ACCOUNTS of chemical research

Article

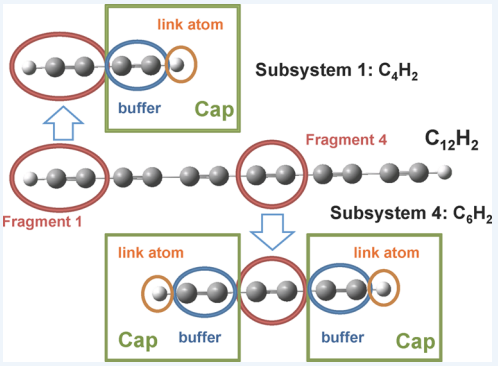
pubs.acs.org/accounts

Localization of Molecular Orbitals: From Fragments to Molecule

Zhendong Li, Hongyang Li, Bingbing Suo, and Wenjian Liu*

Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, State Key Laboratory of Rare Earth Materials Chemistry and Applications, College of Chemistry and Molecular Engineering, and Center for Computational Science and Engineering, Peking University, Beijing 100871, People's Republic of China

CONSPECTUS: Localized molecular orbitals (LMO) not only serve as an important bridge between chemical intuition and molecular wave functions but also can be employed to reduce the computational cost of many-body methods for electron correlation and excitation. Therefore, how to localize the usually completely delocalized canonical molecular orbitals (CMO) into confined physical spaces has long been an important topic: It has a long history but still remains active to date.



While the known LMOs can be classified into (exact) orthonormal and nonorthogonal, as well as (approximate) absolutely localized MOs, the ways for achieving these can be classified into two categories, *a posteriori* top-down and *a priori* bottom-up, depending on whether they invoke the global CMOs (or equivalently the molecular density matrix). While the top-down approaches have to face heavy tasks of minimizing or maximizing a given localization functional typically of many adjacent local extrema, the bottom-up ones have to invoke some tedious procedures for first generating a local basis composed of well-defined occupied and unoccupied subsets and then maintaining or resuming the locality when solving the Hartree−Fock/Kohn−Sham (HF/KS) optimization condition.

It is shown here that the good of these kinds of approaches can be combined together to form a very efficient hybrid approach that can generate the desired LMOs for any kind of gapped molecules. Specifically, a top-down localization functional, applied to individual small subsystems only, is minimized to generate an orthonormal local basis composed of functions centered on the preset chemical fragments. The familiar notion for atomic cores, lone pairs, and chemical bonds emerges here automatically. Such a local basis is then employed in the global HF/KS calculation, after which a least action is taken toward the final orthonormal localized molecular orbitals (LMO), both occupied and virtual. This last step is very cheap, implying that, after the CMOs, the LMOs can be obtained essentially for free. Because molecular fragments are taken as the basic elements, the approach is in the spirit of "from fragments to molecule". Two representatives of highly conjugated molecules, that is, $C_{12}H_2$ and C_{60}, are taken as showcases for demonstrating the success of the proposed approach. The use of the so-obtained LMOs will lead naturally to low-order scaling post-HF/KS methods for electron correlation or excitation. In addition, the underlying fragment picture allows for easy and pictorial interpretations of the correlation/excitation dynamics.

1. INTRODUCTION

It has long been known that localized molecular orbitals (LMOs) serve as an important bridge between chemical intuition (Lewis structure, functional group, etc.) and molecular wave functions, which are otherwise delocalized throughout the physical space. In addition, LMOs can be utilized to reduce the computational scaling of correlated wave function methods, since the number of virtual orbitals needed for the correlation of each electron pair may then be reduced dramatically. A large number of unitary localization schemes[1−24] and optimization algorithms[25−28] have been proposed in the past. They all work well for the occupied space. However, only few of them[13−23] are robust also for the full virtual space. At variance with such unitary transformations, a nonsymmetric transformation was also proposed[29,30] to construct nonorthogonal LMOs (NOLMOs). Due to the absence of "orthogonalization tails", such NOLMOs are more localized than orthonormal LMOs. Yet, they cannot readily be utilized in many-body methods for correlation and excitation, except for those that involve only the manifold of single excitations, for example, time-dependent density functional theory (TD-DFT). There exists also a third type of approach, the so-called locally projected self-consistent field (SCF),[31−33] which by construction generates "absolutely localized MOs" (ALMOs), each of which is expanded in terms only of the basis functions belonging to a given molecular component. Although they are very useful in interpreting various physical phenomena, such ALMOs do not satisfy the Brillouin condition and are hence intrinsically approximate.

From an algorithmic point of view, the available unitary localization schemes can be classified into two classes, *a posteriori* top-down[1−21] and *a priori* bottom-up.[21−24] The former starts with the global CMOs or molecular density matrix already satisfying the Brillouin condition and tries to obtain the

Special Issue: Beyond QM/MM: Fragment Quantum Mechanical Methods

Received: February 28, 2014
Published: July 14, 2014

   dx.doi.org/10.1021/ar500082t | Acc. Chem. Res. 2014, 47, 2758−2767

LMOs via some locality criteria, $\Omega[\{\psi_i\}_{i=1}^N]$ defined only by the CMOs $\{\psi_i\}_{i=1}^N$ and intrinsic quantities such as multipole moments, overlap matrix, etc. In contrast, the latter starts with an initial local basis $\{\phi_\mu\}_{\mu=1}^K$ ($K \geq N$) and tries to preserve the locality as much as possible when the Brillouin condition is solved iteratively. Notwithstanding this distinction, the LMOs by the two classes of approaches are localized in similar physical spaces.

The locality criteria invoked by the popular top-down localization schemes can be written into a generic form

$$\Omega[\{\psi_i\}] = \sum_i^N |\vec{A}_{ii}|^2, \quad \vec{A}_{ii} = \langle\psi_i|\vec{A}_i|\psi_i\rangle \tag{1}$$

where $\vec{A}_i^{\mathrm{BF}} = (x, y, z)$ corresponds to the well-known Foster−Boys (FB) functional,[1,2] maximization of which gives rise to the LMOs that have maximal separation between the orbital centers. It can also be recast into a form that minimizes the sum of the second moments (variances) $(\Omega_2)_i$, namely,

$$\begin{aligned} {}_1\Omega_2 &= \sum_i^N (\Omega_2)_i, \quad (\Omega_{2m})_i = \langle\psi_i|(\vec{r} - \langle\psi_i|\vec{r}|\psi_i\rangle)^{2m}|\psi_i\rangle \\ &= [(\sigma_{2m})_i]^{2m} \end{aligned} \tag{2}$$

where $(\sigma_{2m})_i$ is the $2m$-th moment spread of orbital ψ_i. The problems with the FB localization are twofold. First, a small ${}_1\Omega_2$ does not imply that every second moment $(\Omega_2)_i$ is small. Second, even if an orbital has a thick tail, a small ${}_1\Omega_2$ can still occur. For these reasons, Jørgensen[19,20] introduced a modified FB functional

$${}_n\Omega_{2m} = \sum_i^N [(\Omega_{2m})_i]^n = \sum_i^N [(\sigma_{2m})_i]^{2mn} \tag{3}$$

where the power n is called "penalty exponent". For a given moment $2m$, minimization of ${}_n\Omega_{2m}$ with $n > 1$ will tighten those orbitals with largest orbital spreads (OS), as compared with $n = 1$. However, this comes at the expense that the orbitals with smaller OSs will become loosened. In practice, $n = 2$ turns out to be a good choice.[19] For comparison, $n < 1$ does the opposite. The case of $n = 0.5$ and $m = 1$, not yet considered before, is particularly interesting, because it amounts to directly minimizing the standard (second moment) OSs $(\sigma_2)_i$. Likewise, for a given power n, $m > 1$ will tighten the tails of the LMOs, as compared with $m = 1$.

The choice of $\vec{A}_i^{\mathrm{ER}} = \int d^3\vec{r}'|\psi_i(r')|^2/|\vec{r} - \vec{r}'|$ in eq 1 corresponds to the Edmiston−Ruedenberg (ER) functional,[3,4] maximization of which yields orbitals that have maximal self-repulsion and hence minimal exchange energy. Similarly, the choice of $\vec{A}_i = |\psi_i(r)|^2$ corresponds to the von Neissen functional,[6] which amounts to maximizing the sum of the charge density self-overlaps. The Pipek−Mezey (PM) functional[8] is defined by $\vec{A}_i^{\mathrm{PM}} = (P_{\mathrm{A}}, P_{\mathrm{B}}, \ldots)$, with P_{A} being the projector of atom A. While the Mulliken population analysis was used originally, other population analyses are also possible.[9] In essence, the numbers of atoms constituting the LMOs are minimized here.

