THE JOURNAL OF PHYSICAL CHEMISTRY A
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCA Article

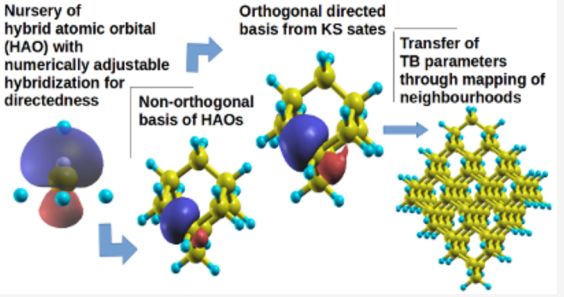
Hybrid Atomic Orbital Basis from First Principles: Bottom-Up Mapping of Self-Energy Correction to Large Covalent Systems

Manoar Hossain, Joydev De, and Joydeep Bhattacharjee*

Cite This: J. Phys. Chem. A 2021, 125, 6805−6817 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Construction of hybrid atomic orbitals is proposed as the approximate common eigenstates of finite first moment matrices. Their hybridization and orientation can be a priori tuned as per their anticipated neighborhood. Their Wannier function counterparts constructed from the Kohn−Sham (KS) single particle states constitute an orthonormal multiorbital tight binding (TB) basis resembling hybrid atomic orbitals locked to their immediate atomic neighborhood, while spanning the subspace of KS states. The proposed basis thus renders predominantly single TB parameters from first principles for each nearest neighbor bond involving no more than two orbitals irrespective of their orientation and also facilitates an easy route for the transfer of such TB parameters across isostructural systems exclusively through mapping of neighborhoods and projection of orbital charge centers. With hybridized 2s, 2p and 3s, 3p valence electrons, the spatial extent of the self-energy correction (SEC) to TB parameters in the proposed basis is found to be localized mostly within the third nearest neighborhood, thus allowing effective transfer of self-energy-corrected TB parameters from smaller reference systems to much larger target systems, with nominal additional computational cost beyond that required for explicit computation of SEC in the reference systems. The proposed approach promises inexpensive estimation of the quasi-particle structures of large covalent systems with workable accuracy.



1. INTRODUCTION

Setting a minimal tight binding (TB) basis for a given systems of atoms calls for the appropriate orientation of orbitals at each atomic site in accordance with their immediate atomic neighborhood so that the nearest neighbor interactions can be represented by the least number of orbitals. In this direction, hybrid atomic orbitals have been used by quantum chemists since their introduction[1,2] almost a century ago. Rational approaches for their construction[3−7] over the past several decades have been primarily focused on partitioning systems into substructures which are spanned by groups of hybrid orbitals, leading to unambiguous partitioning of electrons into bonding orbitals and lone pairs, and further into atomic orbitals. For such partitioning, notionally similar several approaches[4,6,8−11] have been proposed grossly based on the maximum overlap condition which in effect leads to localization of orbitals within the chosen subspace of molecular orbitals. In these approaches, either the overlap matrix[4,6] or the first-order density matrices,[10,12] both of which are calculated typically on the basis of either Slater type orbitals (STOs)[13] or Gaussian type orbitals (GTOs),[14,15] are generally transformed into block diagonal forms, each spanned by orbitals centered on nearest neighbor atoms. The resultant hybrid orbitals involving atomic orbitals centered on more than one atom[7,16] render unambiguous bonding orbitals and bond orders, while the ones like natural hybrid orbitals (NHOs),[12] the effective atomic orbital (EAO),[11] oriented quasi-atomic orbitals,[17] or the ones constructed using the maximal orbital analysis[18] approach, attempt to describe the state of the orbitals of the atoms as they participate in bonds. Hybrid orbitals in the line of NHOs have been popularly constructed ab initio at the HF level.[19,20]

A more explicit approach[21,22] has been to construct generalized hybrid orbitals (GHOs) as combinations of STOs with common Slater exponents and fixed positions of nodes along bonds to assign their orientation. It is expedient to clarify that in this paper we refer to bonds simply as the linear connectivity between atoms which are primarily nearest neighbors if not mentioned specifically. Many of these efforts were undertaken in aid to molecular mechanics calculation[23,24] where the description of interactions between substructures eases with the use of orbitals which are directed along bonds. Effective analytical models for such interactions have also been developed[25] recently for inexpensive deductive computation of

Received: January 13, 2021
Revised: July 17, 2021
Published: July 29, 2021

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

properties of bulk as well as clusters of sp^x hybridized covalent systems. Notably, unlike GHOs, NHOs or EAOs by construction may not be oriented exactly along the bonds. In general for all such hybrid orbitals, their directed nature, maximal localization, and orthonormality are not guaranteed simultaneously by construction. In a part of this work we explore the simultaneity of these conditions in construction of hybrid atomic orbitals from first principles proposed in this work.

Instead of overlap or density matrices, in this work we take recourse to first moment matrices (FMM) due to their direct correspondence to localization. FMMs are known not to commute among each other in more than one dimension if projected on to a finite subspace of orthonormal states. We propose construction of hybrid atomic orbitals (HAOs) as approximate eigenstates of the FMMs within a finite subspace of Kohn−Sham (KS) states of isolated atoms. Orientation and hybridization of the proposed orbitals can be a priori naturalized as per their anticipated neighborhood. This substantially eases the effort of orienting them appropriately while transferring them from isolated atoms to real systems, which eventually eases the interpretation of elements of the Hamiltonian. An orthonormal set of localized Wannier orbitals resembling HAOs is further constructed on the basis of KS single particle states of the given system. These Wannier orbitals, which we refer to in this paper as hybrid atomic Wannier orbitals (HAWOs), constitute a multiorbital tight binding (TB) basis locked to their immediate atomic neighborhood by construction and render hopping parameters involving effectively only two orbitals per bond. HAWOs thus offer easy transfer of the corresponding TB parameters to other isostructural systems exclusively through mapping of neighborhoods and projection of charge centers learned from HAOs. Effective transfer of TB parameters is demonstrated in nanoribbons of graphene and hexagonal boron nitride, C_{60}, and nanodiamonds and their silicon based counterparts. In particular, we show in the HAWO basis that it is possible to effectively transfer the self-energy (SE) correction (SEC) of single particle levels from smaller reference systems to much larger isostructural systems through TB parameters with minimal additional computational expense through the proposed mapping of multiorbital TB parameters beyond the nearest neighborhood.

2. METHODOLOGICAL DETAILS

2.A. Construction of Hybrid Orbitals. In a given direction, for example along $\hat{x}$, the most localized orbitals $\{\phi\}$ would diagonalize the corresponding FMM:

$$X_{ij} = \langle \phi_i | x | \phi_j \rangle \tag{1}$$

This becomes clear by noting that the total spread of a finite set of N number of orbitals along $\hat{x}$ is given by

$$\Omega_x = \sum_{i=1,N} [\langle \phi_i | x^2 | \phi_i \rangle - |\langle \phi_i | x | \phi_i \rangle|^2] \tag{2}$$

which can be expressed as

$$\begin{aligned}\Omega_x &= \sum_{i=1,N} \left(\sum_{j=1}^{\infty} X_{ij} X_{ji} - X_{ii} X_{ii} \right) \\ &= \sum_{i=1,N} \sum_{j\neq i}^{\infty} |X_{ij}|^2 \\ &= \sum_{i=1,N} \left(\sum_{j\neq i}^{N} |X_{ij}|^2 + \sum_{j=N+1}^{\infty} |X_{ij}|^2 \right)\end{aligned} \tag{3}$$

Diagonalization of X in the $N \times N$ subspace would therefore set the first term in eq 3 to zero, leading to minimization of the total spread. Notably, X can be calculated directly as in eq 1 only for isolated systems well separated from their periodic images. For a periodic system with nonzero crystal momentum, computation of X would essentially involve evaluation of geometric phases[26] of Bloch electrons evolved across the Brillouin zone.[27,28] Nevertheless, there exists therefore a unique set of orbitals which completely diagonalize X and would also thereby have maximum localization along $\hat{x}$. Similar unique sets exist for the $\hat{y}$ and $\hat{z}$ directions as well. However, the matrices X, Y, and Z, when projected into a finite subspace of orthonormal states, do not commute with each other in general unless mandated by symmetries. This implies that a unique set of orbitals with maximum localization simultaneously in all three orthogonal directions would not exist in general. The same is true for Wannier functions (WFs) in the case of periodic systems with nonzero wave vectors. Numerically localized Wannier functions[29,30] therefore are not unique, and the choice of gauge used for their construction depends on the chosen criteria of localization.

We chose to look for the possibility to construct a set of localized orbitals which will be a reasonable compromise between the three unique sets of orbitals having maximum localization along the three orthogonal directions. We thus resorted to the condition of simultaneous approximate joint diagonalization[31] of the three FMMs: X, Y, and Z. To compute such an approximate eigensubspace of the three FMMs, we adopted an iterative scheme based on generalization of the Jacobi method of matrix diagonalization,[32] wherein off-diagonal elements are iteratively minimized by applying rotation of coordinates by an optimally chosen angle. The extension of the method to more than one square matrix irrespective of whether they are commuting or not, is based on a proposed[31] choice of angle of rotation leading to the complex rotation matrix U which has been proven[31] to minimize the composite objective function (off) defined as

$$\text{off}(UXU^{\dagger}) + \text{off}(UYU^{\dagger}) + \text{off}(UZU^{\dagger}) \tag{4}$$

where $\text{off}(A) = \sum_{1\leq i\neq j\geq N} |A_{ij}|^2$ for an $N \times N$ matrix A. N is the number of orthonormal states used to compute X, Y, and Z. U is a product of all the $N(N-1)/2$ complex plane rotations, one each for each pair of (ij) for $i \neq j$. For a given (ij) the plane rotation $R(i,j)$ is an $N \times N$ identity matrix except for

$$\begin{pmatrix} r_{ii} & r_{ij} \\ r_{ji} & r_{jj} \end{pmatrix} = \begin{pmatrix} c & \bar{s} \\ -s & \bar{c} \end{pmatrix} \tag{5}$$

where $c, s \in C$ and $|c|^2 + |s|^2 = 1$.

