Direct Unconstrained Variable-Metric Localization of One-Electron Orbitals

Ziling Luo* and Rustam Z. Khaliullin*

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Spatially localized one-electron orbitals, orthogonal and non-orthogonal, are widely used in electronic structure theory to describe chemical bonding and speed up calculations. In order to avoid linear dependencies of localized orbitals, the existing localization methods either constrain orbital transformations to be unitary, that is, metric preserving, or, in the case of variable-metric methods, fix the centers of non-orthogonal localized orbitals. Here, we developed a different approach to orbital localization, in which these constraints are replaced with a single restriction that specifies the maximum allowed deviation from the orthogonality for the final set of localized orbitals. This reformulation, which can be viewed as a generalization of existing localization methods, enables one to choose the desired balance between the orthogonality and locality of the orbitals. Furthermore, the approach is conceptually and practically simple as it obviates the necessity in unitary transformations and allows one to determine optimal positions of the centers of non-orthogonal orbitals in an unconstrained and straightforward minimization procedure. It is demonstrated to produce well-localized orthogonal and non-orthogonal orbitals with the Berghold and Pipek−-Mezey localization functions for a variety of molecules and periodic materials including large systems with nontrivial bonding.
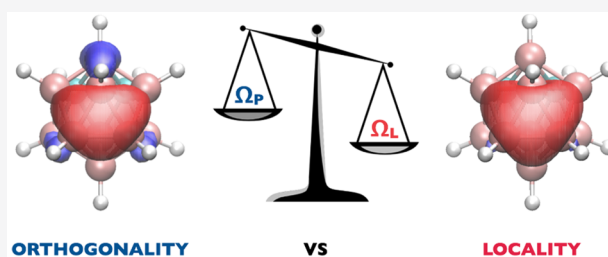
INTRODUCTION

Spatially localized orbitals are of paramount importance in one-electron theories such as the Hartree−Fock method and Kohn−Sham density functional theory as well as in post-Hartree−Fock wave function-based electron-correlation methods. Localized orbitals are widely used to describe and visualize chemical bonding between atoms, thus helping classify bonds and understand electronic-structure origins of observed properties of atomistic systems.[1−5] Furthermore, localized orbitals are the key ingredient in multiple local electronic-structure methods[6−13] that dramatically reduce the computational cost of modeling electronic properties of large systems.[14−18] Spatially localized orbitals are known as localized molecular orbitals (LMOs) in the field of molecular quantum chemistry and maximally localized Wannier functions (MLWFs) in solid state physics and materials science.[19] Here, they will be collectively referred to as LMOs, whereas the eigenstates of the effective one-electron Hamiltonian will be called canonical molecular orbitals (CMOs) regardless of whether the system is isolated or treated with periodic boundary conditions.

In traditional localization methods, LMOs are constructed by finding a unitary transformation of CMOs that minimizes a localization function that effectively measures the spread of individual orbitals. Since CMOs are orthogonal and a unitary transformation preserves the orbital metric, LMOs obtained in this way are orthogonal (OLMOs) by construction.[20] Multiple localization functions have been proposed for molecular systems

including Boys−Foster,[1] Edmiston−Ruedenberg,[2,21,22] Pipek−Mezey,[3] and von Niessen.[4] The Boys−Foster localization[1] is perhaps the most popular because of the simplicity of its physical interpretation, low computational complexity, and ease of implementation. The Pipek−Mezey localization,[3] which maximizes atomic charges[23−25] of each orbital, is also widely used because it does not mix LMOs representing σ and π bonds and thus gives a clear picture of bonding patterns. In the past decade, these well-established localization functions have been modified to reduce orbital tails and produce more uniform localization across the set of LMOs.[25−30] For large supercells of condensed-phase periodic systems, where electronic structure can be described with the Γ-point sampling of the Brillouin zone, several efficient optimization methods have been proposed to construct MLWF.[19,31−35] While the methods listed above rely on different localization criteria, they have been formulated to ensure that orbitals remain orthogonal in a localization procedure. This is typically achieved through the exponential[34]

or Cayley parametrization of unitary transformations or through simple Jacobi rotations.[2]

Due to the imposed orthogonality condition, orthogonal LMOs exhibit small nonzero values even far away from the localization center. These orthogonalization tails can complicate the interpretation of chemically relevant electronic-structure information and make its transferability from one system to another more difficult. More importantly, the tails can reduce orbital locality making orbital-based local correlation methods less computationally efficient. To mitigate the undesirable orthogonality effects, metric-preserving unitary transformation has been applied to non-orthogonal orbitals.[36] Furthermore, it has been proposed to replace unitary transformation with more general variable-metric nonsingular transformations. This generalization lifts the orthogonality constraint in the localization procedure and increases the number of degrees of freedom available to LMOs.[37−43] It has been found that non-orthogonal localized molecular orbitals (NLMOs) are indeed about 10−30% more localized than OLMOs if measured by the value of the Boys−Foster function.[42,44]

Substantial recent efforts have been made to develop reliable algorithms to construct NLMOs.[36,42,44,45] Despite noticeable progress the existing methods produce NLMOs that are either still fairly similar to OLMOs[46] or lead to the linear dependence between the orbitals.[42] To overcome the linear dependence problem, Yang and co-workers[42,43] have developed a localization method in which the centers of NLMOs are fixed during the minimization of the localization function. The positions of the centers can be taken from the corresponding OLMOs[42] or simply guessed based on the knowledge of bonding patterns in a system.[43] While this method solves severe linear dependence issues, it requires either the computational effort to find the OLMOs centroids or good understanding of bonding properties in advance, which may limit the application of the method to relatively simple systems with unambiguous Lewis structures.

In this work, we propose a variable-metric method to simplify the construction and improve the locality of LMOs. The key new component of the method is a simple penalty function that prevents LMOs from becoming linearly dependent and allows one to choose the desirable balance between the non-orthogonality and locality of LMOs. The penalty function replaces all constraints that are normally imposed on localized orbitals with a single restriction that specifies the allowed deviation from the orthogonality for the final localized orbitals. For OLMOs, the method is advantageous because it optimizes orbital mixing coefficients directly, obviating complicated parametrization of unitary transformations and simplifying orbital optimization algorithms. For NLMOs, the new approach allows one to determine the optimal positions of their centers automatically in an unconstrained and straightforward optimization procedure, without *a priori* knowledge of bonding patterns in the system.

■ METHODOLOGY

Theory. The localization procedure starts with a set of occupied one-electron states $|i_0\rangle$. These orbitals are not assumed to be canonical or even orthogonal. However, they are assumed to be normalized, which does not reduce the generality of the method. Furthermore, the initial orbitals must be linearly independent; that is, their overlap matrix $\sigma_{ji}^0 \equiv \langle j_0|i_0\rangle$ must be invertible. The trial NLMOs are expressed as a linear combination of these initial states

$$|j\rangle = |i_0\rangle A^i{}_j \tag{1}$$

The conventional tensor notation is used to work with the non-orthogonal orbitals:[47] covariant quantities are denoted with subscripts, contravariant quantities with superscripts, and summation is implied over the same covariant and contravariant orbital indices. Summation is not implied if two indices are both covariant.