Once a localization functional is chosen, any optimization algorithm can be employed.[25−27] Given a good start, asymptotically linear scaling can eventually be achieved,[28] demonstrated only for the occupied space though. Still, however, one has to remember that any localization functional for the virtual space has too many adjacent local extrema, implying that the optimization will become more and more difficult as the size (measured by the number of atoms or basis functions) of the systems increases. The question to be asked is then: Is it possible to alleviate the optimization? It will be shown here that this is indeed possible by combining the good of the top-down and bottom-up schemes. Briefly, the top-down localization is applied only to small subsystems for generating a well-behaved local basis. A least-change action is then taken toward the final LMOs, both occupied and virtual. The computational overhead in preparing the local basis is overcompensated by the reduced number of SCF iterations, whereas the least-change step costs essentially nothing. That is, after the CMOs, the LMOs can be obtained for free. Because molecular fragments are taken here as the basic elements, the approach is in the spirit of "from fragments to molecule".[21] Different from other fragmentation schemes where only the energy is "conquered", both the energy and wave function are here divided and conquered.

2. THE LEAST-CHANGE ALGORITHM

Suppose we have an orthonormal local basis $\{|\phi_\mu\rangle\}_{\mu=1}^K$ whose occupied components can be projected out as

$$\begin{aligned} |\bar{\phi}_\mu\rangle &= \hat{P}_{\mathrm{o}}|\phi_\mu\rangle = \sum_i^N |\psi_i\rangle\langle\psi_i|\phi_\mu\rangle = \sum_i^N |\psi_i\rangle C_{\mu i}^*, \\ C_{\mu i} &= \langle\phi_\mu|\psi_i\rangle \end{aligned} \tag{4}$$

where $\{|\psi_i\rangle\}_{i=1}^N$ are the occupied CMOs. The functions $\{|\bar{\phi}_\mu\rangle\}_{\mu=1}^K$ are generally nonorthogonal, with the overlap being

$$S_{\mu\nu} = \langle\bar{\phi}_\mu|\bar{\phi}_\nu\rangle = [\mathbf{C}\mathbf{C}^\dagger]_{\mu\nu} = \langle\phi_\mu|\hat{P}_{\mathrm{o}}|\phi_\nu\rangle \tag{5}$$

which is nothing but the matrix representation of $\hat{P}_{\mathrm{o}}$ in the local basis $\{|\phi_\mu\rangle\}_{\mu=1}^K$. A singular value decomposition can be applied to decompose $\mathbf{C}$ as

$$\mathbf{C} = \mathbf{L}\lambda\mathbf{R}^\dagger, \quad \mathbf{L}^\dagger\mathbf{L} = \mathbf{I}_K, \quad \mathbf{R}^\dagger\mathbf{R} = \mathbf{I}_N \tag{6}$$

where λ is a $K \times N$ real-valued diagonal matrix, whereas $\mathbf{L}$ and $\mathbf{R}$ are the $K \times K$ and $N \times N$ unitary eigenvector matrices of $\mathbf{C}\mathbf{C}^\dagger$ and $\mathbf{C}^\dagger\mathbf{C}$, respectively,

$$\mathbf{C}\mathbf{C}^\dagger\mathbf{L} = \mathbf{L}\lambda^2 \tag{7}$$

$$\mathbf{C}^\dagger\mathbf{C}\mathbf{R} = \mathbf{R}\lambda^2 \tag{8}$$

It is always possible to choose the phases of the column vectors of $\mathbf{L}$ and $\mathbf{R}$ such that $\forall\lambda_i \geq 0$. Two sets of orthonormal orbitals can now be introduced

$$|\psi_i'\rangle = \sum_{j=1}^N |\psi_j\rangle R_{ji}, \quad i = 1, \ldots N \tag{9}$$

$$|\phi_\mu'\rangle = \sum_{\nu=1}^K |\phi_\nu\rangle L_{\nu\mu}, \quad \mu = 1, \ldots K \tag{10}$$

Since $\mathbf{L}$ diagonalizes the density matrix $\mathbf{S}$, $\{|\phi_\mu'\rangle\}$ can be interpreted as natural orbitals represented in the local basis $\{|\phi_\mu\rangle\}$. For instance, if $\{|\phi_\mu\rangle\}$ are chosen to be the orthonormal atomic orbitals (OAOs) centered on an atom, $\{|\phi_\mu'\rangle\}$ would be the natural atomic orbitals (NAOs).[34] Likewise, if $\{|\phi_\mu\rangle\}$ are the OAOs centered on two atoms, $\{|\phi_\mu'\rangle\}$ would be the natural bond orbitals (NBOs).[35] However, $\{|\phi_\mu'\rangle\}$ are not yet within the occupied space. Only those ϕ_μ' associated with $\lambda_i \approx 1$ are

dx.doi.org/10.1021/ar500082t | Acc. Chem. Res. 2014, 47, 2758−2767

good candidates for the occupied orbitals, while those associated with $\lambda_i \approx 0$ belong to the unoccupied space. In contrast, the functions $\{|\psi_i'\rangle\}$ all stay within the occupied space. For those column eigenvectors of **L** and **R** of nonzero eigenvalues, the corresponding functions $\{|\phi_\mu'\rangle\}$ and $\{|\psi_i'\rangle\}$ form pairs, namely,

$$\langle\phi_\mu'|\psi_i'\rangle = [\mathbf{L}^\dagger\mathbf{C}\mathbf{R}]_{\mu i} = \delta_{\mu i}\lambda_i \tag{11}$$

In this case, $\{|\psi_i'\rangle\}$ can be recognized as canonical orthonormalization of the projected functions $\{|\bar{\phi}_\mu\rangle\}$ (eq 4),

$$|\psi_i'\rangle = \sum_j^N |\psi_j\rangle(\mathbf{C}^\dagger\mathbf{L}\lambda^{-1})_{ji} = \sum_\mu |\bar{\phi}_\mu\rangle L_{\mu i}\lambda_i^{-1} \tag{12}$$

Therefore, even if $\{|\bar{\phi}_\mu\rangle\}$ projected from $\{|\phi_\mu\rangle\}$ are local, $\{|\psi_i'\rangle\}$ are not necessarily local to the same extent, since canonical orthonormalization often ruins the locality. Nevertheless, if $\{|\phi_\mu\rangle\}$ are all located in a particular region of space (e.g., some fragment of a molecule), the resulting $\{|\psi_i'\rangle\}$ functions would strictly be localized in the same region. This point was employed by Zoboko and Mayer[12] to construct occupied LMOs $\{|\psi_i'\rangle\}$ that are orthonormal and delocalized within a fragment but are nonorthogonal between fragments. A least change from $\{|\bar{\phi}_\mu\rangle\}$ is only achieved by the well-known Löwdin symmetric orthonormalization. However, some constraints must be imposed on the initial local basis $\{|\phi_\mu\rangle\}_{\mu=1}^K$, namely, $K = N$ and $\forall\lambda_i > 0$. Under such conditions, the desired occupied orthonormal LMOs $\{|\tilde{\phi}_\mu\rangle\}$ can be obtained as follows

$$|\tilde{\phi}_\mu\rangle = \sum_{\nu=1}^N |\bar{\phi}_\nu\rangle[\mathbf{S}^{-1/2}]_{\nu\mu} = \sum_{i=1}^N |\psi_i\rangle T_{i\mu},$$
$$\mathbf{T} = \mathbf{C}^\dagger\mathbf{S}^{-1/2} = \mathbf{R}\mathbf{L}^\dagger \tag{13}$$

The least change is of course defined in a least-squares sense. To see this, we can minimize the following functional

$$\sum_{\mu=1}^N \langle\tilde{\phi}_\mu - \phi_\mu|\tilde{\phi}_\mu - \phi_\mu\rangle = 2N - \mathrm{Tr}(\mathbf{C}\mathbf{T} + \mathbf{T}^\dagger\mathbf{C}^\dagger) \tag{14}$$

$$= 2N - 2\mathcal{R}\sum_i^N A_{ii}\lambda_i, \quad \mathbf{A} = \mathbf{R}^\dagger\mathbf{T}\mathbf{L} \tag{15}$$

Since **A** is unitary (i.e., $|A_{ii}| \le 1$) and $\lambda_i > 0$, the minimum is achieved[36] at $\mathbf{A} = 1$, thereby leading to $\mathbf{T} = \mathbf{R}\mathbf{L}^\dagger$ (eq 13). As such, the so-obtained $\{|\tilde{\phi}_\mu\rangle\}_{\mu=1}^N$ are most similar to the initial local basis $\{|\phi_\mu\rangle\}_{\mu=1}^N$.