It has been shown[31] that the objective function defined in eq 4 is minimized if U is a product of $R(i,j)$ matrices as shown in eq 5 whose elements are given as

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

$$c = \sqrt{\frac{x + r}{2r}}; \qquad s = \frac{y - iz}{\sqrt{2r(x + r)}} \tag{6}$$

where

$$r = \sqrt{x^2 + y^2 + z^2}$$

and $[x,y,z]^\dagger$ is the eigenvector corresponding to the highest eigenvalue of a 3 × 3 matrix:

$$\begin{aligned} G(i, j) = {} & \text{Real}(h^\dagger(X, i, j)\, h(X, i, j)) \\ & + \text{Real}(h^\dagger(Y, i, j)\, h(Y, i, j)) \\ & + \text{Real}(h^\dagger(Z, i, j)\, h(Z, i, j)) \end{aligned}$$

with

$$h(A, i, j) = [a_{ii} - a_{jj}, a_{ij} + a_{ji}, i(a_{ji} - a_{ij})] \tag{7}$$

Notably, given the form of $R(i,j)$, for a rotated matrix $A' = R(i,j)\, AR^\dagger(i,j)$ corresponding to plane rotation for the (ij)th pair of elements of A, it is easily seen that $a'_{kk} = a_{kk}$ for $k \neq i$ and $k \neq j$, leading to the invariance:

$$\text{off}(A') + |a'_{ii}|^2 + |a'_{jj}|^2 = \text{off}(A) + |a_{ii}|^2 + |a_{jj}|^2$$

owing to preservation of the norm in the similarity transformation. Therefore, minimizing off(A') would naturally imply maximizing $|a'_{ii}|^2 + |a'_{jj}|^2$, which further implies maximizing $|a'_{ii} - a'_{jj}|^2$ since

$$2(|a'_{ii}|^2 + |a'_{jj}|^2) = |a'_{ii} + a'_{jj}|^2 + |a'_{ii} - a'_{jj}|^2$$

and

$$a'_{ii} + a'_{jj} = a_{ii} + a_{jj}$$

owing to invariance of the trace under the similarity transformation. Therefore, in our case the minimization of the objective function (eq 4) implies maximizing the separation between the charge centers of the ith and the jth orbitals, which is thus similar to the principle of the Foster and Boys[33] scheme of orbital localization. This becomes clear by rewriting the total spread (eq 3) for the N orbital $\{\phi_i, i = 1,N\}$ as

$$\Omega = \sum_{k=1,3} \sum_{i=1,N} \left(\sum_{j \neq i}^{N} |a_{ij}^k|^2 + \sum_{j=N+1}^{\infty} |a_{ij}^k|^2 \right) \tag{8}$$

where $A^{k=1,2,3} = X, Y, Z$. Equation 8 clearly suggests that minimization of the objective function in eq 4 minimizes the first term in eq 8, leading to minimization of the total spread. Equation 8 also suggests that the total spread will reduce with increasing number of states (N) on the basis of which the first moment matrices are constructed.

We test the proposed approach first with FMMs computed on the basis of GTOs constructed for Ti with parameters from ref 34. In Figure 1 we plot the charge centers $(\langle\phi|\vec{r}|\phi\rangle)$ of the approximate eigenstates of the first moment matrices.

Evidently, the charge centers constitute coordination polyhedra around isolated atoms which are consistent in shape with those tabulated in Figures 6−8 in ref 35. This agreement confirms the identity of the resultant orbitals as the hybrid orbitals and numerically establishes the connection between maximal localization and hybridization. Such a connection between sp^3 hybridization and minimization of

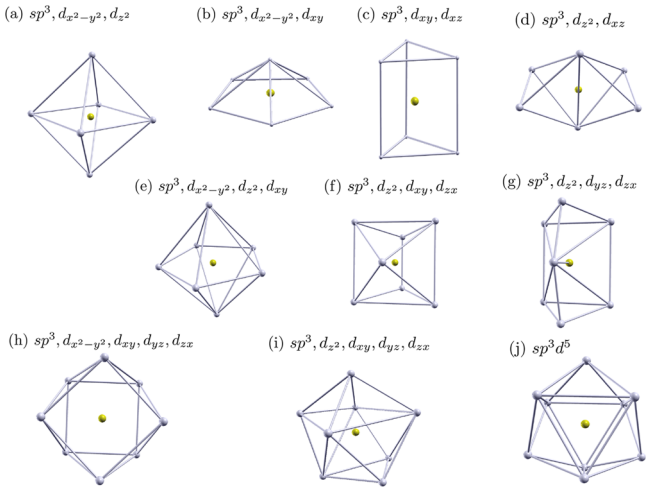


Figure 1. Plots of charge centers (shown in gray) of the hybrid orbitals formed by the group of GTOs representing 3s, 3p, and 3d orbitals of Ti (shown in yellow) constructed as per ref 34.

the total quadratic spread of s and the three p orbitals has been analytically proven.[36] In this work, however, we do not use GTOs further and rather resort to KS states of isolated atoms. For example, for atoms of the p block, such as boron, carbon, nitrogen, and silicon dealt with in this work, if the first moment matrices are constructed on the basis of three (four) KS states with the lowest energies, namely, the one s-like nondegenerate having the lowest energy and two (three) of the three p-like degenerate states above the s-like state, the approximate eigensubspace would render three (four) $2sp^2$ $(2sp^3)$ hybridized orbitals. Notably, for isolated systems such as molecules, clusters, and nanostructures, the approximate common eigenspectrum of the FMMs computed within the manifold of occupied KS states results in partitioning[37−39] of the ground state charge density into bonding and localized orbitals.

2.A.1. Orientation and Transfer of Orbitals. Although, as evident above, construction of HAOs for an isolated atom as such does not require any predefined directionality, the orientation of the HAOs associated with an atom can be nevertheless locked to their anticipated neighborhood by placing the isolated atom within an external potential which represents the generic or exact atomic neighborhood of the given atom in the actual system in which the HAOs are to be used. We construct such external potentials by placing weakly confining spheres with small constant negative potentials inside the spheres in place of exact or generic locations of neighboring atoms as present in the actual system. For example, to lock sp^3 HAOs to a four-coordinated tetrahedral neighborhood, a tetrahedra of confining spheres is placed around the host C atom, leading to orientation of the sp^3 orbitals maximally in the direction of the confining spheres as seen in Figure 5a. Typically we find confining potential amplitudes on the order of 0.01 eV and radius 0.5 Å to be sufficient for the purpose. Such a weak confinement in the vicinity causes a change of KS energy eigenvalues of isolated atoms on the order of 0.001 eV and retains the shape of the lowest KS states which are used for construction of the HAOs, effectively unaltered. For sp^3 HAOs, the tetrahedra of the confining spheres can be an exact tetrahedron, as in the case of bulk Si, or a strained tetrahedron, as in the case of cyclopropane. As is evident in Figure 2a for cyclopropane

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

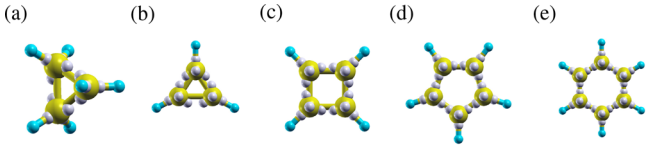


Figure 2. Projected charge centers of HAOs shown by gray spheres depicting their orientations around their host C atom shown in yellow.

and in Figure 2b−e for planar molecules C_nH_n, the projected charge centers of the HAOs (shown in gray) symmetrically deviate from the C−C bonds with decreasing C−C−C angle as we go from C_6H_6 to C_3H_3. For all of these molecules the HAOs were constructed with the weakly confining spheres placed around the host C atom exactly as per their nearest neighbors in the molecules, resulting in HAOs largely retaining their pure sp^3 nature but oriented symmetrically about the directions of the confining spheres from the host atoms. The placement of confining potential spheres thus provides a gross directional reference for orientation of the full set of the HAOs.

The positions of the charge centers of the HAOs are learned in terms of the directions of the confining spheres from the isolated host atom. Such learnings are subsequently used in projecting centers of HAOs around the corresponding atom in a given system, as seen for the molecules in Figure 2 and the nanodiamonds in Figure 4. While HAOs are transferred from their nursery of isolated host atoms to their matching host atoms in a given system, HAOs are rotated such that their actual charge centers align along the direction of their projected centers from the matching host atoms.