The loss function minimized in this work contains two terms: a conventional localization function Ω_L and a term that penalizes unphysical states with linearly dependent occupied orbitals Ω_P

$$\Omega(\mathbf{A}) = \Omega_L(\mathbf{A}) + c_P\Omega_P(\mathbf{A}),$$
$$\Omega_P(\mathbf{A}) = - \ln \det[\sigma(\mathbf{A})] \tag{2}$$

where $c_P > 0$ is the penalty strength and σ is the NLMO overlap matrix

$$\sigma_{kl} = \langle k|l\rangle = A^j{}_k\sigma_{ji}^0 A^i{}_l \tag{3}$$

If the NLMOs are normalized, the determinant of σ varies from 1 for orthogonal NLMOs to 0 for linearly dependent NLMOs. The penalty function—the key ingredient of the proposed method—varies from 0 to $+\infty$ for these two extreme cases, making the linearly dependent state inaccessible in the localization procedure with finite penalty strength c_P. A normalization constraint can be imposed on NLMOs if their coefficients are expressed in terms of independent optimization parameters denoted with lowercase $\mathbf{a}$

$$A^i{}_j = a^i{}_j(a^k{}_j\sigma_{kl}^0 a^l{}_j)^{-1/2} \equiv a^i{}_j N_j \tag{4}$$

where the $\mathbf{a}$-dependent normalization coefficient N_j is defined for brevity. It should be noted that the normalization requirement does not reduce the flexibility of the localization procedure and can be removed altogether if the penalty function is replaced with its more general version

$$\Omega_P^G(\mathbf{A}) = \Omega_P(\mathbf{A}) + \ln \det[\text{diag}(\sigma(\mathbf{A}))] \tag{5}$$

with $\mathbf{A}$ now being independent optimization parameters. It is worth mentioning that the determinant of the NLMO overlap is equal to the square of the volume of the multidimensional parallelepiped spanned by the NLMO vectors, defined by matrix $\mathbf{A}$.

The inclusion of the penalty term converts the localization procedure into an unconstrained, straightforward optimization problem. Additionally, adjusting the strength of the penalty c_P enables one to achieve the right balance between the non-orthogonality and the locality of the orbitals (see below).

Penalty Strength. If the penalty strength c_P is extremely large, Ω_L is negligible compared to the penalty term and the minimization of Ω is numerically equivalent to a trivial orbital orthogonalization. In the opposite case of extremely small c_P, the minimization of Ω may result in linear dependence between NLMOs reported earlier.[43]

As we show below, there is a wide range of c_P values between the two extremes that produce NLMOs that are substantially more localized than OLMOs and linearly independent. A simple strategy to find an appropriate penalty strength is to minimize Ω with a sufficiently large initial c_P value and then gradually decrease c_P until the determinant of the overlap of the *optimal* NLMOs drops below a desired threshold $D_{tar} \in (0, 1]$. The initial value of c_P should be chosen to balance approximately the localization and penalty components of Ω. Thus, a reasonable

initial value of c_P can be estimated by assuming that the reduction in Ω_L upon the minimization of Ω has the same order of magnitude as the change in the penalty $c_P \Omega_P$:

$$c_P^{\text{init}} \sim \frac{\Omega_L(\mathbf{I}) - \Omega_L(\mathbf{A}^*)}{\Omega_P(\mathbf{A}^*) - \Omega_P(\mathbf{I})} = \frac{\beta_L^*}{\beta_P^*}\Omega_L(\mathbf{I}) \tag{6}$$

where $\mathbf{A}^*$ denotes the (yet unknown) solution to the minimization problem,

$$\beta_L^* \equiv \frac{\Omega_L(\mathbf{I}) - \Omega_L(\mathbf{A}^*)}{\Omega_L(\mathbf{I})} \in [0, 1] \tag{7}$$

is the (positive) expected relative reduction in the localization function, and

$$\beta_P^* \equiv \ln \frac{\det \sigma(\mathbf{I})}{\det \sigma(\mathbf{A}^*)} \approx \ln \frac{\det \sigma(\mathbf{I})}{D_{\text{tar}}} \equiv \beta_P > 0 \tag{8}$$

is the logarithm of the ratio of the initial and final determinants. The importance of eqs 6−8 is that they allow one to estimate the initial value of c_P as a product of $\Omega_L(\mathbf{I})$, which can be easily calculated in the beginning of the optimization procedure, and a dimensionless constant α

$$c_P^{\text{init}} = \frac{\beta_L^*}{\beta_P}\Omega_L(\mathbf{I}) \equiv \alpha\Omega_L(\mathbf{I}) \tag{9}$$

Equation 9 makes clear that the penalty component is an extensive function of a system with the units that are consistent with the localization component. Although the optimal dimensionless parameter β_L^* is not known *a priori*, its magnitude can be easily estimated to obtain a sufficiently large initial guess for c_P. For example, an optimization of canonical orbitals det $\sigma(\mathbf{I}) = 1$ that should produce the NLMO overlap determinant $D_{\text{tar}} \approx 0.1$ can be initialized by adopting the maximum possible value of $\beta_L^* = 1$, giving $\alpha = \ln^{-1} 10$.

The procedure for tuning c_P is shown as the outer loop of the optimization algorithm in Figure 1. Its only required input is D_{tar}. It is worth mentioning that the algorithm can be modified to treat the penalty strength as a Lagrange multiplier that imposes the det $\sigma(\mathbf{A}^*) = D_{\text{tar}}$ constraint rigorously.

Implementation. In this work, we adopted the localization function proposed by Resta[31,32] and generalized by Berghold *et al.*[34] to three dimensions and simulation cells of general shape and symmetry:

$$\Omega_L(\mathbf{A}) = -\sum_K \sum_i \omega_K |z_i^K|^2,$$
$$z_i^K = A^m{}_i B_{mn}^K A^n{}_i,$$
$$B_{mn}^K = \langle m_0 | e^{i\mathbf{G}_K \cdot \hat{\mathbf{r}}} | n_0 \rangle \tag{10}$$

where $\hat{\mathbf{r}}$ is the position operator in three dimensions and ω_K and $\mathbf{G}_K$ are suitable sets of weights and reciprocal lattice vectors, respectively, labeled by index K.[33,34] We chose to write the summation over K explicitly because K is not an orbital index. The function in eq 10 can be used for both gas-phase and periodic systems.[34] In the former case, the function is equivalent to the Boys−Foster localization.[32,34] In the latter case, its applicability is limited to the electronic states described within the Γ-point approximation.

We also considered the Pipek−Mezey localization function[3,25] that has the advantage of preserving the separation of σ and π bonds and is commonly employed for molecular system