Minimization of functional 14 is equivalent to maximization of $\sum_{\mu=1}^N \mathcal{R}\langle\tilde{\phi}_\mu|\phi_\mu\rangle$, which can further be generalized to maximization of the following functional

$$\Xi = \sum_{\mu=1}^N [\langle\tilde{\phi}_\mu|\phi_\mu\rangle + \langle\phi_\mu|\tilde{\phi}_\mu\rangle]^n \tag{16}$$

This is in the spirit of the modified FB functional 3. Use of $n = 2$ was made by Angeli et al.[10] to construct LMOs. However, at variance with $n = 1$, which permits an analytic solution (eq 14), the cases of $n > 1$ require some optimization algorithm, just like the aforementioned top-down localization schemes. Therefore, the least-change algorithm (eq 14) is preferred.

The above least-change algorithm can also be applied to the virtual CMOs by replacing $\hat{P}_o$ (eq 4) with $\hat{P}_v = \sum_{i=N+1}|\psi_i\rangle\langle\psi_i|$. Since the unitary transformations going from CMOs to LMOs are to be carried out for the occupied and virtual spaces separately, the overall transformation matrix **T** is block-diagonal

$$\mathbf{T} = \begin{pmatrix} \mathbf{C}_{oo}^\dagger\mathbf{S}_{oo}^{-1/2} & 0 \\ 0 & \mathbf{C}_{vv}^\dagger\mathbf{S}_{vv}^{-1/2} \end{pmatrix}$$
$$= \begin{pmatrix} \mathbf{C}_{oo}^\dagger(\mathbf{C}_{oo}\mathbf{C}_{oo}^\dagger)^{-1/2} & 0 \\ 0 & \mathbf{C}_{vv}^\dagger(\mathbf{C}_{vv}\mathbf{C}_{vv}^\dagger)^{-1/2} \end{pmatrix} \tag{17}$$

Further expanding the CMOs $\{|\psi_i\rangle\}$ in terms of the initial local basis $\{|\phi_\mu\rangle\}$,

$$(|\psi\rangle_o, |\psi\rangle_v) = (|\phi\rangle_o, |\phi\rangle_v)\begin{pmatrix} \mathbf{C}_{oo} & \mathbf{C}_{ov} \\ \mathbf{C}_{vo} & \mathbf{C}_{vv} \end{pmatrix} \tag{18}$$

we can express the LMOs $\{|\tilde{\phi}_\mu\rangle\}$ as

$$(|\tilde{\phi}\rangle_o, |\tilde{\phi}\rangle_v) = (|\phi\rangle_o, |\phi\rangle_v)\mathbf{U} \tag{19}$$

where

$$\mathbf{U} = \mathbf{C}\mathbf{T} = \begin{pmatrix} \mathbf{C}_{oo} & \mathbf{C}_{ov} \\ \mathbf{C}_{vo} & \mathbf{C}_{vv} \end{pmatrix}\begin{pmatrix} \mathbf{C}_{oo}^\dagger(\mathbf{C}_{oo}\mathbf{C}_{oo}^\dagger)^{-1/2} & 0 \\ 0 & \mathbf{C}_{vv}^\dagger(\mathbf{C}_{vv}\mathbf{C}_{vv}^\dagger)^{-1/2} \end{pmatrix} \tag{20}$$

which can also be rewritten as

$$\mathbf{U} = \begin{pmatrix} \mathbf{I} & -\mathbf{X}^\dagger \\ \mathbf{X} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{N}_+^{-1} & 0 \\ 0 & \mathbf{N}_-^{-1} \end{pmatrix} \tag{21}$$

$$\mathbf{N}_+ = (\mathbf{I} + \mathbf{X}^\dagger\mathbf{X})^{1/2}, \quad \mathbf{N}_- = (\mathbf{I} + \mathbf{X}\mathbf{X}^\dagger)^{1/2} \tag{22}$$

$$\mathbf{X} = \mathbf{C}_{vo}\mathbf{C}_{oo}^{-1} = -(\mathbf{C}_{vv}^\dagger)^{-1}\mathbf{C}_{ov}^\dagger \tag{23}$$

Equations 18 and 19, presented differently in ref 17, imply that the **F** matrix in the local basis $\{|\phi_\mu\rangle\}$ is first diagonalized via **C** and then block-diagonalized via **T**,

$$(\mathbf{U}^\dagger\mathbf{F}\mathbf{U})_{vo} = (\mathbf{T}^\dagger\mathbf{C}^\dagger\mathbf{F}\mathbf{C}\mathbf{T})_{vo} = 0 = (\mathbf{U}^\dagger\mathbf{F}\mathbf{U})_{ov} \tag{24}$$

The resulting full LMOs $\phi^{LMO} = \{|\tilde{\phi}_\mu\rangle\}$ are related to the CMOs $\phi^{CMO} = \{|\psi_p\rangle\}$, initial local functions $\phi^{pLMO} = \{|\phi_\mu\rangle\}$, and atomic orbitals (AOs) $\phi^{AO} = \{|\chi_\mu\rangle\}$ as follows:

$$\phi^{LMO} = \phi^{pLMO}\mathbf{U} = \phi^{AO}\mathbf{L}^{pLMO}\mathbf{U} = \phi^{CMO}\mathbf{T}^\dagger \tag{25}$$

At variance with the above "top-down, least-change" algorithm, we can also design a "bottom-up, least-change" algorithm without recourse to the CMOs. Specifically, the requirement of eq 24 dictates that the decoupling matrix **X** should satisfy the following Riccatti equation

$$\mathbf{F}_{vo} + \mathbf{F}_{vv}\mathbf{X} - \mathbf{X}\mathbf{F}_{oo} - \mathbf{X}\mathbf{F}_{ov}\mathbf{X} = 0. \tag{26}$$

For a given **F** matrix, eq 26 can be solved iteratively with the efficient algorithms proposed when reducing the matrix Dirac equation down to the exact two-component equation.[37] Such (micro)iterations usually converge very fast, as long as there exists a gap between the largest eigenvalue of $\mathbf{F}_{oo}$ and the smallest eigenvalue of $\mathbf{F}_{vv}$. The resulting **X** matrix can be employed to construct the unitary matrix **U** (eq 21) and then the density matrix $\mathbf{P} = \mathbf{U}\mathbf{n}\mathbf{U}^\dagger$ that is needed for constructing a new **F** matrix in the (macro-) SCF iterations. The global CMOs are hence avoided completely. If needed, they can be recovered

dx.doi.org/10.1021/ar500082t | Acc. Chem. Res. 2014, 47, 2758−2767

by virtue of the relation $\mathbf{C} = \mathbf{U}\mathbf{T}^{\dagger}$, with $\mathbf{T}^{\dagger}$ being the solutions of the following eigenvalue equations

$$\tilde{\mathbf{F}}_{oo}\mathbf{T}^{\dagger}_{oo} = \mathbf{T}^{\dagger}_{oo}\varepsilon_{o}, \quad \mathbf{T}^{\dagger}_{oo} = \mathbf{N}_{+}\mathbf{C}_{oo} \tag{27}$$

$$\tilde{\mathbf{F}}_{vv}\mathbf{T}^{\dagger}_{vv} = \mathbf{T}^{\dagger}_{vv}\varepsilon_{v}, \quad \mathbf{T}^{\dagger}_{vv} = \mathbf{N}_{-}\mathbf{C}_{vv} \tag{28}$$

where $\tilde{\mathbf{F}}_{oo}$ and $\tilde{\mathbf{F}}_{vv}$ are the diagonal blocks of $\mathbf{U}^{\dagger}\mathbf{F}\mathbf{U}$

$$\tilde{\mathbf{F}}_{oo} = \mathbf{N}_{+}^{-1}(\mathbf{F}_{oo} + \mathbf{X}^{\dagger}\mathbf{F}_{vo} + \mathbf{F}_{ov}\mathbf{X} + \mathbf{X}^{\dagger}\mathbf{F}_{vv}\mathbf{X})\mathbf{N}_{+}^{-1} \tag{29}$$

$$\tilde{\mathbf{F}}_{vv} = \mathbf{N}_{-}^{-1}(\mathbf{F}_{vv} - \mathbf{X}\mathbf{F}_{ov} - \mathbf{F}_{vo}\mathbf{X}^{\dagger} + \mathbf{X}\mathbf{F}_{oo}\mathbf{X}^{\dagger})\mathbf{N}_{-}^{-1} \tag{30}$$

