled cluster doubles at the multireference level. *The Journal of Chemical Physics* **2020**, *152*, 204114.

[58] Szabados, Á.; Margócsy, Á.; Surján, P. R. Pivot invariance of multiconfiguration perturbation theory via frame vectors. *The Journal of Chemical Physics* **2022**, *157*, 174118.

[59] Sizov, G. N.; Staroverov, V. N. Properties and applications of basis sets with linearly dependent products. *Theoretical Chemistry Accounts* **2025**, *144*, 24.

[60] Sizov, G. N.; Staroverov, V. N. Reconstruction of Local Potentials from Matrix Representation in Arbitrary Basis Sets. *Journal of Chemical Theory and Computation* **2026**, *22*, 6525–6532.

[61] Szabados, Á.; Daru, J.; Gombás, A.; Surján, P. R. Is a Bit over Okay? – Overcomplete Sets of Localized Molecular Orbitals. *Journal of Chemical Theory and Computation* **2026**, *22*, 6622–6634.

[62] Christensen, O. *An Introduction to Frames and Riesz Bases*; Applied and Numerical Harmonic Analysis; Birkhäuser, 2016; Vol. 27.

[63] Casazza, P. G.; Kutyniok, G.; Philipp, F. In *Finite Frames: Theory and Applications*; Casazza, P. G., Kutyniok, G., Eds.; Birkhäuser Basel, 2012; Chapter 1, pp 1–54.

[64] Morgenshtern, V. I.; Bölcskei, H. In *Mathematical Foundations for Signal Processing, Communications, and Networking*; E. Serpedin, T. C., Rajan, D., Eds.; CRC Press, 2012; Chapter 20, pp 737–789.

[65] Mayer, I.; Bakó, I.; Stirling, A. Are There Atomic Orbitals in a Molecule? *The Journal of Physical Chemistry A* **2011**, *115*, 12733–12737.

[66] Israel, A. B.; Greville, T. *Generalized Inverses: Theory and Applications, (2nd ed.)*; Springer,: New York, 2003.

[67] Eldar, Y. C.; Forney, G. D. Optimal tight frames and quantum measurement. *IEEE Transactions on Information Theory* **2001**, *48*, 599–610.

[68] Löwdin, P.-O. Quantum Theory of Cohesive Properties of Solids. *Advances in Physics* **1966**, *5*, 1.

[69] Löwdin, P.-O. On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. *The Journal of Chemical Physics* **1950**, *18*, 365.

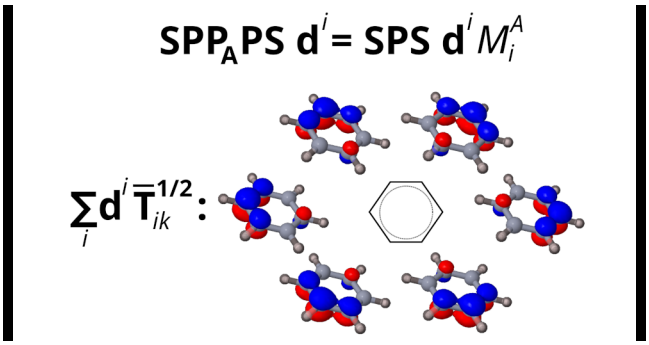
[70] Behler, J. Perspective: Machine learning potentials for atomistic simulations. *The Journal of Chemical Physics* **2016**, *145*, 170901.

[71] Stolte, N.; Daru, J.; Forbert, H.; Marx, D.; Behler, J. Random Sampling Versus Active Learning Algorithm for Machine Learning Potentials of Quantum Liquid Water. *Journal of Chemical Theory and Computation* **2025**, *21*, 886–899.

[72] Surján, P. R. *Program BP-MUNGAUSS*; 2026.

[73] Poirier, R.; Peterson, M. *Program MUNGAUSS*; 1989.

[74] Jmol Development Team Jmol: An open-source Java viewer for chemical structures in 3D. `http://www.jmol.org/` Date of access: 2026. 10. 11. .


$\mathbf{SPP_APS\ d}^i = \mathbf{SPS\ d}^i M_i^A$
$\sum_i \mathbf{d}^i \bar{\mathbf{T}}_{ik}^{1/2}$:

TOC Graphic