---

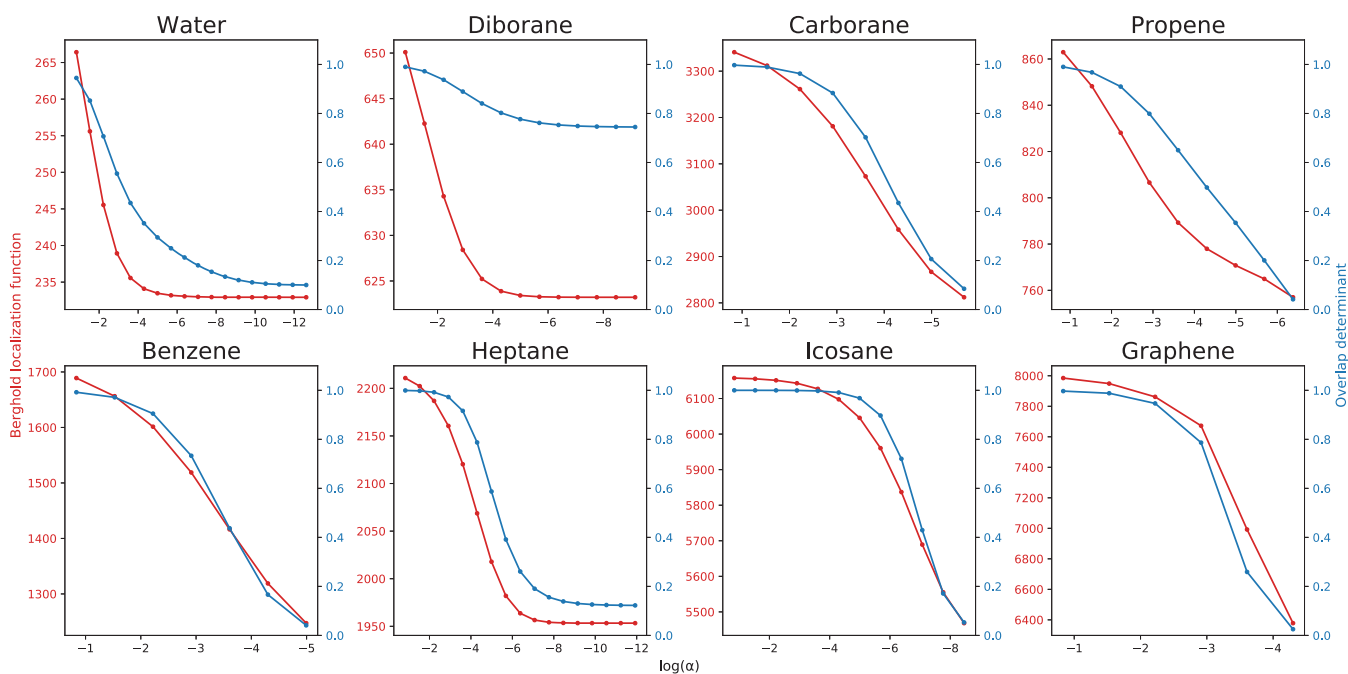**Algorithm 1** Conjugate gradient minimization of Ω

1: Input ϵ_{CG} ▷ Localization convergence threshold
2: Input D_{tar} ▷ Minimum allowed NLMO determinant
3: Input $\mathbf{T}_0$ ▷ Initial basis set coefficients for $|i_0\rangle$
4: Input $\mathbf{S}$ ▷ Basis set overlap
5: Input $\mathbf{L}^K$ ▷ Basis set representation of the localization operator
6: $\sigma_0 \leftarrow \mathbf{T}_0^\dagger \mathbf{S} \mathbf{T}_0$ ▷ Initial orbital overlap
7: $\mathbf{B}^K \leftarrow \mathbf{T}_0^\dagger \mathbf{L}^K \mathbf{T}_0$ ▷ Initial localization matrix, Eq. (10)
8: $\mathbf{a} \leftarrow \mathbf{I}$ ▷ Initial guess on variational parameters
9: StopOuter $\leftarrow$ False ▷ Flag to exit the outer loop
10: $i_{\text{Outer}} \leftarrow 0$ ▷ Iteration counter
11: **repeat** ▷ Loop to change the penalty strength
12: $i_{\text{Outer}} \leftarrow i_{\text{Outer}} + 1$
13: StopCG $\leftarrow$ False ▷ Flag to exit the CG loop
14: $i_{\text{CG}} \leftarrow 0$ ▷ Iteration counter
15: **repeat** ▷ Fixed-penalty localization loop
16:   $i_{\text{CG}} \leftarrow i_{\text{CG}} + 1$
17:   $\mathbf{A} \leftarrow \mathbf{a} \left[\text{diagonal}(\mathbf{a}^\dagger \sigma_0 \mathbf{a})\right]^{-\frac{1}{2}}$ ▷ Update NLMOs
18:   $\sigma \leftarrow \mathbf{A}^\dagger \sigma_0 \mathbf{A}$ ▷ Update overlap
19:   $\text{Det} \leftarrow \text{determinant}(\sigma)$ ▷ Determinant
20:   $\Omega_P \leftarrow -\ln[\text{Det}]$ ▷ Orthogonalization function
21:   $\mathbf{P} \leftarrow$ Eq (15) and (13) ▷ Orthogonalization gradient
22:   $\Omega_L \leftarrow$ Eq (10) ▷ Localization function
23:   $\mathbf{L} \leftarrow$ Eq (14) and (13) ▷ Localization gradient
24:   **if** $i_{\text{Outer}} = 1$ **and** $i_{\text{CG}} = 1$ **then**
25:    $c_P \leftarrow \Omega_L(\ln[\text{Det}/D_{\text{tar}}])^{-1}$ ▷ Initial strength
26:   **end if**
27:   $\Omega \leftarrow \Omega_L + c_P \Omega_P$
28:   **if** $i_{\text{CG}} > 1$ **then**
29:    $\boldsymbol{\Gamma} \leftarrow \mathbf{G}$ ▷ Save old gradient
30:   **end if**
31:   $\mathbf{G} \leftarrow \mathbf{L} + c_P \mathbf{P}$
32:   **if** $\|\mathbf{G}\|_{\text{max}} < \epsilon_{\text{CG}}$ **then**
33:    StopCG $\leftarrow$ True
34:   **end if**
35:   **if not** StopCG **then**
36:    **if** $i_{\text{CG}} > 1$ **then**
37:     $\mathbf{O} \leftarrow \mathbf{D}$ ▷ Save old direction
38:    **end if**
39:    $\mathbf{D} \leftarrow -\mathbf{G}$ ▷ Initial direction
40:    **if** $i_{\text{CG}} > 1$ **then**
41:     $\beta \leftarrow \text{Tr}(\mathbf{G}^\dagger \mathbf{D})/\text{Tr}(\boldsymbol{\Gamma}^\dagger \mathbf{O})$
42:     $\mathbf{D} \leftarrow \mathbf{D} + \beta\mathbf{O}$ ▷ Search direction
43:    **end if**
44:    $\alpha \leftarrow \text{argmin}_\alpha \Omega(\mathbf{a} + \alpha\mathbf{D})$ ▷ Line search
45:    $\mathbf{a} \leftarrow \mathbf{a} + \alpha\mathbf{D}$ ▷ Update variational DOFs
46:   **end if**
47: **until** StopCG
48: **if** $\text{Det} < D_{\text{tar}}$ **or** $i_{\text{Outer}} > N_{\text{Outer}}^{\text{Max}}$ **then**
49:   StopOuter $\leftarrow$ True
50: **end if**
51: **if** $i_{\text{Outer}} > 1$ **then**
52:   $c_P \leftarrow c_P \cdot r$ ▷ Reduce c_P by factor $r \in (0, 1)$
53: **end if**
54: **until** StopOuter
55: **return** $\mathbf{T} \leftarrow \mathbf{T}_0 \mathbf{A}$ ▷ Return NLMOs coefficients

Figure 1. Algorithm for the optimization of NLMOs.

$$\Omega_L^{\text{PM}}(\mathbf{A}) = -\sum_{K=1}^{\text{Atoms}} \sum_i |z_i^K|^2,$$
$$z_i^K = A^m{}_i B_{mn}^K A^n{}_i,$$
$$B_{mn}^K = \frac{1}{2} \sum_{\mu \in K} \langle m_0 | (|\chi_\mu\rangle\langle\chi^\mu| + |\chi^\mu\rangle\langle\chi_\mu|) | n_0 \rangle \tag{11}$$

Figure 2. Dependence of the optimal localization function and determinant of the NLMO overlap on α—the adjustable part of the penalty strength. The first point on the left is $\alpha = \ln^{-1} 10 \approx 0.434$.

where z_i^K is the (real) contribution of orbital i to the Mulliken charge of atom K and $|\chi_\mu\rangle$ and $|\chi^\mu\rangle$ are atom-centered covariant and contravariant basis set functions.[33,34] The summation over μ is written explicitly to emphasize that it is restricted to the basis functions centered on atom K.