Instead of solving eq 26 for a given $\mathbf{F}$, the unitary transformation matrix (denoted as $\tilde{\mathbf{U}}$) can also be constructed[21] by solving the simpler linear Sylvester equation

$$\mathbf{F}^{(i)}_{vo} + \mathbf{F}^{(i)}_{vv}\mathbf{X}^{(i)} - \mathbf{X}^{(i)}\mathbf{F}^{(i)}_{oo} = 0 \tag{31}$$

by diagonalizing both $\mathbf{F}^{(i)}_{oo}$ and $\mathbf{F}^{(i)}_{vv}$. The starting point is $\mathbf{F}^{(0)} = \mathbf{F}$. In terms of the so-obtained $\mathbf{X}^{(i)}$, the $\mathbf{U}^{(i)}$ and hence $\mathbf{F}^{(i+1)} = \mathbf{U}^{(i)\dagger}\mathbf{F}^{(i)}\mathbf{U}^{(i)}$ matrices can be constructed. The iterations continue until $\mathbf{F}^{(i+1)}_{vo} = \mathbf{X}^{(i+1)} = 0$. The final unitary matrix $\tilde{\mathbf{U}}$ is the product of all the $\mathbf{U}^{(i)}$ and may differ from the $\mathbf{U}$ (eq 21) resulting from eq 26 by some unitary matrix $\mathbf{U}'$, i.e., $\tilde{\mathbf{U}} = \mathbf{U}\mathbf{U}'$. Numerical experiences[21] showed that $\mathbf{U}'$ is very close to unit. That is, the LMOs constructed by solving eqs 26 and 31 are very similar to each other. The former are of course identical to those by the top-down localization (eq 20). The present "bottom-up, least-change" localization is an alternative to that proposed by Malrieu and co-workers[22] through the diagonalization of the (dressed) CIS (single configuration interaction) matrix. Although only two blocks (occupied and virtual) have been considered here, the formulation can also be extended to multiple blocks,[36] such that ROHF/ROKS (restricted open-shell HF/KS) or CASSCF (complete active space SCF) inactive, active, and virtual orbitals can also be localized in a bottom-up manner, again alternative to that proposed by Malrieu and co-workers.[23]

3. THE LOCAL BASIS

The above least-change algorithm ensures that the transformed set of orbitals $\{\tilde{\phi}_{\mu}\}$ is most similar to whatever initial basis $\{\phi_{\mu}\}$. As such, it is absolutely essential to first prepare an initial local basis that is linearly independent and (optionally) orthonormal and, in particular, can clearly be decomposed into occupied and virtual subsets. The least-change localization is bound to fail if some occupied (virtual) ϕ_{μ} functions become instead the dominant components of the final virtual (occupied) LMOs. There can be many different ways to construct such a local basis, such as

(1) the orthonormal NBOs[35] obtained through a series of diagonalizations of the one- and two-centered blocks of the molecular density matrix in the orthonormal NAO basis,[34]
(2) the nonorthogonal bond orbitals constructed with maximally localized hybrid AOs,[10]
(3) the hierarchically orthonormalized atomic cores, bond orbitals, and virtual OAOs,[22,23]
(4) the fragment orbitals (FOs) from separate diagonalizations of the fragment density matrices projected from the molecular density matrix in the AO basis[12] (note, the FOs are orthonormal within a fragment but nonorthogonal between fragments),
(5) the orthonormal regional orbitals (ROs)[14,15] from separate diagonalizations of the regional blocks of the molecular density matrix in the OAO basis,
(6) the orthonormal "primitive fragment LMO" (pFLMO)[21] obtained from subsystem SCF calculations,
(7) the decomposition of the AOs into minimal Boys LMOs (including the occupied and valence virtual orbitals) and protohard atomic virtual orbitals,[16]
(8) the decomposition of the AOs into minimal least-change LMOs (including the core/valence occupied and valence virtual orbitals) and atomic virtual orbitals,[17,18]
(9) the Cholesky decomposition of the molecular density matrix in the AO basis.[11]

Scheme 9 is already a noniterative approach for constructing occupied LMOs with locality inherited from the sparsity of the density matrix. However, the so-obtained LMOs are usually less local than, for example, the Boys LMOs and should hence be regarded as prelocal. They may be taken as the initial guess to speed up subsequent top-down localizations.[28] The idea can also be extended to virtual LMOs, not yet done though. Both schemes 7 and 8 amount to manipulating directly the AOs without explicit recourse to chemical intuition. These two kinds of local bases are very similar in structure and dimension. However, unlike the former,[16] the latter[17,18] does not rely on the quality of a minimal AO set (e.g., STO-3G). Schemes 1−6 all take chemical intuition (Lewis structure, functional group, or fragment in general) as input. Apart from some algorithmic differences, they differ mainly in the definition of fragments. For instance, the chemical distinction of core, lone-pair, and few-center bond orbitals is imposed from the outset in schemes 1−4, while they are realized in the end by localizing the subsystem orbitals in scheme 6. In contrast, the ROs by scheme 5 are delocalized throughout the prechosen regions.

Given so many ways for preparing the initial local basis, the preference can only be judged from whether the scheme is universal (i.e., effective for all kinds of gapped molecules and AOs), operationally simple, and populationally transferable (i.e., no significant change in orbital occupation when transformed to the final LMOs). Taking all the criteria into account, we think scheme 6 is most preferred. Here, the whole molecule is first divided into N_F disjoint fragments. Each fragment is then capped with a buffer, which is just those parts of the molecule directly bonded to the considered fragment. The dangling bonds are finally saturated with hydrogen atoms to form N_F closed subsystems. Taking 1,3,5,7,9,11-dodecahexayne ($C_{12}H_2$) as an example, one triple bond can be chosen as a fragment, with its neighboring triple bonds as the buffer. We then have two identical subsystems of C_4H_2 and four identical subsystems of C_6H_2, see Figure 1. The conventional SCF calculations can be carried out for the unique subsystems in parallel. Since the subsystems are very small, any top-down localization schemes can be applied to generate subsystem LMOs. By virtue of the Löwdin population, the LMOs centered on every fragment can be identified. The basis functions of the buffer atoms are retained since they are necessary to describe the local environment of the fragments. However, the basis functions of the added link atoms have to be projected out. The resulting fragment orbitals $\{\xi^{I}_{\mu}\}$ are nonorthogonal and linearly dependent. Instead of performing a canonical orthonormalization, we consider a two-step procedure: The linear independency is first achieved by eliminating those ξ^{I}_{μ} that have largest overlaps with the eigenvectors $\{v_p\}$ of the overlap

dx.doi.org/10.1021/ar500082t | Acc. Chem. Res. 2014, 47, 2758−2767

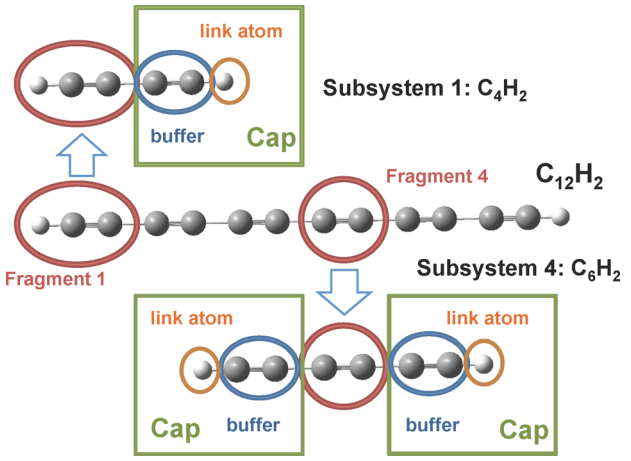


Figure 1. Fragmentation of $C_{12}H_2$.

matrix $\langle \xi_\mu^I | \xi_\nu^J \rangle$ with eigenvalues smaller than a threshold (e.g., 10^{-6}). In case of degeneracy, those ν_p and ξ_μ of larger OSs are first eliminated. The surviving $\{\xi_\mu^I\}$ are finally symmetrically orthonormalized, leading to the desired linearly independent and orthonormal pFLMO $\phi^{\mathrm{pLMO}} = \{\phi_\mu\}$ in eq 25. The same procedure can also be applied to the AOs. In the end, we have the same number of local pFLMOs and OAOs but the former have a clear separation of occupied and virtual subsets. Taking the pFLMOs as the basis, the SCF calculation for the whole molecule can be carried out, which usually requires fewer iteration cycles to converge than that starting with the superposition of atoms. The computational overhead for preparing the pFLMOs is hence compensated. The one-step transformation **U** (eq 20 or 21) costs essentially nothing. Therefore, compared with the conventional top-down localizations, both occupied and virtual LMOs can be obtained for free. Since they are localized on the same fragments as the pFLMOs, the so-obtained LMOs have been called "fragment LMOs" (FLMO).[21] At variance with such a one-step top-down-like localization, we can also perform an iterative bottom-up localization by solving eq 26 or 31 as discussed above. In either case, the global SCF in the pFLMO basis can be regarded as a "conquer" step. Because of this, the "divide" step for generating the pFLMOs is much less critical than other fragmentation schemes that "conquer" only the energy but not the wave function.