In addition to providing reference for orientation, the confining spheres can have an important role in deciding the level of hybridization of the HAOs. This becomes evident by noting that, if we use four KS states and three confining spheres coplanar with the host atom, then instead of forming four sp^3 orbitals the HAOs separate into three $2sp^2$ orbitals and one $2p_z$ orbital, as is evident from the unhybridized shape of the $2p_z$ orbital in Figure 3a. Figure 3 shows the evolution of the

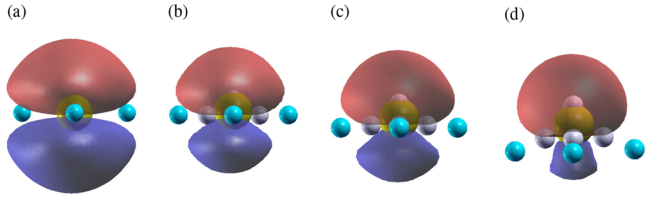


Figure 3. (a−d) Evolution of a pure $2p_z$ orbital (a) from sp^2 hybridization background, to an sp^3 hybridized orbital due to increased deviation of the centers (cyan spheres) of the three confining potential spheres from coplanarity with the host atom (yellow sphere). Centers of HAOs are shown by gray spheres.

$2p_z$ HAO from a pure orbital perpendicular to the plane of sp^2 hybridization, toward a $2sp^3$ hybridized orbital, with increasing noncoplanarity of the confining spheres with the host atom. HAOs with such intermediate hybridization ($2sp^{2+} + 2p_z^{+}$) have been used for C_{60} (Figure 8). However, stronger confining potentials are found necessary to influence the hybridization of KS states, typically on the order of 1 eV for C atoms, such that the orbitals align along the confining spheres. The confining potentials in this case therefore do lead to minor modification of the shape of the KS states, and thereby of the HAOs as well, although not quite obviously at the isosurfaces plotted in Figure 3. However, the values of TB parameters calculated on the basis of their Wannierized counterparts in C_{60} suggests that the overall shape of those orbitals are largely retained close to the sp^2 orbitals. Notably, we could have used stronger confinement to align the HAOs in C_3H_6, C_3H_3, or C_4H_4 as we did for C_{60}, but the degree of confinement would have to be much higher than that used for C_{60}, which would have substantially altered the shape of the HAOs themselves, since it is obvious that with pure s, p_x, p_y, and p_z orbitals it is impossible to form any set of hybrid orbitals in which two orbitals can have relative orientation less than 90°.

2.B. Wannier Functions Based on HAOs. The next step is to construct orthonormalized Wannier functions from the KS states following the HAOs transferred to a given system. The transferred HAOs constitute a nonorthogonal basis of hybridized atomic orbitals. In the general framework of periodic systems with nonzero wave vectors ($\vec{k}$), we begin with constructing a nonorthogonal set of quasi-Bloch states as

$$\tilde{\psi}_{\vec{k},j}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \phi_{\vec{R},j}(\vec{r}) \tag{9}$$

where $\phi_{\vec{R},j}(\vec{r})$ is the jth HAO localized in the unit cell denoted by the lattice vector $\vec{R}$ spanning over N unit cells defining the Born−von Karman periodicity. The projections of the nonorthogonal quasi-Bloch states on the orthonormal Bloch states constructed from the KS single particle states at all allowed $\vec{k}$ are calculated as

$$O_{\vec{k},m,j} = \langle \psi^{KS}_{\vec{k},m} | \tilde{\psi}_{\vec{k},j} \rangle \tag{10}$$

Elements of O thus record the representation of the HAOs within the manifold of KS bands considered. Overlaps between the nonorthogonal quasi-Bloch states within the manifold of the considered KS states are therefore calculated as

$$S_{\vec{k},m,n} = \sum_{l} O^{*}_{\vec{k},l,m} O_{\vec{k},l,n} \tag{11}$$

The degree of representability of HAO ϕ_n, within the set of KS states considered, is guaranteed by setting a lower cutoff on individual $S_{\vec{k},n,n}$ values to be typically more than 0.85. For all the systems studied in this work, the above criterion is found to be satisfied by the lower bound on the number KS states, which is set by the total number of valence orbitals of all atoms of a given system. A new set of orthonormal Bloch states from the KS single particle states are subsequently constructed using the Löwdin symmetric orthogonalization[40] scheme as

$$\Psi_{\vec{k},n}(\vec{r}) = \sum_{m} S_{\vec{k},m,n}^{-1/2} \sum_{l} O_{\vec{k},l,m} \psi^{KS}_{\vec{k},l}(\vec{r}) \tag{12}$$

where the sum over l spans the KS states considered and the sum over m takes care of the orthonormalization. Subsequently, a localized set of orthonormal Wannier functions are constructed as

$$\Phi_{\vec{R}',j}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{k}} e^{-i\vec{k}\cdot\vec{R}'} \Psi_{\vec{k},j}(\vec{r}) \tag{13}$$

In this process the Löwdin symmetric orthogonalization clearly provides a choice of gauge for linear combination of KS states such that the resultant Wannier functions $\{\Phi_{\vec{R}',j}(\vec{r})\}$ resemble the corresponding HAOs $[\{\phi_{\vec{R},j}(\vec{r})\}]$ as much

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

possible within the manifold of KS states considered. Hence we refer to these Wannier functions as the hybrid atomic Wannier orbitals ("HAWOs"). In Figure 5 we show an HAO

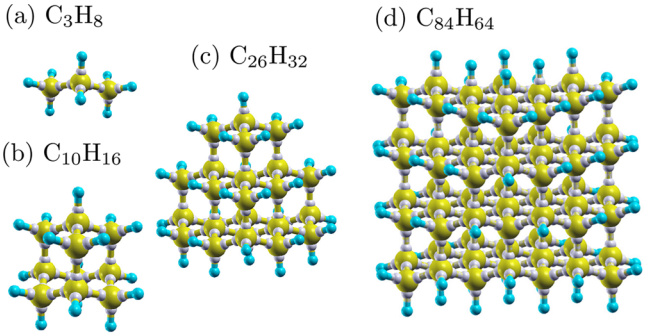


Figure 4. C_nH_m systems with projected charge centers of HAOs shown as gray spheres, used in this work as example of sp^3 hybridized covalent systems.

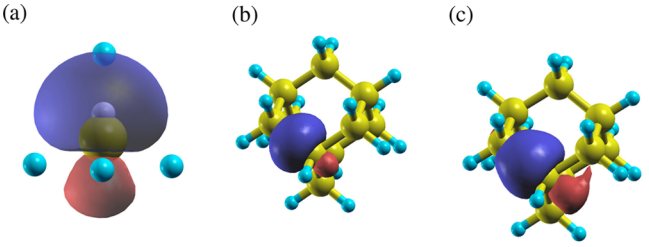


Figure 5. (a) HAO representing an sp^3 orbital of an isolated C atom (yellow sphere) used in this work. Charge center of the orbital is shown in gray. Centers of the confining spheres used to determine gross orientation are shown in cyan. (b) HAO shown in (a) transferred to a C atom an adamantane ($C_{10}H_{16}$) molecule; (c) the corresponding HAWO.

before and after transfer to adamantane and the corresponding HAWO constructed from the KS states of adamantane. HAWOs can thus be considered as analogue of NHOs constructed from a given set of KS states with acceptable representability. Notably, following the same approach, templates of bonding and lone pair orbitals made of HAOs can be used to construct localized Wannier functions rendering orbital resolved description of the valence band.[29,30] Similarly, templates of antibonding orbitals made of HAOs can be used to extract a meaningful description of the unoccupied bands in the line of the valence virtual orbitals.[41]

2.B.1. TB Parameters in HAWO Basis. TB parameters in the HAWO basis are computed from energetics of KS single particle states as

$$t_{\vec{R}',\vec{R},i,j} = \langle \Phi_{\vec{R}',i} | H^{\mathrm{KS}} | \Phi_{\vec{R},j} \rangle = \sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{R}'-\vec{R})} \sum_{l} (OS^{-1/2})^{*}_{li} (OS^{-1/2})_{lj} E^{\mathrm{KS}}_{\vec{k},l} \quad (14)$$

Notably, similar TB parameters have been derived in the past two decades from first principles based on either the maximally localized Wannier function[42−47] or atomic orbitals[48,49] constructed from KS states. Much effort has been reported in deriving TB parameters through projection of KS states on pseudoatomic orbitals[50,51] as well. However, attempts to calculate TB parameters in a hybrid atomic orbital basis constructed from first principles, as proposed in this work, has been limited so far primarily to analytical models.[52,53]

In Figure 6a for cyclopropane, we plot the TB parameters calculated as per eq 14 for two HAOs participating dominantly

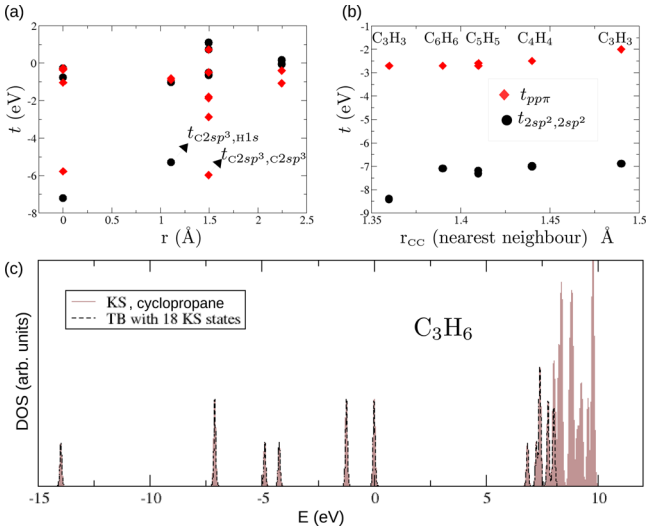


Figure 6. (a) TB parameter calculated for cyclopropane. (b) Nearest neighbor TB parameters between in-plane and out-of-plane orbitals in C_3H_3, C_4H_4, C_5H_5, and C_6H_6 molecules (shown in Figure 2) arranged as a function of C−C bond lengths available in the molecules. (c) DOS calculated from 50 lowest KS eigenvalues, compared with DOS from eigenvalues of TB Hamiltonian constructed from 18 lowest KS states, 18 being the total number of valence orbitals of cyclopropane.