The unconstrained minimization of function Ω with fixed c_P can be carried out with a variety of algorithms. In this work, we used a simple conjugate gradient algorithm summarized in Figure 1. The gradient $G_i^{\ j} \equiv \frac{\partial \Omega}{\partial a^i_{\ j}}$ required in the algorithm is a sum of the localization $L_i^{\ j} \equiv \frac{\partial \Omega_L}{\partial a^i_{\ j}}$ and penalty $P_i^{\ j} \equiv \frac{\partial \Omega_P}{\partial a^i_{\ j}}$ components:

$$G_k^{\ l} = L_k^{\ l} + c_P P_k^{\ l} \tag{12}$$

These components can be readily expressed in terms of the derivatives with respect to the transformation coefficients $\tilde{X}_k^{\ l} \equiv \frac{\partial \Omega_X}{\partial A^k_{\ l}}$, where X is either L or P:

$$X_i^{\ j} = \tilde{X}_k^{\ l} \frac{\partial A^k_{\ l}}{\partial a^i_{\ j}} = [\tilde{X}_i^{\ j} - (\sigma_{in}^0 A^n_{\ j})(A^m_{\ j} \tilde{X}_m^{\ j})] N_j \tag{13}$$

$$\tilde{L}_k^{\ l} = -\sum_K 4\omega_K [\mathrm{Re}(B_{kn}^K) A^n_{\ l} \, \mathrm{Re}(z_l^K) + \mathrm{Im}(B_{kn}^K) A^n_{\ l} \, \mathrm{Im}(z_l^K)] \tag{14}$$

$$\tilde{P}_k^{\ l} = -2\sigma_{km}^0 A^m_{\ n} \sigma^{nl} \tag{15}$$

If the gradient in eq 14 is computed for the Pipek−Mezey localization function, weights ω_K are equal to one and the imaginary part of z_i^K is equal to zero.

The logarithm of the determinant of the symmetric positive definite overlap matrix σ is computed using the trace of the matrix logarithm *via* Mercator series:

$$\ln \det(\sigma) = \ln[\det(\mathbf{s}) \det(\mathbf{I} + \mathbf{X}) \det(\mathbf{s})],$$

$$\ln \det(\mathbf{I} + \mathbf{X}) = \mathrm{Tr} \ln[\mathbf{I} + \mathbf{X}],$$

$$\ln[\mathbf{I} + \mathbf{X}] = \mathbf{X} - \frac{1}{2}\mathbf{X}^2 + \frac{1}{3}\mathbf{X}^3 - \frac{1}{4}\mathbf{X}^4 + \dots \tag{16}$$

where $\mathbf{s} = [\mathrm{diag}(\sigma)]^{1/2}$. The series converges only if $\|\mathbf{X}\|_F < 1$. If it is not the case, the square root of $\mathbf{I} + \mathbf{X}$ is computed recursively until the norm requirement is satisfied. The advantage of this algorithm is that it relies exclusively on matrix−matrix multiplication and can be readily implemented in any matrix library. Its computational cost grows cubically with the number of orbitals for dense matrices and linearly for sparse overlap matrices of localized orbitals.

The localization procedure was implemented in the CP2K software package.[48] The DBCSR library[49] that handles sparse matrices efficiently on massively parallel computing platforms is utilized to enable orbital localization in large systems.

Computational Details. To test the localization procedure, we constructed NLMOs for several systems ranging from a simple water molecule to complex molecules with nontrivial bonding patterns and to large periodic systems. For all systems, the initial CMOs were obtained using the conventional diagonalization-based SCF procedure implemented in the electronic-structure module of CP2K. The Becke−Lee−Yang−Parr generalized gradient approximation[50,51] was used as the exchange−correlation functional. Goedecker−Teter−Hutter pseudopotentials[52] were used together with a triple-ζ atom-centered Gaussian basis set with two sets of polarization functions for all atoms. The energy cutoff was set at 600 Ry to define the auxiliary plane-wave basis set in the construction of the effective Hamiltonian. The integration over the Brillouin zone was performed using the Γ-point approximation.

■ RESULTS AND DDISCUSSION

Compromise between Locality and Orthogonality. For all test systems, the conjugate gradient localization procedure in

Table I. Relative Reduction in the Localization Function and the Final Determinant of the NLMO Overlap

molecules	$\Delta_{OLMOs/CMOs}$	$\Delta_{NLMOs/CMOs}$	$\Delta_{NLMOs/OLMOs}$	$\det(\sigma)$
H_2O	22	36	18	0.100
CO_2	65	76	30	0.025
diborane (B_2H_6)	62	64	6.2	0.745
borazine ($B_3N_3H_6$)	73	78	20	0.026
carborane ($C_2B_{10}H_{12}$)	72	76	17	0.085
propene (C_3H_6)	61	67	14	0.042
1-butyne (C_4H_6)	62	70	19	0.063
benzene (C_6H_6)	69	78	28	0.041
heptane (C_7H_{16})	89	90	12	0.122
icosane ($C_{20}H_{42}$)	98	98	11	0.053
decacyclene ($C_{72}H_{24}$)	94	95	16	0.042
graphene	77	82	21	0.025
average	70	76	18	0.114

Figure 1 is stable and efficient. The numerical precision of the implemented code allows one to treat NLMOs with $\det(\sigma)$ as low as 10^{-8} as distinct. However, a visual inspection of NLMOs with such a tiny overlap determinant reveals that many orbitals become almost identical (for example, localized on the same bonds) with only minor physically irrelevant differences. At the same time, NLMOs with $\det(\sigma) > 10^{-1}$ are found to highlight bonding patterns in all test systems correctly. Therefore, we set the minimum allowed NLMO determinant D_{tar} to 10^{-1} in all tests. The initial value of $\alpha = \ln^{-1} 10$ was set according to eq 9. The value of α was decreased by a factor of 2 in the outer loop of the algorithm (Figure 1) until the overlap determinant fell below D_{tar} or until the optimal Ω_L stopped changing with c_P appreciably (see water, diborane, and heptane examples in Figure 2).

Since the determinant of the overlap is proportional to the square of the volume of the parallelepiped spanned by the NLMO vectors, $D_{tar} = 10^{-1}$ corresponds to the volume of $\det(\mathbf{A}) \approx 0.32$, which is large enough to produce distinct NLMOs.

Figure 2 demonstrates how the penalty strength affects the optimal orbital localization and the determinant of the orbital overlap. In all tests, the initial penalty strength is sufficiently large to produce almost perfectly orthogonal localized orbitals. At the same time it is not too large to yield more localized non-orthogonal localized orbitals after just several steps of c_P adjustment.