Finally, we remark that the present projection (eq 5) and symmetric orthonormalization (eq 13) procedures for transforming the pFLMOs to the FLMOs can also be viewed as a special kind of block-diagonalization of the molecular density matrix in the pFLMO basis. However, at variance with the series of Jacobi rotations employed in the construction of the natural[13] and regional (RLMO)[14,15] LMOs, the locality is preserved here by the least-change algorithm, which is unique and rigorous. It will be shown below that the present FLMOs are indeed very similar to those by the top-down localization. In contrast, in the case that the locality of the density matrix is very low, for example, for highly conjugated systems,[38] the locality of the RLMOs is also very low. This explains why some frozen RLMOs have to be reactivated in the elongation approach.[39]

4. ILLUSTRATIONS

Except for $C_{12}H_2$ shown in Figure 1, C_{60} is taken as another representative of highly conjugated systems. Here, one five-member ring is chosen as a fragment, with the five adjacent six-member rings as the buffer. After adding in the link hydrogen atoms, we obtain 12 identical subsystems $C_{20}H_{10}$. The trust region algorithm[26] for minimizing the modified FB functional (eq 3) was implemented into the BDF pakcage.[40] The Dunning cc-pVDZ, aug-cc-pVDZ, aug-cc-pVTZ and cc-pVQZ basis sets[41] were employed in the HF calculations.

The first point to be addressed is the quality of the initial pFLMOs. This can be monitored by their occupation numbers, $n_\mu = 2\langle \phi_\mu | \hat{P}_o | \phi_\mu \rangle$. Ideally, n_μ is equal to 2 (0) for all the occupied (virtual) pFLMO orbitals. Deviations from such ideal occupations are plotted in Figure 2 for the pFLMOs obtained by minimizing the subsystem functionals ${}_n\Omega_2$ with $n = 0.5$, 1, 2. It is seen that, for both molecules, the n_μ of the pFLMOs are very close to their ideal values. The deviations from the total number of electrons are only −0.11 ($n = 0.5$), −0.09 ($n = 1$), and −0.08 ($n = 2$) for $C_{12}H_2$ and −0.24 ($n = 0.5$), −0.24 ($n = 1$), and −0.25 ($n = 2$) for C_{60}. In this regard, the pFLMOs are much better than the symmetrically orthogonalized local bases carefully optimized by Jorgensen,[17] where the deviations may be up to dozens of electrons for similar systems (the situation may be improved, but only slightly, by further localizing

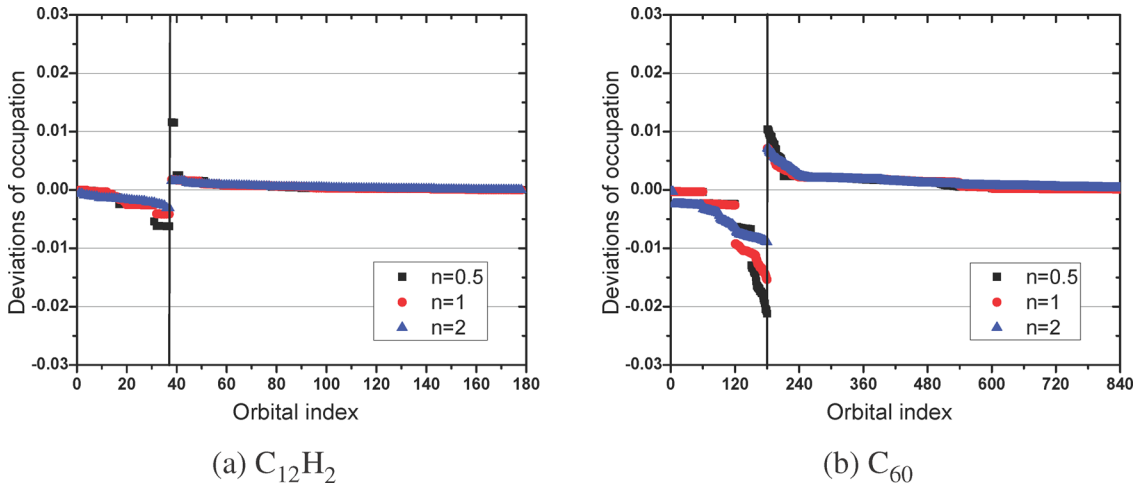


Figure 2. Deviation of individual pFLMO occupations. Vertical lines separate occupied (left) and virtual spaces. cc-pVDZ results.

dx.doi.org/10.1021/ar500082t | Acc. Chem. Res. 2014, 47, 2758−2767

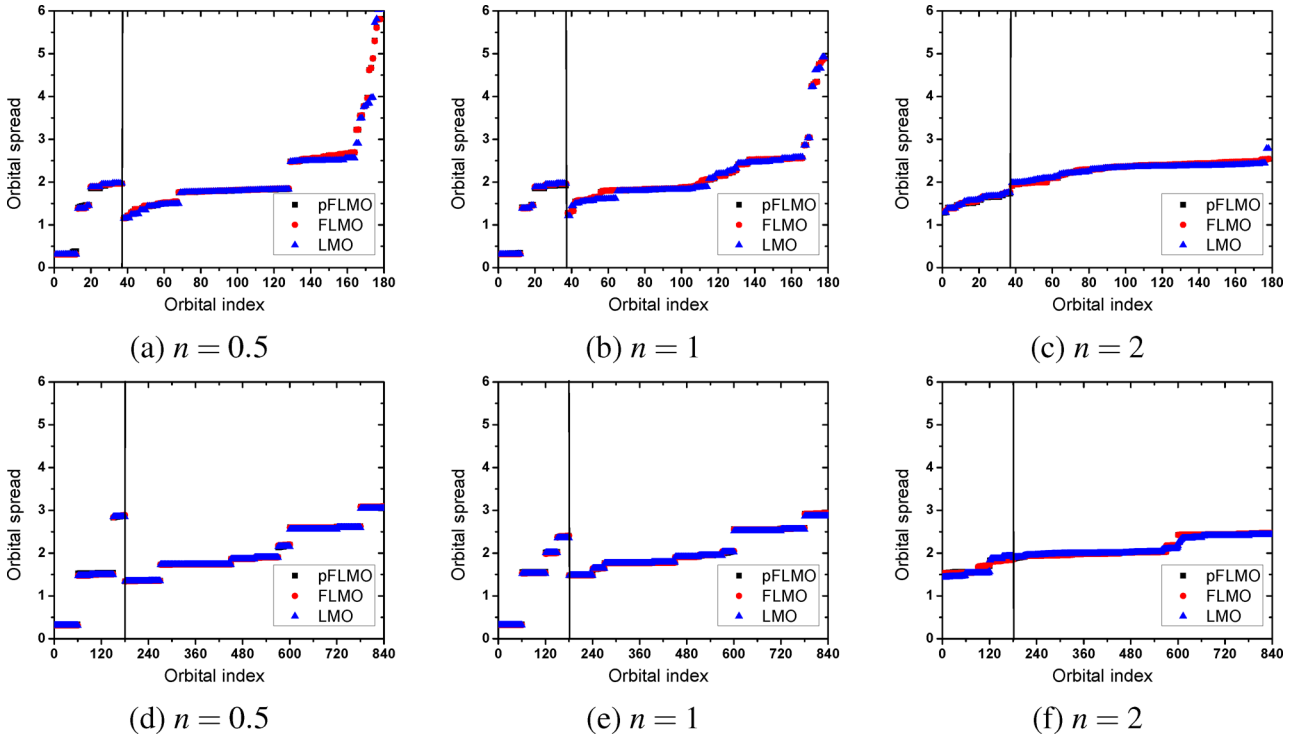


Figure 3. Orbital spreads for the pFLMOs, FLMOs, and LMOs of $C_{12}H_2$ (a–c) and C_{60} (d–f). Vertical lines separate occupied (left) and virtual spaces. cc-pVDZ results.

separately the core, valence, and unoccupied components of the valence basis[18]).