in a C−C bond and a C−H bond. The t_{sp^3,sp^3} value is comparable to that in adamantane ($C_{10}H_{16}$) (Figure 12) despite the substantial misalignment (Figure 2) of HAO and the C−C bond in cyclopropane, while there is perfect alignment of the two in $C_{10}H_{16}$. The hopping parameters are obtained with 18 KS states, which is same as the total number of valence orbitals of all the atoms, resulting thereby in density of states in exact agreement with that obtained from DFT (Figure 6c) as discussed in the next paragraph. In Figure 6b we plot hopping parameters for π and σ bonds as a function of C−C bond lengths available in planar C_3H_3 to C_6H_6 molecules. As is evident in Figure 2, the best alignment of the HAOs along the C−C bond is possible for C_6H_6 and the worst is obviously for the shorter bond of C_3H_3 and similarly for C_3H_6. Yet the highest in-plane hopping parameter in terms of magnitude is found for the shorter bond of C_3H_3, which is about 20% more than that of the C−C in-plane bond of benzene, whereas the C−C bond length in benzene is only about 2.2% more than the shorter bond of C_3H_3. Similarly, the C−C nearest neighbor hoping parameter and the bond length in C_3H_8 are both within 1% of those of C_3H_6, whereas in C_3H_8 the HAOs are almost perfectly aligned along the C−C bond (Figure 4) while in C_3H_6 they are misaligned by more than 20°. These results can possibly be explained by the inherent bent nature of the bonds[54] in C_3H_6 and C_3H_3, reflected by the symmetric misalignment of the HAOs along the two C−C bonds while perfect alignment is maintained along the C−H bonds. We plan to examine this aspect for bent bonds in detail in the future.

As is evident in Figure 7a for $C_{10}H_{16}$, the edge of the valence band is already well described if we consider only the nearest neighbor hopping in the HAWO basis. However, as shown in

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

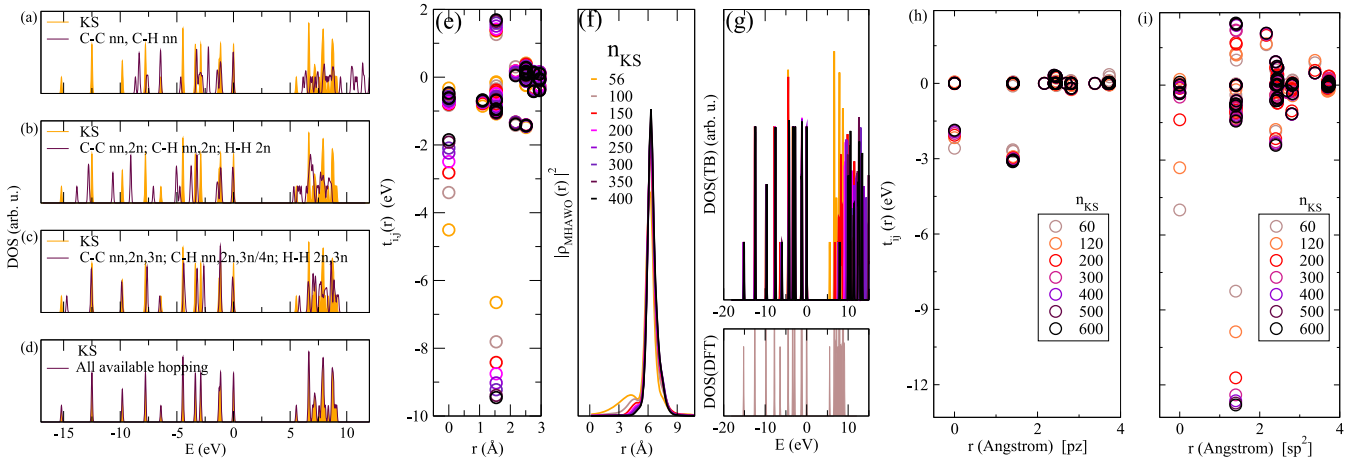


Figure 7. For $C_{10}H_{16}$, (a−d) evolution of density of states (DOS) with increase in range of hopping starting from (a) the nearest neighbor (nn) to (d) all available hopping graduating through hopping between second (2n) and third (3n) nearest neighbors and beyond. Convergence of (e) TB parameters and (f) spatial localization of $2sp^3$ orbitals and (g) TB DOS, in terms of the number of KS states used in construction of HAWOs as mentioned in the legend of (f). KS DOS is shown below (g). Similar convergence of TB parameters for (h) $2p_z$ and (i) $2sp^2$ orbitals in AGNR ($3p + 1$, $p = 2$).

Figure 7b onward, the match of DOS from TB and DFT improves drastically with increasing extent of hopping considered up to the second nearest neigbour. This is immediately understood by noting the non-nominal positive value of the second nearest hopping element plotted in Figure Figure 7e, arising due to the proximity of lobes of different signs of the two HAOs. In Figure 7e−g, we demonstrate the evolution of the TB parameters, HAWOs, and DOS from TB, as a function of the number of KS states considered for the construction of HAWOs. The rationale for this analysis is the possibility that the antibonding subspace may not be adequately represented by the unoccupied KS states if we restrict the total number of KS states to be the same as the total number of HAOs associated with all the atoms, which is same as the total number of valence orbitals of all the atoms. Indeed we see clear convergence of the shape of the HAWO (Figure 7f) as well as the corresponding TB parameters (Figure 7e) if we consider KS states in excess of the total number of HAOs. Figure 7h,i suggests that the convergence can be much quicker for unhybridized orbitals such as $2p_z$ compared to hybridized orbitals such as sp^2 and sp^3, since the unhybridized orbitals primarily constitute the edges of the valence and conduction bands. However, the TB DOS expectedly starts deviating from the DFT DOS more in the conduction band (Figure 7g) if we include more KS states beyond the total number of HAOs, owing to the semiunitary nature of the net transformation matrix ($OS^{1/2}$) implied in eq 12 which will be rectangular in such scenarios. It is thus important to decide on the number of KS states to be considered depending on the purpose. If the aim is to represent only the valence bands through well-localized HAWOs, then it may be prudent to look for convergence of HAWOs in terms of KS states. However, if the band gap needs to be represented accurately by the TB parameters, then the number of KS states should be kept the same as the total number of valence orbitals.

2.C. Bottom-Up Mapping of TB Parameters. The HAWO basis derived from the KS states offers a multiorbital TB basis which is by construction locked to the local coordination as per the atomic neighborhood of each atom. The TB parameters derived in such a basis should therefore be transferable from one system to another with a matching atomic environment. A key aim of this work is to demonstrate such a transferability for effective transfer of multiorbital TB parameters in the HAWO basis from smaller reference systems to larger target systems. The mapping of TB parameters is done in two steps.

1. Pairs of atoms of the target system, not limited to nearest neighbors, are mapped onto pairs of atoms in the reference system based on a collection of criteria.

2. Among the mapped pairs of atoms, pairs of system orbitals are mapped to pairs of reference orbitals through mapping of their respective projected charge centers. In step 1 the criteria to map pairs of atoms include matching structural parameters such as their spatial separation and their individual nearest neighborhoods characterized in terms of the type of neighboring atoms and angles made by nearest neighbors on the atoms. In particular, we use a parameter calculated as

$$\zeta_i = \sum_j^{N_i} Z_j w(r_{i,j}) \tag{15}$$

where N_i is the number of neighbors of the ith atom within a suitably chosen cutoff radius, w is a weight factor which is a function of the distance $r_{i,j}$ of the jth neighbor of the ith atom, and Z_i is a characteristic number to be associated with each type of atom. Z_i can be chosen to be the atomic weight, as we mostly used in this work, or a similar number which can facilitate identification of a type of neighborhood or a region of the system through values of ζ. In this work we chose the weight factor w to be 1.0 within half of the cutoff radius, beyond which the factor is smoothly reduced to zero using a cosine function. The choice of cutoff radius depends on the size of the reference system. It should not be too large for variations to average out, nor should it be too small to become insensitive to morphological variations in the reference system itself. ζ allows us to map atom pairs effectively through a prudent choice of values of $\{Z_i\}$ since it would allow assessment of the proximity of atoms to edges, interfaces, or any kind of structural inhomogeneity without any exhaustive structural relaxation.

In step 1, the minimum of the deviation

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

$$|\zeta_1^{\text{target}} - \zeta_1^{\text{reference}}| + |\zeta_2^{\text{target}} - \zeta_2^{\text{reference}}|$$

obtained within a range of allowed deviations of structural parameters is used as the criterion to choose matching pairs of atoms between target and reference systems.