Figure 2 shows that within a large range of values spanning 3–6 orders of magnitude c_P serves as an adjustable parameter that can be tuned to achieve a desirable locality–orthogonality compromise. Thus, the flexibility of the unconstrained local-ization method presented here allows one to combine the strengths of the existing localization methods that produce either orthogonal orbitals or NLMOs with fixed localization centers. It is also important to emphasize that the localization procedure is unconstrained, does not require complicated parametrization of unitary matrices, and relies on a simple easy-to-implement conjugate gradient optimization algorithm.

NLMOs More Localized than OLMOs. Figure 2 reveals the expected trend: the orbitals become more localized as they are allowed to be less orthogonal. The relative reduction in the localization is quantified by $\Delta_{X/Y} \equiv \frac{\Omega_L(Y) - \Omega_L(X)}{\Omega_L(Y)} \times 100\%$, where X and Y can refer to CMOs, OLMOs (obtained with the conventional unitary transformations), or NLMOs. Table I compares the relative reduction as measured by the Berghold function for OLMO/CMO, NLMO/CMO, and NLMO/

OLMO pairs. Although NLMOs are constructed with D_{tar} set to 10^{-1}, most final values of the NLMO overlap determinants (last column in Table I) are somewhat lower than D_{tar} because the last outer-loop iteration brings $\det(\sigma)$ below D_{tar}.

The relative reduction in localization between OLMOs and CMOs ranges from 22% to 98% and reaches high values for extended systems (e.g., icosane) where the localization has the ability to reduce the spread significantly. The average relative reduction for the data set considered here is 70%. The NLMOs are even more localized than OLMOs. The relative reduction in localization between NLMOs and OLMOs ranges from 6% to 30% with the data set average of 18%. This additional reduction in the orbital spread can lead to a substantial reduction in the number of significant excitation amplitudes in local electron-correlation methods and to noticeable computational savings.

The relative reduction in localization between OLMOs and NLMOs is similar to that obtained with the algorithms that fix NLMO localization centers.[42,43] Since the NLMOs centers were previously fixed at the position of OLMOs centers, this implies that the locations of the centers of NLMOs and OLMOs are very similar. It also suggests that the NLMOs obtained with $D_{tar} = 10^{-1}$ represent the electronic structure of molecules reliably and also serves as an additional verification of the previously employed fixed-center procedures.
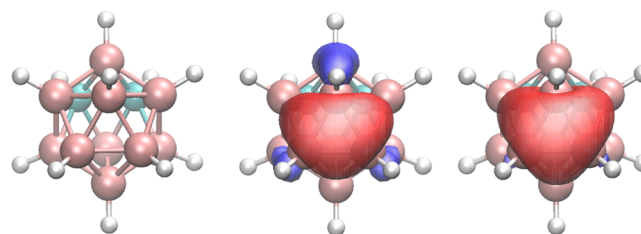

Figure 3. Orthogonal (middle) and non-orthogonal (right) LMOs of three-center–two-electron B–B–B bond in the carborane molecule $C_2B_{10}H_{12}$. The isosurface value is 0.04 au.

Figures 3, 4, 5, and 6 compare the shapes of NLMOs and OLMOs for several representative isolated molecules and periodic systems. Figure 3 shows that the NLMOs and OLMOs of a carborane $C_2B_{10}H_{12}$ molecule representing a three-center–two-electron B–B–B bond have similar positions of their centroids and similar overall shape. The main lobes (red region) of NLMOs are larger than those of OLMOs, whereas
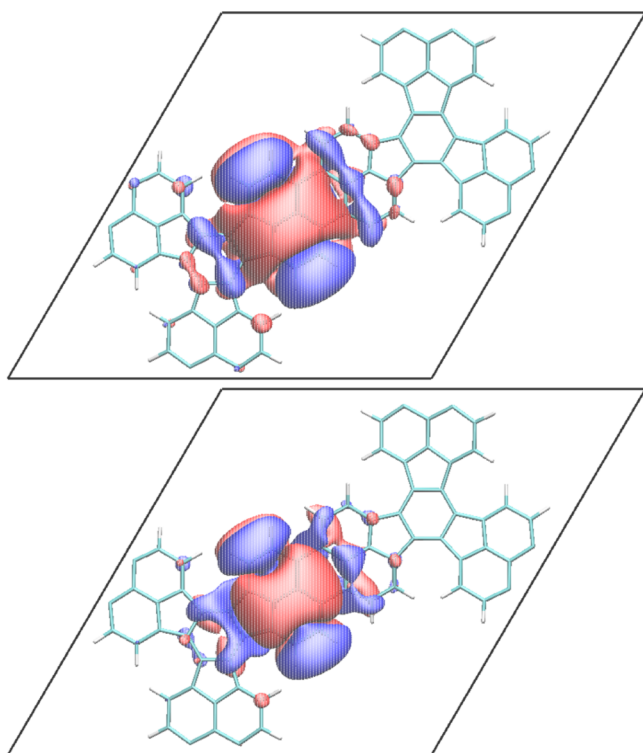
Figure 4. OLMO (top) and NLMO (bottom) representation of a π bond in the π-conjugated 2D polymer decacyclene $C_{72}H_{24}$. The isosurface value is set at a relatively low value of 0.002 au to emphasize the tails of the orbitals.
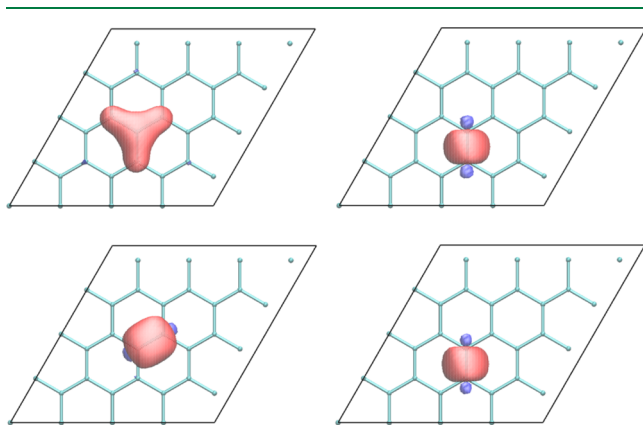
Figure 5. Representative OLMOs (top) and NLMOs (bottom) of graphene. The isosurface value is 0.06 au.

NLMO peripheral tails are smaller. This redistribution of the probability density amplitude toward the center of NLMO is what makes NLMOs more localized than OLMOs—the effect noted previously.[44]

The reduced tails of NLMOs are also visible in Figure 4, which shows LMO representation of a π bond in the extended π-conjugated 2D polymer decacyclene $C_{72}H_{24}$. Without strict orthogonality constraints, these tails can be reduced even further by imposing higher order (e.g., quartic) penalty on amplitudes far away from the orbital center.[27]

σ−π Mixing. Figure 5 compares typical NLMOs and OLMOs of graphene. The single C−C σ bonds are well-reproduced by OLMOs and NLMOs, but both types of LMOs fail to represent the double bonds adequately. OLMOs tend to delocalize over several bonds, complicating analysis of the

electron pairs. Although NLMOs are more localized and extend only over two carbon atoms, they tend to take the shape of a mixture between σ and π bonds producing the so-called τ orbitals. To prevent the σ−π mixing, the Berghold localization function was replaced with the Pipek−Mezey function.