The OSs $(\sigma_2)_i$ for the pFLMOs, FLMOs, and (top-down) LMOs of $C_{12}H_2$ and C_{60} are plotted in Figure 3. It is seen that, for a given ${}_n\Omega_2$, the OS curves for the three sets of orbitals essentially coincide for both the occupied and virtual spaces, confirming again the high quality of the pFLMOs on one hand and the great similarity between the FLMOs and LMOs on the other. It is also clear that, when going from ${}_{0.5}\Omega_2$ via ${}_1\Omega_2$ to ${}_2\Omega_2$, the most local orbitals become slightly loosened, whereas the most delocalized orbitals are strongly tightened, leading to $\sigma_{min}(n = 0.5) < \sigma_{min}(n = 1) \ll \sigma_{min}(n = 2)$ for the minimal (σ_{min}) and $\sigma_{max}(n = 0.5) > \sigma_{max}(n = 1) \gg \sigma_{max}(n = 2)$ for the maximal (σ_{max}) OSs, see Table 1. Such behaviors can well be understood by looking at the corresponding weight functions $\sqrt{x}$, x, and x^2. Obviously, during the minimization process, $\sqrt{x}$ and x^2 provide the largest weights for $x < 1$ and $x > 1$, respectively. As a result, ${}_2\Omega_2$ tends to provide balanced localizations of the virtual and occupied LMOs, such that they are localized roughly in the same regions. It is even more so when the penalty exponent n is further increased.[19] In contrast, the OSs of the LMOs by ${}_{0.5}\Omega_2$ and ${}_1\Omega_2$ have a larger span. However, their averaged OSs (σ_{av}) are markedly smaller than those by ${}_2\Omega_2$, $\sigma_{av}(n = 0.5) < \sigma_{av}(n = 1) < \sigma_{av}(n = 2)$. The reason is that the majority of the LMOs by ${}_{0.5}\Omega_2$ and ${}_1\Omega_2$ have smaller OSs than those by ${}_2\Omega_2$. Note that such penalty behaviors are universal.

To check the basis set dependence of the localization, we further carried out HF calculations for $C_{12}H_2$ with the aug-cc-pVDZ, aug-cc-pVTZ, and cc-pVQZ basis sets. It is seen from Figure 4 that while the pFLMO and FLMO OS curves still overlap each other, there do exist tiny differences between the FLMOs and LMOs, particularly in the high OS end. Specifically, in the case of ${}_2\Omega_2$, some of the virtual FLMOs

Table 1. Orbital Spreads (OS) for Various Orbitals of $C_{12}H_2$ and C_{60}[a]

		occupied			virtual		
molecule	orbital	σ_{min}	σ_{max}	σ_{av}	σ_{min}	σ_{max}	σ_{av}
$C_{12}H_2$	CMO	2.44	13.52	8.25	5.03	16.93	9.26
$n = 0.5$	pFLMO	0.31	1.96	1.32	1.16	5.82	2.18
	FLMO	0.31	1.97	1.31	1.16	5.81	2.18
	LMO	0.31	1.97	1.31	1.15	6.04	2.13
$n = 1$	pFLMO	0.32	1.92	1.30	1.28	4.89	2.20
	FLMO	0.32	1.96	1.31	1.28	4.88	2.20
	LMO	0.32	1.97	1.31	1.21	4.92	2.17
$n = 2$	pFLMO	1.28	1.73	1.55	1.91	2.54	2.31
	FLMO	1.31	1.76	1.57	1.91	2.54	2.32
	LMO	1.28	1.75	1.57	1.99	2.79	2.30
C_{60}	CMO	6.61	7.30	6.90	6.01	9.32	7.51
$n = 0.5$	pFLMO	0.33	2.87	1.35	1.34	3.09	2.09
	FLMO	0.31	2.89	1.33	1.34	3.09	2.09
	LMO	0.31	2.85	1.32	1.34	3.04	2.07
$n = 1$	pFLMO	0.34	2.40	1.37	1.48	2.93	2.10
	FLMO	0.32	2.41	1.35	1.48	2.93	2.10
	LMO	0.32	2.35	1.34	1.48	2.86	2.08
$n = 2$	pFLMO	1.53	1.86	1.67	1.87	2.47	2.16
	FLMO	1.53	1.85	1.66	1.91	2.47	2.16
	LMO	1.43	1.92	1.63	1.89	2.43	2.14

[a] ${}_n\Omega_2$ (eq 3) was used for both pFLMOs and LMOs (cc-pVDZ). σ_{min}, σ_{max}, and σ_{av}: minimal, maximal, and averaged OSs.

are somewhat more delocalized than the corresponding LMOs, σ_{max}(virtual FLMO) > σ_{max}(virtual LMO) (see Table 2). The opposite is true in the case of ${}_{0.5}\Omega_2$. The case of ${}_1\Omega_2$ is intermediate, where the relative ordering of σ_{max}(virtual FLMO) and σ_{max}(virtual LMO) are basis set dependent. Actually, the averaged OS is more instructive: σ_{av}(virtual

dx.doi.org/10.1021/ar500082t | Acc. Chem. Res. 2014, 47, 2758–2767

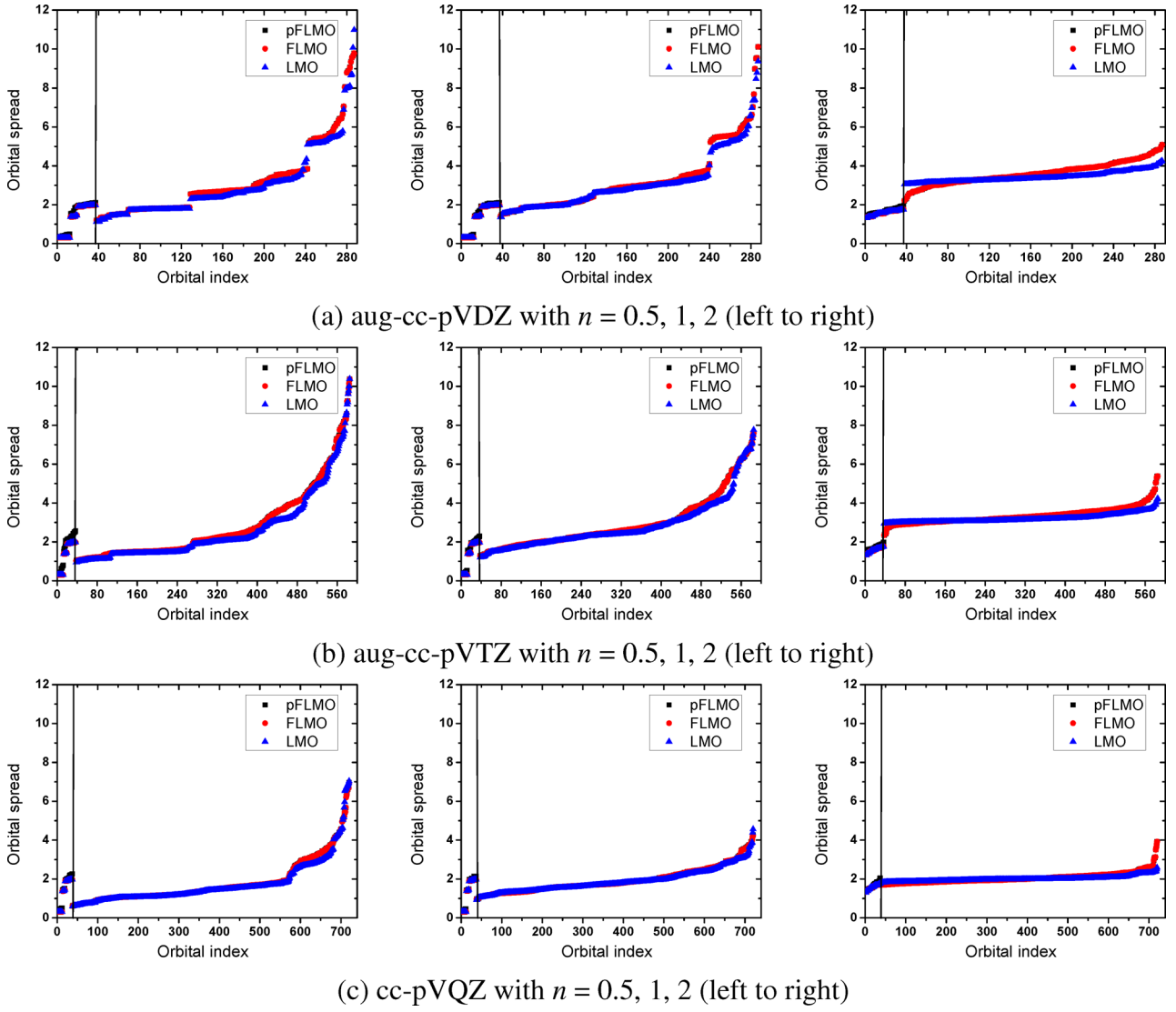


Figure 4. Orbital spreads for the pFLMOs, FLMOs, and LMOs of $C_{12}H_2$ with various basis sets.