As in step 1, in step 2 the mapping of one or a pair of HAOs from the reference to target systems is done on the basis of matching structural parameters, as well as a parameter calculated as

$$\xi_i = \sum_j^{N_i} \zeta_j w^{\text{WC}}(r_{i,j}) \tag{16}$$

where ζ_j corresponds to the jth atom in the neighborhood defined by w^{WC} around the projected charge center of the ith HAO. The angle made by the directions of the projected charge centers of the HAOs from their respective host atoms is a key matching parameter in step 2. Additionally, if the HAOs belong to different atoms, then the dihedral angle made by the centers of the HAOs through the axis connecting their host atoms is also a key parameter. Thus, in step 2, the minimum of the deviation

$$|\xi_1^{\text{target}} - \xi_1^{\text{reference}}| + |\xi_2^{\text{target}} - \xi_2^{\text{reference}}|$$

within acceptable deviations of structural parameters defines matching pairs of HAOs.

As an example, we show mapping from a small curved finite patch (Figure 8a) to C_{60}. Since C_{60} constitutes a curved surface without any edge, mapping should be done from the innermost neighborhood of the chunk. Since in C_{60} the angles made by nearest neighbors at a given atom differ distinctly depending on whether an angle opens inside a pentagon or a hexagon, the matching parameters for mapping are mostly structural, primarily the direct and dihedral angles. The reference patch is cropped from C_{60} and passivated by H. We fix the tolerance of ζ to zero, which implies that C_{60} is getting mapped from only six C atoms of the patch (Figure 8a) having all C neighbors. Given the curvature of C_{60}, we chose to use confining spheres to influence the hybridization of sp^2 HAOs in order to break their coplanarity and align them along nearest neighbor C−C bonds, as shown in Figure 3c, where the placement of confining potential spheres are as per the nearest neighborhood in C_{60}. The projected charge centers of HAOs with intermediate hybridization ($2sp^{2+} + 2p_z^{+}$) between sp^2 + p_z and sp^3, shown in Figure 8c, is used to map from that of the reference shown in Figure 8b. The TB parameters $t_{2sp^{2+},2sp^{2+}}$ for the shorter and longer C−C bonds are about −6.9 and −6.5 eV, whereas $t_{2p_z^{+},2p_z^{+}}$ values are about −2.36 and −2.0 eV. The match of the DFT DOS with the DOS from TB parameters mapped from the reference system is shown in Figure 8d, which can be further improved beyond the valence bond by considering HAOs for excited states, which will be taken up in a subsequent work on optical properties.

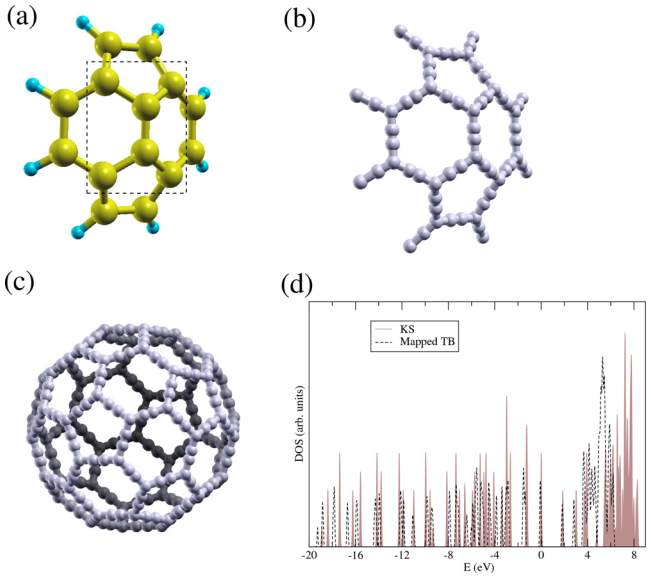


Figure 8. (a) Structure of reference system and (b) corresponding charge centers of HAOs with intermediate hybridization ($2sp^{2+} + 2p_z^{+}$) between sp^2 and sp^3. (c) Projected charge center with similar hybridization for C_{60}. (d) Corresponding matches of DFT DOS with TB DOS with parameters mapped from the reference system.

2.D. Self-Energy Correction of TB Parameters. Self-energy-corrected TB parameters $\{t^{\text{QP}}_{\vec{R}',\vec{R},i,j}\}$ in the HAWO basis are calculated by substituting $E^{\text{KS}}_{k,n}$ in eq 14 by quasiparticle energies $E^{\text{QP}}_{k,n}$ obtained at the G_0W_0 level, which is the first-order non-self-consistent GW approximation of MBPT.[55,56] Within the GW approximation, the quasi-particle energies are approximated as

$$E^{\text{QP}}_{\vec{k},n} = E^{\text{KS}}_{\vec{k},n} + \langle \psi^{\text{KS}}_{\vec{k},n} | \Sigma - V^{\text{KS}}_{\text{xc}} | \psi^{\text{KS}}_{\vec{k},n} \rangle \tag{17}$$

where $V^{\text{KS}}_{\text{xc}}$ is the mean-field exchange−correlation potential and Σ[57] is the self-energy operator derived by considering the many-electron effects as a perturbation treated within a self-consistent framework of Dyson's equation formulated in terms of the one-particle dynamic nonlocal Green's function constructed from the KS states. Similar efforts have been reported in recent years on incorporating the SEC in TB parameters computed in terms of the maximally localized Wannier functions.[58−60] Incorporation of the SEC in TB parameters has also been attempted through matching specific bands of the QP structure.[61−63]

3. COMPUTATIONAL DETAILS

Electronic structures of the ground states of all the systems considered in this work were calculated using the Quantum Espresso (QE) code,[64] which is a plane wave based implementation of DFT. We have used norm-conserving pseudopotentials with the Perdew−Zunger (LDA) exchange−correlation[65] functional and a plane wave cutoff of 60 Ry for wave functions and commensurately more for charge density and potential. Variable cell structural relaxation has been carried out for all periodic systems. We used a 1 × 1 × 15 Monkhorst−Pack grid of k-points for AGNRs and a 1 × 1 × 29 Monkhorst−Pack grid of k-points for ZGNRs as well as for ZBNNRs. Self-energy corrections to single particle levels have been estimated at the non-self-consistent G_0W_0 level of the GW approximation implemented in the BerkeleyGW code.[66] To calculate the static dielectric matrix required for computation of the self-energy operator, the generalized plasmon-pole model[57] is used to extend the static dielectric matrix in the finite frequencies. For all the nanoribbons, parameters are chosen from ref 67. In-house implementation interfaced with the QE code is used for generation of HAOs, HAWOs from KS states, calculation of TB parameters in the

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

HAWO basis, and mapping of TB parameters from reference to target systems.

3.A. Mapping Self-Energy-Corrected TB Parameters in HAWO Basis. *3.A.1. Nanoribbons.* We have recently reported[68] estimation of the quasi-particle (QP) band gap for graphene and hBN nanoribbons based on the self-energy correction to TB parameters mapped from narrower ribbons on the basis of a single $2p_z$ electron per atom. The transfer was made explicitly by identifying equivalent atoms based on proximity to the ribbon edges as shown in Figures 2c−e, 3f, and 5a in ref 68. In this section we begin by systematizing the process of identifying the equivalent atoms through the mapping mechanism proposed in section 2.C. The identification is primarily based on ζ values for atoms and ξ values for HAO charge centers wherever sufficient variations of ζ and ξ are available in the reference systems, as demonstrated in Figures 9c,f,i and 11c,f,i for hBN and ZGNR, respectively.

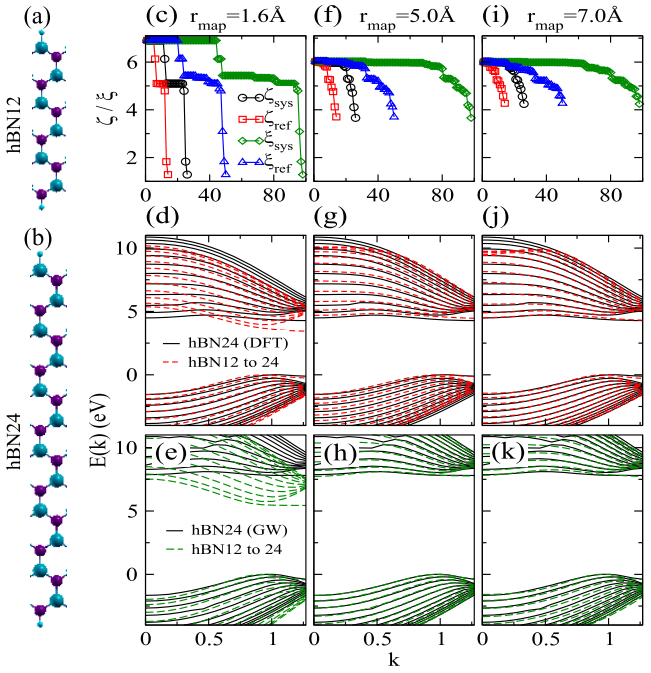


Figure 9. (a, b) Hexagonal zigzag boron nitride nanoribbons (hZBNNR, hBN) hBN12 and hBN24, respectively. (c, f, i) Plots of ζ and ξ values "ref" (reference hBN12) and "sys" (target hBN24) for different spatial ranges of neighborhood considered for mapping. (d, g, j) Matching of DFT band structure and mapped TB band structure for increasing r_{map}. (e, h, k) Matching of DFT+G_0W_0 band structure and mapped self-energy-corrected TB band structure for increasing r_{map}.

Mapping of AGNR from $p = 1$ to $p = 4$, as shown in Figure 10d,g,j,m, calls for matching of structural parameters as the key strategy for mapping, since the width of reference AGNR with $p = 1$ of the $3p + 1$ family is narrow enough and there are only two types of C atoms per unit cell.