The OLMOs and NLMOs obtained with the Pipek−Mezey and Berghold localization schemes were compared for benzene and allyl alcohol (Figure 6). For OLMOs, the Pipek−Mezey scheme preserves the separation between the σ and π bonds.[3] Upon a visual inspection of NLMOs of these systems, Pipek−Mezey orbitals appear to be more localized than those obtained with the Berghold (i.e., Boys−Foster) scheme. A quantitative comparison of the locality of Pipek−Mezey and Berghold LMOs is difficult because there is no universal function that measures orbital locality. If the Pipek−Mezey function is chosen as a common measure to compare the locality of LMOs, then Pipek−Mezey NLMOs appear more local because they minimize the common measure by definition (Figure 7, lower panel). If the common measure is chosen to be the Berghold spread, then Berghold NLMOs appear more local for the same reason (Figure 7, upper panel).

It is important to note that as the determinant of the NLMO overlap decreases the σ−π separation of the Pipek−Mezey scheme is not maintained and a pair of σ and π bonds tend to mix generating a pair of τ and τ' NLMOs (Figure 6). It has been shown that, in the original Pipek−Mezey localization scheme designed for OLMOs, τ and τ' orbitals cannot be more localized than σ and π orbitals[3]

$$\Omega_L^{PM}(\sigma) + \Omega_L^{PM}(\pi) \leq \Omega_L^{PM}(\tau) + \Omega_L^{PM}(\tau') \tag{17}$$

Here we used a simple two-orbital system as an example to demonstrate that, in the case of NLMOs, the Pipek−Mezey localization function can generate τ and τ' as the most localized solution. We consider two atoms A and B, each with p_x and p_z atomic orbitals, separated along the x axis as shown in Figure 8A. The canonical bonding $\sigma(\pi)$ orbital is represented by the positive linear combination of two $p_x(p_z)$ atomic orbitals. Mixing the occupied σ and π orbitals produces τ and τ' orbitals:

$$|\tau\rangle = |\sigma\rangle \cos(\Theta) + |\pi\rangle \sin(\Theta)$$
$$|\tau'\rangle = -|\sigma\rangle \sin(\Phi) + |\pi\rangle \cos(\Phi)$$
$$\tag{18}$$

In the general case of NLMOs, there are two mixing angles Θ and Φ that can be varied independently to minimize Ω_L^{PM}. OLMOs are recovered when $\Theta = \Phi$. Figure 8B shows the dependence of the Pipek−Mezey localization function on Θ and Φ. The red diagonal line corresponds to all possible OLMOs. The metric-preserving yellow line describes the NLMOs with $\det(\sigma) = 10^{-1}$. As shown in the work of Pipek and Mezey,[3] the σ and π orbitals ($\Theta = \Phi = 0$) minimize the localization function along the diagonal line. But as the orbitals are allowed to be less orthogonal, the most localized solution $\Theta + \Phi = \pi/2$ corresponds to τ and τ'. This simple example shows that, generally speaking, σ−π separation cannot be expected for NLMOs even if they are constructed with the Pipek−Mezey method.

■ CONCLUSIONS

In this work, we proposed a new approach to construct localized orthogonal and non-orthogonal one-electron orbitals. In this approach, the catastrophic linear dependence of the orbitals is prevented by augmenting the localization function with a simple
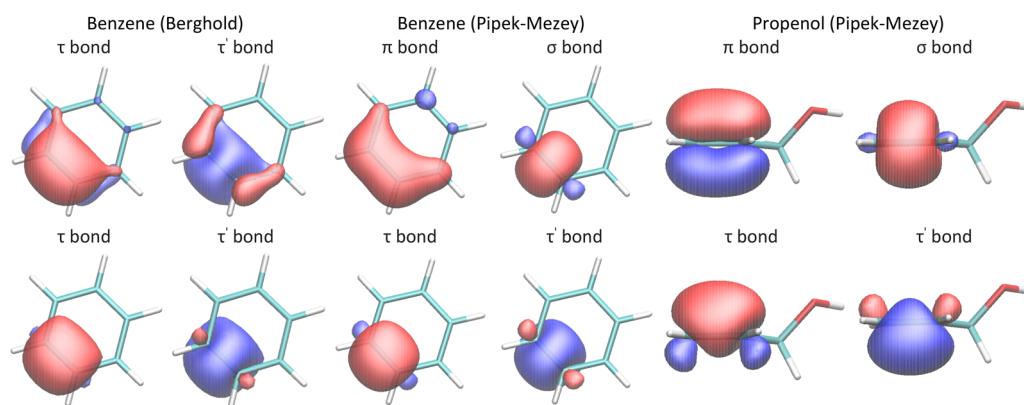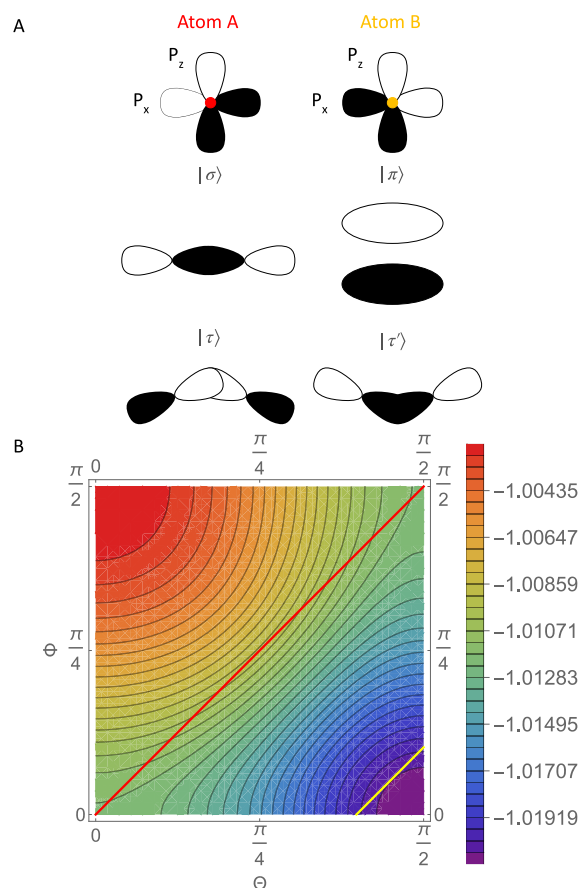
Figure 6. Comparison of OLMOs (top) and NLMOs (bottom) computed with the Berghold and Pipek−Mezey localization functions for benzene and allyl alcohol. The isosurface value is 0.05 au.
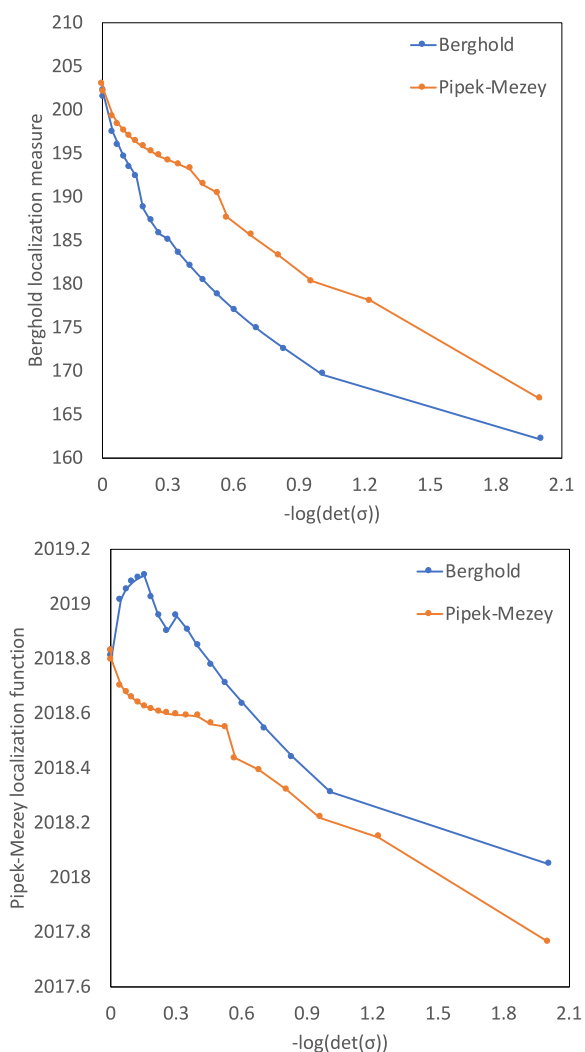
Figure 7. Locality of LMOs of graphene measured by the Berghold spread (upper panel) and Pipek−Mezey localization function (lower panel). LMOs are obtained by minimizing the Berghold (blue) and Pipek−Mezey (orange) localization functions.