FLMO) > σ_{av}(virtual LMO) regardless of the basis set and localization functional. That the FLMOs are somewhat more delocalized than the LMOs for such delocalized electronic structures and such extended basis sets is hardly surprising, since the FLMOs do not fully minimize the localization functional ${}_n\Omega_2$.

The real question is whether the degree of localization of the FLMOs is sufficient. To check this, we take the absolute overlaps, O_{ai},

$$O_{ai} = \int |\phi_a(\vec{r})\phi_i(\vec{r})|\, \mathrm{d}^3\vec{r} \in [0, 1] \tag{32}$$

to measure the significance of particle−hole (p-h) pairs: Those with O_{ai} smaller than a threshold can be neglected. It is seen from Table 3 that essentially no CMO pairs can be neglected, whereas the LMO pairs can significantly be cut off. In particular, the cutoff efficacy follows the ordering ${}_{0.5}\Omega_2 > {}_1\Omega_2 > {}_2\Omega_2$. Therefore, it is ${}_{0.5}\Omega_2$ instead of ${}_n\Omega_2$ ($\geq$2) that should be recommended as a replacement of the original FB functional ${}_1\Omega_2$. In other words, enhancing the locality of the most delocalized virtual LMOs via a large penalty exponent is not really effective for cutting down the p-h pairs. This is an important finding, contrary to the statement[19] that the existence of a few virtual LMOs with large OSs will ruin the efficiency of local correlation/excitation approaches. The cutoff efficacy of the FLMO pairs is marginally lower (by up to 4%) than the LMO ones. This price is acceptable in view of the gained efficiency over the global top-down localization. The use of the FLMOs does lead to very efficient linear-scaling time-dependent density functional theory with a very small prefactor,[21] which can now routinely be applied to capture ca. 400 excited states of any molecules composed of ca. 300 atoms described by ca. 4000 basis functions.[42] In addition, the underlying fragment picture also allows for easy and pictorial interpretations of the excitation dynamics.

5. CONCLUSIONS AND OUTLOOK

The basic requirement for the localizability of CMOs is the existence of a gap between the occupied and virtual spaces. If this is true, no matter how seemingly delocalized the electronic structure looks, the CMOs can always be localized by choosing an appropriate localization functional. The optimization of such a functional is then the task to be accomplished. This is by no means trivial, in particular when the size of systems increases. It has been shown here that such a heavy task of optimization can

dx.doi.org/10.1021/ar500082t | Acc. Chem. Res. 2014, 47, 2758−2767

Table 2. Orbital Spreads for Various Orbitals of $C_{12}H_2$ with Various Basis Sets[a]

			occupied			virtual		
basis		orbital	σ_{min}	σ_{max}	σ_{av}	σ_{min}	σ_{max}	σ_{av}
aug-cc-pVDZ		CMO	0.64	13.52	6.55	4.60	18.43	10.14
	n = 0.5	pFLMO	0.36	2.11	1.45	1.19	9.79	3.17
		FLMO	0.31	1.97	1.31	1.19	9.78	3.17
		LMO	0.31	1.98	1.31	1.13	10.98	2.99
	n = 1	pFLMO	0.35	2.12	1.43	1.39	10.11	3.27
		FLMO	0.32	1.97	1.31	1.39	10.11	3.27
		LMO	0.32	1.98	1.31	1.37	9.38	3.11
	n = 2	pFLMO	1.37	1.95	1.68	2.20	5.08	3.61
		FLMO	1.36	1.77	1.58	2.22	5.09	3.61
		LMO	1.31	1.76	1.58	3.05	4.25	3.42
aug-cc-pVTZ		CMO	0.65	13.52	6.45	5.33	16.09	9.90
	n = 0.5	pFLMO	0.38	2.56	1.64	1.03	10.37	2.82
		FLMO	0.31	1.97	1.31	1.03	10.37	2.82
		LMO	0.31	1.98	1.31	0.95	10.39	2.60
	n = 1	pFLMO	0.39	2.30	1.46	1.28	7.59	2.96
		FLMO	0.32	1.97	1.31	1.27	7.59	2.96
		LMO	0.32	1.97	1.31	1.20	7.77	2.81
	n = 2	pFLMO	1.40	1.98	1.72	2.34	5.38	3.34
		FLMO	1.35	1.77	1.58	2.34	5.38	3.34
		LMO	1.31	1.76	1.58	2.95	4.26	3.20
cc-pVQZ		CMO	2.52	13.53	8.26	5.05	17.24	9.43
	n = 0.5	pFLMO	0.33	2.27	1.43	0.61	6.97	1.80
		FLMO	0.31	1.97	1.31	0.61	6.94	1.80
		LMO	0.31	1.98	1.31	0.59	7.04	1.72
	n = 1	pFLMO	0.34	2.15	1.40	0.96	4.26	1.94
		FLMO	0.32	1.97	1.31	0.96	4.25	1.94
		LMO	0.32	1.97	1.31	0.93	4.56	1.89
	n = 2	pFLMO	2.04	1.32	1.65	1.70	3.92	2.05
		FLMO	1.34	1.76	1.58	1.69	3.92	2.05
		LMO	1.58	1.76	1.31	1.80	2.59	2.00

[a] ${}_n\Omega_2$ (eq 3) was used for both pFLMOs and LMOs (cc-pVDZ). σ_{min}, σ_{max}, and σ_{av}: minimal, maximal, and averaged OSs.

Table 3. Number of Effective p−h Pairs (%) with $O_{ai} > \eta$[a]

molecule	basis	N_{ph}	η	CMO	LMO (0.5)	LMO (1)	LMO (2)	FLMO (0.5)	FLMO (1)	FLMO (2)
$C_{12}H_2$	cc-pVDZ	5217	10^{-1}	84.3	26.2	27.0	34.2	27.3	26.9	34.2
			10^{-2}	99.9	63.7	68.4	88.2	65.7	68.9	84.1
			10^{-3}	100.0	91.1	95.8	99.8	92.5	95.8	99.7
	aug-cc-pVDZ	9250	10^{-1}	74.8	23.0	23.6	29.7	24.0	24.1	31.5
			10^{-2}	98.4	66.3	70.3	83.8	68.5	69.8	84.0
			10^{-3}	100.0	93.6	97.5	100.0	95.4	96.8	99.9
	aug-cc-pVTZ	20276	10^{-1}	74.9	21.8	23.0	29.9	23.3	24.0	31.9
			10^{-2}	98.7	65.5	71.6	86.3	69.5	72.4	85.4
			10^{-3}	100.0	92.7	98.6	100.0	97.0	98.5	100.0
	cc-pVQZ	25271	10^{-1}	80.9	19.9	22.7	27.2	20.8	23.7	28.5
			10^{-2}	98.5	56.9	64.8	80.0	60.3	64.6	77.8
			10^{-3}	100.0	85.7	94.8	100.0	90.0	92.6	99.3
C_{60}	cc-pVDZ	118800	10^{-1}	99.7	8.8	9.0	11.8	9.0	9.2	12.0
			10^{-2}	100.0	44.6	52.0	69.9	46.7	52.7	71.2
			10^{-3}	100.0	84.5	88.2	100.0	87.0	88.8	100.0

[a] N_{ph}: total number of pairs. Penalty exponents are in parentheses.

significantly be alleviated through the generic idea of "from fragments to molecule", where the optimization of the localization functional is applied only to small subsystems for generating fragment-centered pFLMOs. A least action can then be taken to obtain the desired orthonormal LMOs. This is apparently a "synthesis" process, which is conceptually and operationally simple and essentially free. Extensions of the idea to relativistic spinors as well as CASSCF orbitals/spinors are straightforward. It is certainly possible to design a freeze-and-thaw algorithm to alleviate the global SCF/CASSCF problem by decomposing the pFLMOs into frozen and active subsets.

dx.doi.org/10.1021/ar500082t | Acc. Chem. Res. 2014, 47, 2758−2767

Work along these directions is being carried out at our laboratory.

AUTHOR INFORMATION

Corresponding Author

*E-mail: liuwjbdf@gmail.com.

Author Contributions

Z.L. and H.L. contributed equally.

Funding

This work was supported by the NSFC (Project Nos. 21033001, 21273011, and 21290192).

Notes

The authors declare no competing financial interest.

Biographies

Z. Li was born in 1987 in Shanxi, China. He obtained his Ph. D. in Chemistry from Peking University (2014). His research focuses on TD-DFT for open-shell systems.

H. Li was born in 1989 in Shandong, China. He obtained his B.Sc. in Chemistry from Sichuan University (2011). His research focuses on localization of relativistic spinors.