For hBN, an acceptable match (Figure 9g,j and h,k) between explicitly computed band structures, and the same computed from TB parameters with only $2p_z$ orbitals mapped from a narrower ribbon (Figure 9a), is achieved simultaneously at the DFT and DFT+G_0W_0 levels, with hopping considered within the range of no less than 5 Å. A higher spatial range of mapping of self-energy-corrected TB parameters is required for matching of band structures at the DFT+G_0W_0 level for

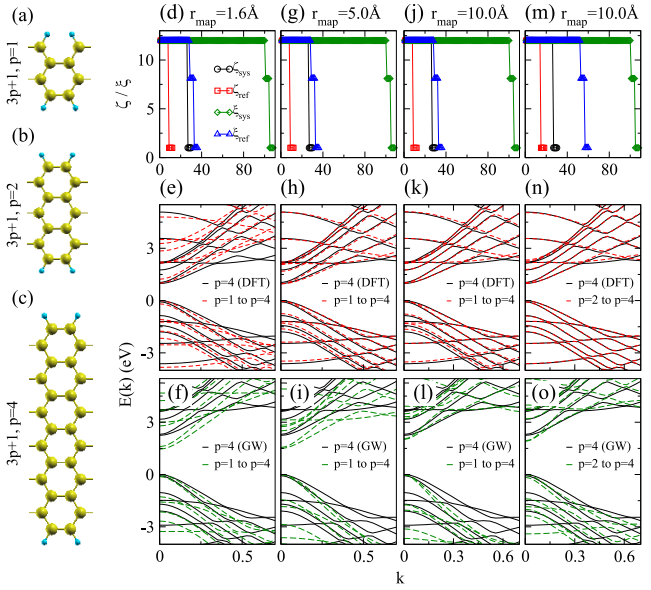


Figure 10. (a−c) Armchair graphene nanoribbons (AGNR) of family $n = 3p + 1$ with (a) $p = 1$, (b) $p = 2$, and (c) $p = 4$, respectively. (d, g, j, m) Plots of ζ and ξ values "ref" (reference $p = 1$ (d, g, j) and $p = 2$ (m)) and "sys" (target $p = 4$) for different spatial ranges (r_{map}) of neighborhood considered for mapping. (e, h, k, n) Matching of DFT band structure and mapped TB band structure for increasing r_{map}. (f, i, l, o) Matching of DFT+G_0W_0 band structure and mapped self-energy-corrected TB band structure for increasing r_{map}.

AGNR (Figure 10l), and more so for ZGNRs (Figure 11k) with smaller band gaps. Notably, for AGNR, the match of self-energy-corrected band edges for $p = 4$ naturally improves with mapping from $p = 2$ (Figure 10o). These trends simply relate

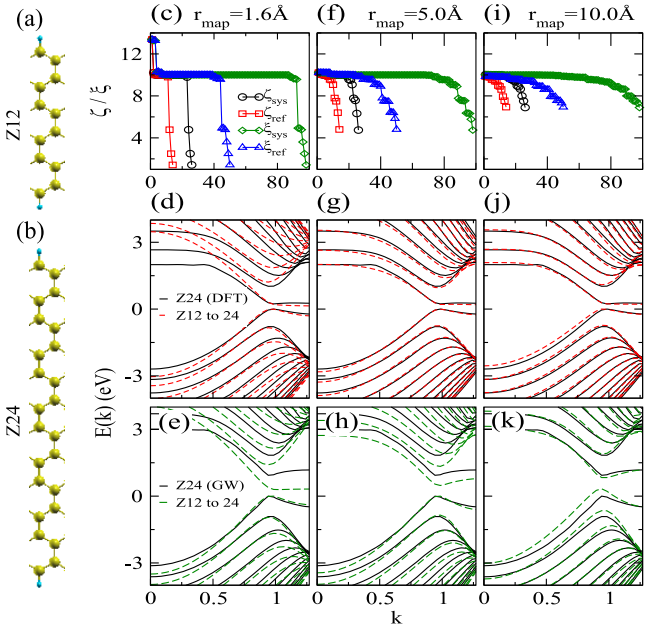


Figure 11. (a, b) Zigzag graphene nanoribbons (ZGNR, Z), Z12 and Z24, respectively. (c, f, i) Plot of ζ and ξ values "ref" (reference Z12) and "sys" (target Z24) for different spatial ranges of neighborhood considered for mapping. (d, g, j) Matching of DFT band structure and mapped TB band structure for increasing r_{map}. (e, h, k) Matching of DFT+G_0W_0 band structure and mapped self-energy-corrected TB band structure for increasing r_{map}.

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

to the degree of localization of the states at the band edges: the more they are delocalized, the larger is the spatial range within which the self-energy correction to TB parameters is to be considered.

3.A.2. Nanodiamonds. Figure 12 suggests that the extent of SEC to TB parameters is spatially limited mostly within the third nearest neighborhood, implying possible transferability of self-energy-corrected TB parameters to large covalent systems from smaller reference systems which are large enough to accommodate the full spatial range of non-nominal SEC to TB parameters. Accordingly, mapping in nanodiamonds is demonstrated with C_3H_8 and $C_{10}H_{16}$ (adamantane) as reference systems to map to nanodiamonds $C_{26}H_{32}$ (pentamantane) and $C_{84}H_{64}$.

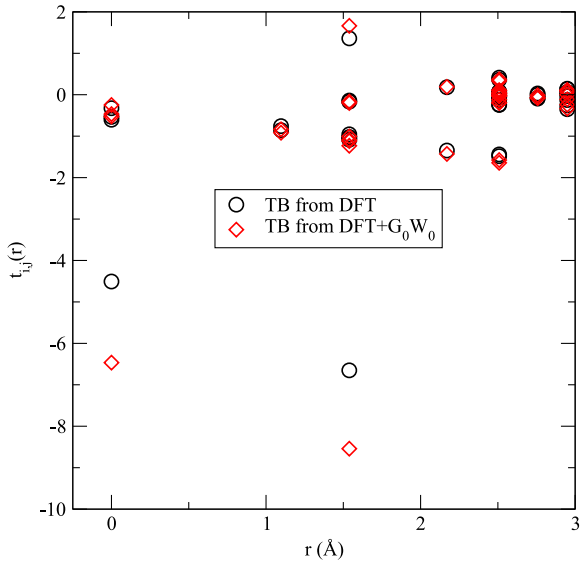


Figure 12. TB parameters involving a C atom in $C_{10}H_{16}$ with three C neighbors, computed using 56 KS states with and without SEC at the G_0W_0 level, and plotted as a function of distance from the atom. TB parameters from DFT are same as those plotted in Figure 7e.

We start with attempts to map $C_{10}H_{16}$, $C_{26}H_{32}$, and $C_{84}H_{64}$ targets from C_3H_8 reference in sp^3 HAO basis. The mapping process starts with plotting the distance of atom pairs (C−H, C−C, H−H) for target and reference systems. As seen in Figure 13a, there is a one-to-one correspondence of C−C bonds between C_3H_8 and all targets up to approximately 2.5 Å, which is the second nearest C−C distance. For C−H and H−H pairs, such correspondence exists up to about 3 and 3.75 Å, respectively. These correspondences decide the range of hopping parameters to be mapped. Notably, C_3H_8 has two varieties of C atoms—one with two (two) C (H) neighbors and the other with one (three) C (H) nearest neighbors—whereas $C_{10}H_{16}$ has C atoms with three (one) C (H) neighbors and two (two) C (H) neighbors. Additionally, $C_{26}H_{32}$ and $C_{84}H_{64}$ have C atoms with all C nearest neighbors (nn). An exact match of ζ among all atoms of reference and target systems is thus impossible in these examples. Matching ζ and ξ will therefore be less effective in mapping from C_3H_8. Also, since there is only one C atom with two (two) C (H) neighbors in C_3H_8, matching ζ can be restrictive in terms of the variety of orientations. We thus opt for matching structural parameters within a tolerance for ζ set to the minimum difference of ζ values between similar types of atoms in reference and target systems to ensure maximal matching of ζ besides finer matching of structural parameters. As is obvious, a better choice of reference system than C_3H_8 with C atoms having all varieties of neighborhoods can be easily made. However, we deliberately chose to test mapping from C_3H_8, which is the smallest possible reference system with just one C atom with two (two) C (H) neighbors, since such C atoms dominate the surfaces of the nanodiamonds and are thereby expected to host the states at the edges of the valence and conduction bands. Surprisingly, as is evident in Figure 13c, with mapping of only the nn-hopping terms from C_3H_8 to $C_{10}H_{16}$, the mapped TB DOS already matches reasonable well