Figure 8. (A) σ and π separation and mixed orbitals for a toy system of two atoms (A and B), each with p_x and p_z atomic orbitals. (B) Contours of the localization function with different values of Θ and Φ: red line, $\det(\sigma) = 1$; yellow line, $\det(\sigma) = 10^{-1}$.

single-value penalty that measures the degree of orbital non-orthogonality and thus allows one to avoid nearly dependent states in the localization procedure. The proposed penalty function enables a more flexible approach to orbital localization as it allows one to replace traditional metric-preserving unitary

transformations of the orbitals with more general variable-metric nonsingular transformations. The approach is also conceptually simpler because a complicated parametrization of unitary transformations[36] is obviated, even in the case of OLMOs, allowing one to optimize orbital mixing coefficients directly using simple unconstrained minimization algorithms.

The new approach is easy to implement as was demonstrated for the Berghold[34] and Pipek−Mezey[3] localization functions. Numerous tests were performed for gas-phase and periodic systems. For gas-phase molecules, the Berghold scheme is

equivalent to the Boys—Foster localization. For periodic systems, the electronic structure of which can be described in the Γ-point approximation, it produces orthogonal and non-orthogonal MLWFs.

The tests show that orbital localization is robust when implemented with a simple conjugate gradient algorithm. The procedure generates NLMOs without *a priori* knowledge of bonding patterns in the system. An additional black-box algorithm is proposed to tune the penalty strength and produce the desired balance between the orthogonality and locality of NLMOs. The desired balance is specified as the minimum allowed determinant of the NLMO overlap matrix—an intuitively clear parameter in the $(0,1]$ range. We found that NLMOs with the minimum allowed determinant of the order of 10^{-1} correctly recover bonding patterns in a variety of molecules and materials, including large systems with nontrivial bonding. NLMOs are approximately 18% more localized than OLMOs and have reduced tails. This observation is consistent with previous results obtained for the NLMOs with the localization centers fixed at the OLMO positions.[42,43] It serves as an additional verification of the previously employed procedures and implies that the NLMOs representation of the electronic structure of molecules is reliable.

To our surprise, the Pipek—Mezey localization, which has been a reliable method to separate orthogonal σ and π orbitals in locally planar systems with double bonds, tends to mix these orbitals when they are allowed to become strongly non-orthogonal. The origins of this effect was explained using a simple two-orbital system as an example.

The broader significance of this work is in the new continuous easy-to-compute non-orthogonality measure that can be adopted by a variety of electronic-structure theories that benefit from non-orthogonal representation of wave functions.

■ AUTHOR INFORMATION

Corresponding Authors

Ziling Luo — *Department of Chemistry, McGill University, Montreal, Quebec QC H3A 0B8, Canada*; Email: ziling.luo@mail.mcgill.ca

Rustam Z. Khaliullin — *Department of Chemistry, McGill University, Montreal, Quebec QC H3A 0B8, Canada;* ⓞ orcid.org/0000-0002-9073-6753; Email: rustam.khaliullin@mcgill.ca

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.9b01286

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ 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) Edmiston, C.; Ruedenberg, K. Localized Atomic and Molecular Orbitals. *Rev. Mod. Phys.* **1963**, *35*, 457—465.

(3) 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.

(4) von Niessen, W. Density Localization of Atomic and Molecular Orbitals. I. *J. Chem. Phys.* **1972**, *56*, 4290—4297.

(5) Weinhold, F. Natural Bond Orbital Analysis: A Critical Overview of Relationships to Alternative Bonding Perspectives. *J. Comput. Chem.* **2012**, *33*, 2363—2379.

(6) Goedecker, S.; Colombo, L. Efficient Linear Scaling Algorithm for Tight-Binding Molecular Dynamics. *Phys. Rev. Lett.* **1994**, *73*, 122—125.

(7) Bowler, D.; Miyazaki, T. Methods in Electronic Structure Calculations. *Rep. Prog. Phys.* **2012**, *75*, 036503.

(8) Zaleśny, R.; Papadopoulos, M.; Mezey, P.; Leszczynski, J., Eds. *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*; Springer Science & Business Media, 2011.

(9) Pulay, P.; Saebø, S. Orbital-Invariant Formulation and Second-Order Gradient Evaluation in Møller-Plesset Perturbation Theory. *Theor. Chim. Acta* **1986**, *69*, 357—368.

(10) Saebø, S.; Pulay, P. A Low-Scaling Method for Second Order Møller-Plesset Calculations. *J. Chem. Phys.* **2001**, *115*, 3975—3983.

(11) Pisani, C.; Busso, M.; Capecchi, G.; Casassa, S.; Dovesi, R.; Maschio, L.; Zicovich-Wilson, C.; Schütz, M. Local-MP2 Electron Correlation Method for Nonconducting Crystals. *J. Chem. Phys.* **2005**, *122*, 094113.

(12) Hampel, C.; Werner, H. Local Treatment of Electron Correlation in Coupled Cluster Theory. *J. Chem. Phys.* **1996**, *104*, 6286—6297.

(13) Förner, W.; Knab, R.; Čížek, J.; Ladik, J. Numerical Application of the Coupled Cluster Theory with Localized Orbitals to Polymers. IV. Band Structure Corrections in Model Systems and Polyacetylene. *J. Chem. Phys.* **1997**, *106*, 10248—10264.

(14) Saebo, S.; Pulay, P. Local Treatment of Electron Correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213—236.

(15) Schütz, M.; Hetzer, G.; Werner, H. Low-Order Scaling Local Electron Correlation Methods. I. Linear Scaling Local MP2. *J. Chem. Phys.* **1999**, *111*, 5691—5705.

(16) Hetzer, G.; Schütz, M.; Stoll, H.; Werner, H. Low-Order Scaling Local Correlation Methods II: Splitting the Coulomb Pperator in Linear Scaling Local Second-Order Møller-Plesset Perturbation Theory. *J. Chem. Phys.* **2000**, *113*, 9443—9455.

(17) Schütz, M.; Werner, H. Low-Order Scaling Local Electron Correlation Methods. IV. Linear Scaling Local Coupled-Cluster (LCCSD). *J. Chem. Phys.* **2001**, *114*, 661—681.