B. Suo was born in 1977 in Shaanxi, China. He obtained his Ph. D. in physics from Northwest University (2005). His research focuses on efficient multireference configuration interaction.

W. Liu is Cheung Kong Professor at Peking University. His research focuses on relativistic theories and methods for molecular electronic structure and magnetic properties. He is recipient of the Bessel Research Award of Humboldt Foundation, Annual Medal of International Academy of Quantum Molecular Science, and Pople Medal of Asia-Pacific Association of Theoretical and Computational Chemists. He is an elected member of International Academy of Quantum Molecular Science.

REFERENCES

(1) Boys, S. F. Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Rev. Mod. Phys.* **1960**, *32*, 296−299.
(2) Foster, J. M.; Boys, S. F. Canonical configurational interaction procedure. *Rev. Mod. Phys.* **1960**, *32*, 300−302.
(3) Edmiston, C.; Ruedenberg, K. Localized atomic and molecular orbitals. *Rev. Mod. Phys.* **1963**, *35*, 457−465.
(4) Edmiston, C.; Ruedenberg, K. Localized atomic and molecular orbitals. II. *J. Chem. Phys.* **1965**, *43*, S97−S116.
(5) Magnasco, V.; Perico, A. Uniform localization of atomic and molecular orbitals. I. *J. Chem. Phys.* **1967**, *47*, 971−981.
(6) Von Niessen, W. Density localization of atomic and molecular orbitals. I. *J. Chem. Phys.* **1972**, *56*, 4290−4297.
(7) Kleier, D. A.; Halgren, T. A.; Hall, J. H., Jr; Lipscomb, W. N. Localized molecular orbitals for polyatomic molecules. I. A comparison of the Edmiston-Ruedenberg and Boys localization methods. *J. Chem. Phys.* **1974**, *61*, 3905−3919.
(8) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916−4926.
(9) Lehtola, S.; Jónsson, H. Pipek−Mezey orbital localization using various partial charge estimates. *J. Chem. Theory Comput.* **2014**, *10*, 642−649.
(10) Angeli, C.; Del Re, G.; Persico, M. Quasi-bond orbitals from maximum-localization hybrids for ab initio CI calculations. *Chem. Phys. Lett.* **1995**, *233*, 102−110.
(11) Aquilante, F.; Pedersen, T. B.; de Merás, A. S.; Koch, H. Fast noniterative orbital localization for large molecules. *J. Chem. Phys.* **2006**, *125*, No. 174101.
(12) Zoboko, T.; Mayer, I. Extremely localized nonorthogonal orbitals by the pairing theorem. *J. Comput. Chem.* **2011**, *32*, 689−695.
(13) Reed, A. E.; Weinhold, F. Natural localized molecular orbitals. *J. Chem. Phys.* **1985**, *83*, 1736−1740.
(14) Gu, F. L.; Aoki, Y.; Korchowiec, J.; Imamura, A.; Kirtman, B. A new localization scheme for the elongation method. *J. Chem. Phys.* **2004**, *121*, 10385−10391.
(15) de Silva, P.; Giebułtowski, M.; Korchowiec, J. Fast orbital localization scheme in molecular fragments resolution. *Phys. Chem. Chem. Phys.* **2012**, *14*, 546−552.
(16) Subotnik, J. E.; Dutoi, A. D.; Head-Gordon, M. Fast localized orthonormal virtual orbitals which depend smoothly on nuclear coordinates. *J. Chem. Phys.* **2005**, *123*, No. 114108.
(17) Ziółkowski, M.; Jansík, B.; Jørgensen, P.; Olsen, J. Maximum locality in occupied and virtual orbital spaces using a least-change strategy. *J. Chem. Phys.* **2009**, *131*, No. 124112.
(18) Høyvik, I.-M.; Jansík, B.; Kristensen, K.; Jørgensen, P. Local Hartee-Fock orbitals using a three-level optimization strategy for the energy. *J. Comput. Chem.* **2013**, *34*, 1311−1320.
(19) Jansík, B.; Høst, S.; Kristensen, K.; Jørgensen, P. Local orbitals by minimizing powers of the orbital variance. *J. Chem. Phys.* **2011**, *134*, No. 194104.
(20) Høyvik, I.-M.; Jansík, B.; Jørgensen, P. Orbital localization using fourth central moment minimization. *J. Chem. Phys.* **2012**, *137*, No. 224114.
(21) Wu, F.; Liu, W.; Zhang, Y.; Li, Z. Linear-scaling time-dependent density functional theory based on the idea of 'from fragments to molecule'. *J. Chem. Theory Comput.* **2011**, *7*, 3643−3660.
(22) Rubio, J.; Povill, A.; Malrieu, J. P.; Reinhardt, P. Direct determination of localized Hartree-Fock orbitals as a step toward N scaling procedures. *J. Chem. Phys.* **1997**, *107*, 10044−10050.
(23) Maynau, D.; Evangelisti, S.; Guihéry, N.; Calzado, C. J.; Malrieu, J.-P. Direct generation of local orbitals for multireference treatment and subsequent uses for the calculation of correlation energy. *J. Chem. Phys.* **2002**, *116*, 10060−10068.
(24) Seijo, L.; Barandiarán, Z. Parallel, linear-scaling building-block and embedding method based on localized orbitals and orbital-specific basis sets. *J. Chem. Phys.* **2004**, *121*, 6698−6709.
(25) Subotnik, J. E.; Shao, Y.; Liang, W.; Head-Gordon, M. An efficient method for calculating maxima of homogeneous functions of orthogonal matrices: Applications to localized occupied orbitals. *J. Chem. Phys.* **2004**, *121*, 9220−9229.
(26) Høyvik, I.-M.; Jansik, B.; Jørgensen, P. Trust region minimization of orbital localization functions. *J. Chem. Theory Comput.* **2012**, *8*, 3137−3146.
(27) Lehtola, S.; Jónsson, H. Unitary optimization of localized molecular orbitals. *J. Chem. Theory Comput.* **2013**, *9*, 5365−5372.
(28) Guo, Y.; Li, W.; Li, S. An efficient linear scaling procedure for constructing localized orbitals of large molecules based on the one-particle density matrix. *J. Chem. Phys.* **2011**, *135*, No. 134107.
(29) Liu, S.; Perez-Jorda, J. M.; Yang, W. Nonorthogonal localized molecular orbitals in electronic structure theory. *J. Chem. Phys.* **2000**, *112*, 1634−1644.
(30) Peng, L.; Gu, F.; Yang, W. Effective preconditioning for ab initio ground state energy minimization with non-orthogonal localized molecular orbitals. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15518−15527.
(31) Stoll, H.; Wagenblast, G.; Preuß, H. On the use of local basis sets for localized molecular orbitals. *Theor. Chim. Acta* **1980**, *57*, 169−178.
(32) Gianinetti, E.; Raimondi, M.; Tornaghi, E. Modification of the Roothaan equations to exclude BSSE from molecular interaction calculations. *Int. J. Quantum Chem.* **1996**, *60*, 157−166.
(33) Mo, Y.; Peyerimhoff, S. D. Theoretical analysis of electronic delocalization. *J. Chem. Phys.* **1998**, *109*, 1687−1697.
(34) Reed, A. E.; Weihold, F. Natural bond orbital analysis of near-Hartree-Fock water dimer. *J. Chem. Phys.* **1983**, *78*, 4066−4073.

(35) Foster, J. P.; Weihold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.
(36) Cederbaum, L.; Schirmer, J.; Meyer, H.-D. Block diagonalisation of Hermitian matrices. *J. Phys. A: Math. Gen.* **1989**, *22*, 2427−2439.
(37) Liu, W.; Kutzelnigg, W. Quasirelativistic theory. II. Theory at matrix level. *J. Chem. Phys.* **2007**, *126*, No. 114107.
(38) Høyvik, I.-M.; Kristensen, K.; Kjægaard, T.; Jørgensen, P. A perspective on the localizability of Hartree-Fock orbitals. *Theor. Chem. Acc.* **2014**, *133*, 1417.
(39) Gu, F. L.; Aoki, Y. An elongation method for large systems toward bio-systems. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7640−7668.
(40) Liu, W.; Hong, G.; Dai, D.; Li, L.; Dolg, M. The Beijing four-component density functional program package (BDF) and its application to EuO, EuS, YbO and YbS. *Theor. Chem. Acc.* **1997**, *96*, 75−83.
(41) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(42) Liu, J.; Zhang, Y.; Liu, W. Photoexcitation of light-harvesting C−P−C_{60} triads: A FLMO-TD-DFT study. *J. Chem. Theory Comput.* **2014**, *10*, 2436−2448.