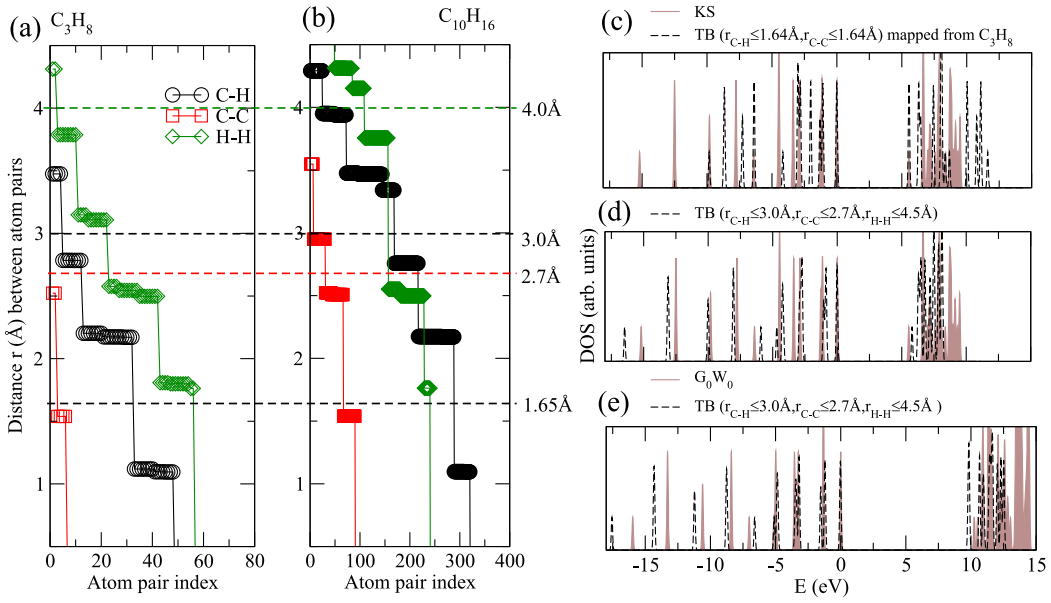


Figure 13. Distribution of distance between pairs of atoms in (a) reference (C_3H_8) and (b) target ($C_{10}H_{16}$) systems. (c, d) Match between DFT DOS and mapped TB DOS as demonstration of efficacy of mapping of TB parameters from C_3H_8 to $C_{10}H_{16}$ with increasing spatial range of neighborhood considered for mapping. (e) Match between DFT+G_0W_0 DOS and mapped self-energy-corrected TB DOS.

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

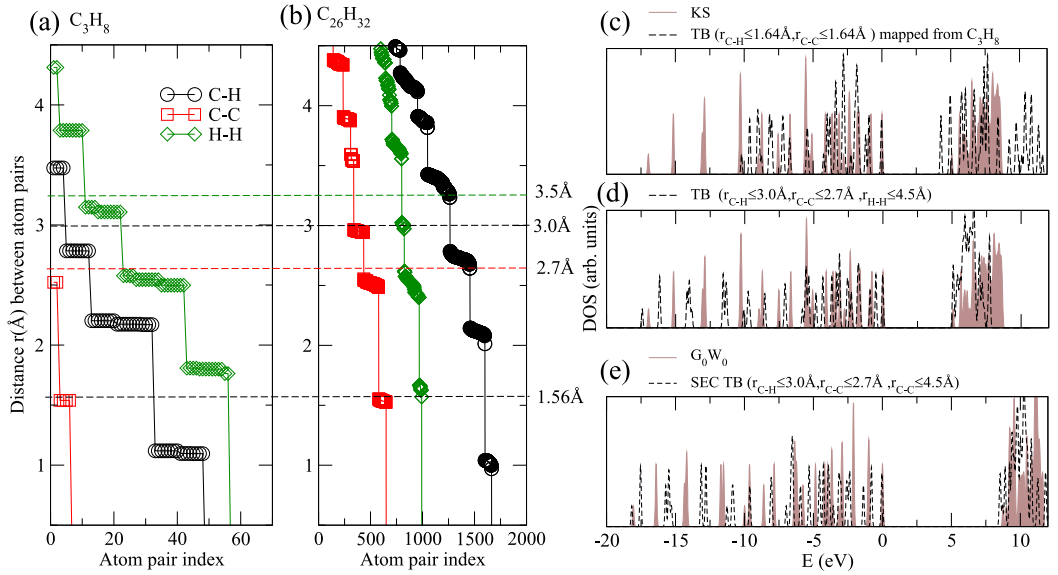


Figure 14. Distribution of distance between pairs of atoms in (a) reference (C_3H_8) and (b) target ($C_{26}H_{32}$) systems. (c, d) Match between DFT DOS and mapped TB DOS as with an increasing spatial range of neighborhood considered for mapping. (e) Match between DFT+G_0W_0 DOS and mapped self-energy-corrected TB DOS.

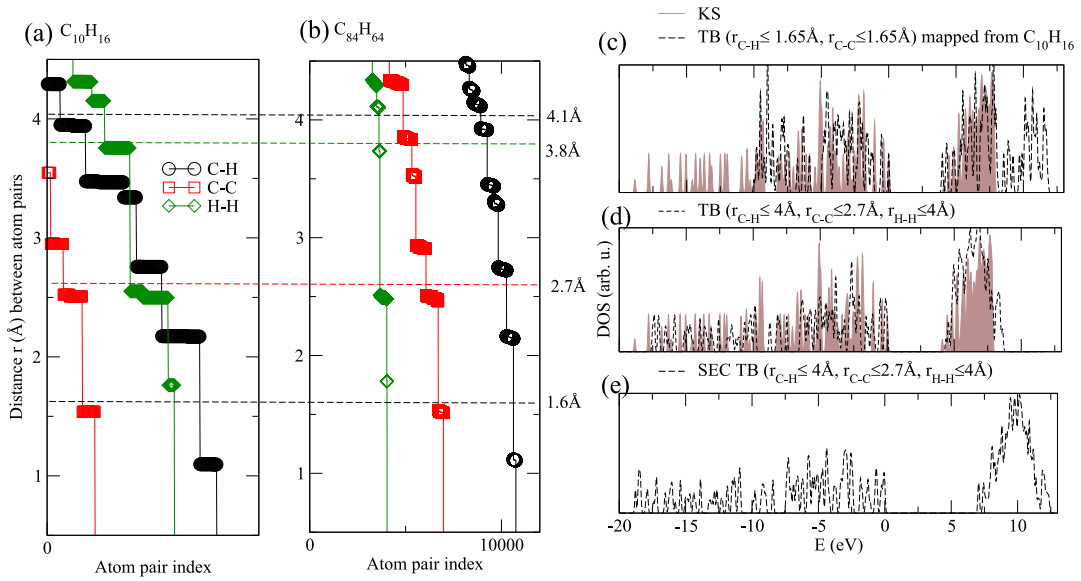


Figure 15. Distribution of distance between pairs of atoms in (a) reference ($C_{10}H_{16}$) and (b) target ($C_{84}H_{64}$) systems. (c, d) Match between DFT DOS and mapped TB DOS as with an increasing spatial range of neighborhood considered for mapping. (e) Match between DFT+G_0W_0 DOS and mapped self-energy-corrected TB DOS.

with the DFT DOS of $C_{10}H_{16}$ in terms of the band gap and the DOS around band edges. With an increase in the range of hopping to 2.7 Å (nn, 2n), 3 Å (nn, 2n, 3n), and 4 Å (2n, 3n, 4n)) for C−C, C−H, and H−H pairs based on the availability of one-to-one mapping (Figure 13a,b), the match of mapped TB DOS and DFT DOS (Figure 13d) extends deeper into the valence band. The quality of match improves further with additional mapping of C−H and H−H atom pairs up to 4.5 Å (Figure 13d) without compromising on tolerance factors. Notably, the ranges of hopping of C−H and H−H, although more than that of C−C, are actually consistent with the range of C−C hopping, since the farthest H atoms considered are associated with two second nearest C atoms. The same mapping parameters are then used to map self-energy-corrected TB (SEC-TB) parameters of C_3H_8 to $C_{10}H_{16}$, leading to a good match of not only the SEC-TB mapped band gap and the QP band gap calculated at the G_0W_0 level, but also the self-energy-corrected DOS of the valence band (Figure 13e). Next, we attempt mapping $C_{26}H_{32}$ from smaller references, starting with mapping from C_3H_8 to $C_{26}H_{32}$, which is an about 5 times increase in system size. Mapping of only nearest neighbor C−C and C−H hopping underestimates the band gap by about 15% (Figure 14c). Mapping all hopping parameters up to 4.5 Å, which is the maximum range of hopping available in the reference, drastically improves the overall match of not only mapped TB DOS and DFT DOS (Figure 14d) but also mapped SEC-TB DOS and DFT+G_0W_0 DOS (Figure 14e), as is seen in the case of mapping $C_{10}H_{16}$ from C_3H_8.

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

Finally, we demonstrate mapping to $C_{84}H_{64}$ from $C_{10}H_{16}$, which is an about 6 times enhancement in system size. Mapping of only the nearest neighbor C−C and C−H bonds results in a good match of the mapped TB band gap (Figure 15c) with the DFT band gap. With further mapping of hopping parameters up to 2.75 Å (nn, 2n), 4 Å (nn, 2n+), and 4 Å (2n, 3n+) (Figure 15a,b) for C−C, C−H, and H−H pairs, a satisfactory match of the entire valence band and a good match (Figure 15d) of the band gap is achieved. Mapping of the SEC of TB parameters from $C_{10}H_{16}$ to $C_{84}H_{64}$ results in a QP band gap of about 7.2 eV, which is within 5% deviation from the QP band gap implied in the literature.[69−72]

In Figure 16 we show similar mapping of TB parameters at the DFT and DFT+G_0W_0 levels for Si based nanodiamonds. As

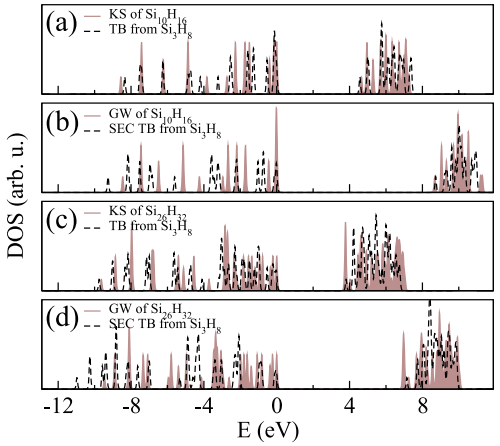


Figure 16. (a, c) Match between DFT DOS and TB DOS with parameters mapped from Si_3H_8. (b, d) Match between DFT+G_0W_0 DOS and SEC-TB DOS using mapped self-energy-corrected TB parameters from Si_3H_8.

in the case of nanodiamonds, mapping of hopping up to second nearest Si neighbors and H atoms associated with them from Si_3H_8 renders a good match of the SEC-TB band gap with the explicitly estimated DFT+G_0W_0 band gap almost up to 6 times escalation of the system size. These results imply consistency in transferability of self-energy-corrected TB parameters with increasing principal quantum number of valence orbitals.