(18) Scheiber, H.; Shi, Y.; Khaliullin, R. Z. Compact orbitals enable low-cost linear-scaling ab initio molecular dynamics for weakly-interacting systems. *J. Chem. Phys.* **2018**, *148*, 231103.

(19) Marzari, N.; Mostofi, A.; Yates, J.; Souza, I.; Vanderbilt, D. Maximally Localized Wannier functions: Theory and Applications. *Rev. Mod. Phys.* **2012**, *84*, 1419—1475.

(20) Weinstein, H.; Pauncz, R. M.; Cohen, M. Localized Molecular Orbitals. *Adv. At. Mol. Phys.* **1971**, *7*, 97—140.

(21) Bytautas, L.; Ruedenberg, K. Electron Pairs, Localized Orbitals and Electron Correlation. *Mol. Phys.* **2002**, *100*, 757—781.

(22) Bytautas, L.; Ivanic, J.; Ruedenberg, K. Split-Localized Orbitals Can Yield Stronger Configuration Interaction Convergence Than Natural Orbitals. *J. Chem. Phys.* **2003**, *119*, 8217—8224.

(23) Mulliken, R. Electronic Population Analysis on LCAO-MO Molecular Wave Functions. *J. Chem. Phys.* **1955**, *23*, 1833—1840.

(24) Löwdin, P. On the Non-orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. *J. Chem. Phys.* **1950**, *18*, 365—375.

(25) Lehtola, S.; Jónsson, H. Pipek-Mezey Orbital Localization Using Various Partial Charge Estimates. *J. Chem. Theory Comput.* **2014**, *10*, 642—649.

(26) Jansík, B.; Høst, S.; Kristensen, K.; Jørgensen, P. Local Orbitals by Minimizing Powers of the Orbital Variance. *J. Chem. Phys.* **2011**, *134*, 194104.

(27) Høyvik, I.; Jansík, B.; Jørgensen, P. Orbital Localization Using Fourth Central Moment Minimization. *J. Chem. Phys.* **2012**, *137*, 224114.

(28) Høyvik, I.; Jansík, B.; Jørgensen, P. Trust Region Minimization of Orbital Localization Functions. *J. Chem. Theory Comput.* **2012**, *8*, 3137−3146.

(29) Høyvik, I.-M.; Jansík, B.; Jørgensen, P. Pipek−Mezey localization of occupied and virtual orbitals. *J. Comput. Chem.* **2013**, *34*, 1456−1462.

(30) Lehtola, S.; Jónsson, H. Unitary Optimization of Localized Molecular Orbitals. *J. Chem. Theory Comput.* **2013**, *9*, 5365−5372.

(31) Resta, R. Quantum-Mechanical Position Operator in Extended Systems. *Phys. Rev. Lett.* **1998**, *80*, 1800−1803.

(32) Resta, R.; Sorella, S. Electron Localization in the Insulating State. *Phys. Rev. Lett.* **1999**, *82*, 370−373.

(33) Silvestrelli, P. Maximally Localized Wannier Functions for Simulations with Supercells of General Symmetry. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59*, 9703−9706.

(34) Berghold, G.; Mundy, C.; Romero, A.; Hutter, J.; Parrinello, M. General and Efficient Algorithms for Obtaining Maximally Localized Wannier Functions. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2000**, *61*, 10040−10048.

(35) Jónsson, E. O.; Lehtola, S.; Puska, M.; Jónsson, H. Theory and Applications of Generalized Pipek-Mezey Wannier Functions. *J. Chem. Theory Comput.* **2017**, *13*, 460−474.

(36) Høyvik, I.; Olsen, J.; Jørgensen, P. Generalising Localisation Schemes of Orthogonal Orbitals to the Localisation of Non-orthogonal Orbitals. *Mol. Phys.* **2017**, *115*, 16−25.

(37) Anderson, P. W. Self-Consistent Pseudopotentials and Ultra-localized Functions for Energy Bands. *Phys. Rev. Lett.* **1968**, *21*, 13−16.

(38) Diner, S.; Malrieu, J. P.; Claverie, P.; Jordan, F. Fully Localized Bond Orbitals and The Correlation Problem. *Chem. Phys. Lett.* **1968**, *2*, 319−323.

(39) Magnasco, V.; Musso, G. Localized Orbitals and Short-Range Molecular Interactions. I. Theory. *J. Chem. Phys.* **1974**, *60*, 3744−3748.

(40) Payne, P. The Hartree-Fock Theory of Local Regions in Molecules. *J. Am. Chem. Soc.* **1977**, *99*, 3787−3794.

(41) Mehler, E. Self-Consistent, Nonorthogonal Group Function Approximation for Polyatomic Systems. I. Closed Shells. *J. Chem. Phys.* **1977**, *67*, 2728−2739.

(42) Feng, H. S.; Bian, J.; Li, L. M.; Yang, W. T. An Efficient Method for Constructing Nonorthogonal Localized Molecular Orbitals. *J. Chem. Phys.* **2004**, *120*, 9458−9466.

(43) Cui, G.; Fang, W.; Yang, W. Efficient Construction of Nonorthogonal Localized Molecular Orbitals in Large Systems. *J. Phys. Chem. A* **2010**, *114*, 8878−8883.

(44) Liu, S. B.; Perez-Jorda, J. M.; Yang, W. T. Nonorthogonal Localized Molecular Orbitals in Electronic Structure Theory. *J. Chem. Phys.* **2000**, *112*, 1634−1644.

(45) Peng, L.; Gu, F. L.; Yang, W. T. Effective Preconditioning for Ab initio Ground State Energy Minimization with Non-orthogonal Localized Molecular Orbitals. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15518−15527.

(46) Sundberg, K.; Bicerano, J.; Lipscomb, W. Variationally Deorthogonalized Localized Molecular Orbitals. *J. Chem. Phys.* **1979**, *71*, 1515−1524.

(47) Head-Gordon, M.; Maslen, P.; White, C. A Tensor Formulation of Many-Electron Theory in a Nonorthogonal Single-Particle Basis. *J. Chem. Phys.* **1998**, *108*, 616−625.

(48) Kühne, T. D.; Iannuzzi, M.; Ben, M. D.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F.; Golze, D.; Wilhelm, J.; Chulkov, S.; Bani-Hashemian, M. H.; Weber, V.; Borstnik, U.; Taillefumier, M.; Jakobovits, A. S.; Lazzaro, A.; Pabst, H.; Müller, T.; Schade, R.; Guidon, M.; Andermatt, S.; Holmberg, N.; Schenter, G. K.; Hehn, A.; Bussy, A.; Belleflamme, F.; Tabacchi, G.; Glö, A.; Lass, M.; Bethune, I.; Mundy, C. J.; Plessl, C.; Watkins, M.; VandeVondele, J.; Krack, M.; Hutter, J. CP2K: An Electronic Structure and Molecular Dynamics Software Package - Quickstep: Efficient and Accurate Electronic Structure Calculations. *arXiv (Chemical Physics)*, Mar. 8, 2020, https://arxiv.org/abs/2003.03868 (accessed 2020-04-16).

(49) Borštnik, U.; VandeVondele, J.; Weber, V.; Hutter, J. Sparse Matrix Multiplication: The Distributed Block-Compressed Sparse Row Library. *Parallel Comput.* **2014**, *40*, 47−58.

(50) Becke, A. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.

(51) Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti Correlation-Energy Formula into A Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.

(52) Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 1703−1710.