4. CONCLUSION

In conclusion, construction of naturalized hybrid atomic orbitals (HAOs) is proposed as the common eigenstates of the noncommuting set of finite first moment matrices corresponding to the orthogonal directions. Hybridization and orientations of HAOs are numerically naturalized as per their anticipated immediate atomic neighborhood. Choice of gauge based on the HAOs leads to the construction of the hybrid atomic Wannier orbitals (HAWOs) from Kohn−Sham (KS) single particle states, rendering a multiorbital orthonormal tight binding (TB) basis locked to the nearest neighborhood. HAWO basis allows calculation of single TB parameters per bond from first principles and facilitates their easy transfer across isostructural systems through mapping of immediate atomic neighborhoods and projection of charge centers learned in the process of naturalization of the HAOs. The mapping allows effective bottom-up transfer of self-energy-corrected TB parameters estimated within the GW approximation of many-body perturbation theory in the HAWO basis, from smaller reference systems to much larger target systems having similar covalent atomic neighborhoods, suggesting a possible route toward computationally inexpensive estimation of quasi-particle structures of large covalent systems within an acceptable range of accuracy, with extra computational cost scaling as N^2, beyond the explicit computation of self-energy correction for smaller reference systems which typically scale as N^4. Demonstrated in nanoribbons and nanodiamond systems, the transferability of self-energy-corrected multiorbital TB parameters in the HAWO basis is rooted at the spatial localization of the extent of self-energy correction predominantly within the third nearest neighborhood, which appears to be robust for σ bonds but less so with π bonds and unpaired electrons.

AUTHOR INFORMATION

Corresponding Author

Joydeep Bhattacharjee − *National Institute of Science Education and Research, Homi Bhaba National Institute, Bhubaneswar 752050 Odisha, India*; orcid.org/0000-0001-6919-1931; Email: jbhattacharjee@niser.ac.in

Authors

Manoar Hossain − *National Institute of Science Education and Research, Homi Bhaba National Institute, Bhubaneswar 752050 Odisha, India*

Joydev De − *National Institute of Science Education and Research, Homi Bhaba National Institute, Bhubaneswar 752050 Odisha, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.1c00320

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

All computations have been performed in computing clusters supported by the Nanomission of the Department of Science & Technology and Department of Atomic Energy of the Government of India.

REFERENCES

(1) Pauling, L. The nature of the chemical bond. Application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. *J. Am. Chem. Soc.* **1931**, *53*, 1367−1400.

(2) Slater, J. C. Directed valence in polyatomic molecules. *Phys. Rev.* **1931**, *37*, 481.

(3) Doggett, G. Excited electronic states of diamond. *Proc. Phys. Soc., London* **1965**, *86*, 393.

(4) Del Re, G. Hybridization and localization in the tight-binding approximation. *Theor. Chim. Acta* **1963**, *1*, 188−197.

(5) McWeeny, R. Some recent advances in density matrix theory. *Rev. Mod. Phys.* **1960**, *32*, 335.

(6) Del Re, G.; Esposito, U.; Carpentieri, M. Bent bonds, hybridization, and the maximum localization criterion. *Theor. Chim. Acta* **1966**, *6*, 36−44.

(7) Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **1988**, *88*, 899−926.

(8) McWeeny, R.; Del Re, G. Criteria for bond orbitals and optimum hybrids. *Theor. Chim. Acta* **1968**, *10*, 13−22.

(9) Murrell, J. Construction of hybrid orbitals. *J. Chem. Phys.* **1960**, *32*, 767−770.
(10) Mayer, I. Atomic orbitals from molecular wave functions: The effective minimal basis. *J. Phys. Chem.* **1996**, *100*, 6249−6257.
(11) Mayer, I. Non-orthogonal localized orbitals and orthogonal atomic hybrids derived from Mulliken's population analysis. *Chem. Phys. Lett.* **1995**, *242*, 499−506.
(12) Foster, a. J.; Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.
(13) Slater, J. C. Atomic shielding constants. *Phys. Rev.* **1930**, *36*, 57.
(14) Hehre, W. J.; Stewart, R. F.; Pople, J. A. self-consistent molecular-orbital methods. i. use of gaussian expansions of Slater-type atomic orbitals. *J. Chem. Phys.* **1969**, *51*, 2657−2664.
(15) Dunning, T. H., Jr Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(16) Zhan, C.-G.; Hu, Z.-M. Maximum bond order hybrid orbitals. *Theoretica chimica acta* **1993**, *84*, 511−520.
(17) West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. A comprehensive analysis of molecule-intrinsic quasi-atomic, bonding, and correlating orbitals. I. Hartree-Fock wave functions. *J. Chem. Phys.* **2013**, *139*, 234107.
(18) Dupuis, M.; Nallapu, M. Maximal orbital analysis of molecular wavefunctions. *J. Comput. Chem.* **2019**, *40*, 39−50.
(19) Rives, A.; Weinhold, F. Natural hybrid orbitals: Ab initio SCF and CI results for CO and NiCO. *Int. J. Quantum Chem.* **1980**, *18*, 201−209.
(20) Gopinathan, M. Determination of atomic hybridization in molecular orbital theory: A valency method. *J. Mol. Struct.: THEOCHEM* **1988**, *169*, 379−388.
(21) Baxter, C. A.; Cook, D. B. Molecular fragments and the hybrid basis. *Int. J. Quantum Chem.* **1996**, *60*, 173−183.
(22) Kirkwood, E.; Cook, D. Generalized hybrid orbitals. *Theoretica chimica acta* **1977**, *44*, 139−149.
(23) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.
(24) Pu, J.; Gao, J.; Truhlar, D. G. Generalized hybrid orbital (GHO) method for combining ab initio Hartree- Fock wave functions with molecular mechanics. *J. Phys. Chem. A* **2004**, *108*, 632−650.
(25) Popov, I.; Tchougreeff, A.; Dronskowski, R. Deductive molecular mechanics of carbon allotropes. *Low Temp. Phys.* **2020**, *46*, 655−670.
(26) Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. A* **1984**, *392*, 45−57.
(27) Resta, R. Macroscopic polarization in crystalline dielectrics: the geometric phase approach. *Rev. Mod. Phys.* **1994**, *66*, 899.
(28) King-Smith, R.; Vanderbilt, D. Theory of polarization of crystalline solids. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47*, 1651.
(29) Marzari, N.; Vanderbilt, D. Maximally localized generalized Wannier functions for composite energy bands. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1997**, *56*, 12847.
(30) Bhattacharjee, J.; Waghmare, U. V. Localized orbital description of electronic structures of extended periodic metals, insulators, and confined systems: Density functional theory calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2006**, *73*, 121102.
(31) Cardoso, J.-F.; Souloumiac, A. Jacobi angles for simultaneous diagonalization. *SIAM journal on matrix analysis and applications* **1996**, *17*, 161−164.
(32) Goldstine, H. H.; Murray, F. J.; Von Neumann, J. The Jacobi method for real symmetric matrices. *J. Assoc. Comput. Mach.* **1959**, *6*, 59−96.
(33) Foster, J.; Boys, S. Canonical configurational interaction procedure. *Rev. Mod. Phys.* **1960**, *32*, 300.
(34) Tatewaki, H.; Huzinaga, S. A systematic preparation of new contracted Gaussian type orbital set. I. Transition metal atoms from Sc to Zn. *J. Chem. Phys.* **1979**, *71*, 4339−4348.
(35) King, R. B. Atomic orbitals, symmetry, and coordination polyhedra. *Coord. Chem. Rev.* **2000**, *197*, 141−168.
(36) Nacbar, D. R.; Ribeiro, A. V.; Bruno-Alfonso, A. A simple geometrical path towards hybrid orbitals. *Mater. Res.* **2014**, *17*, 1474−1476.
(37) Singh, D. D. J.; Pradeep, T.; Bhattacharjee, J.; Waghmare, U. Novel cage clusters of MoS2 in the gas phase. *J. Phys. Chem. A* **2005**, *109*, 7339−7342.
(38) Bhattacharjee, J. Activation of graphenic carbon due to substitutional doping by nitrogen: Mechanistic understanding from first principles. *J. Phys. Chem. Lett.* **2015**, *6*, 1653−1660.
(39) Maji, R.; Bhattacharjee, J. Synergistic View of Magnetism, Chemical Activation, and Oxygen Reduction Reaction as Well as Oxygen Evolution Reaction Catalysis of Carbon-Doped Hexagonal Boron Nitride from First Principles. *J. Phys. Chem. C* **2019**, *123*, 16731−16740.
(40) Löwdin, P.-O. 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.
(41) Schmidt, M. W.; Hull, E. A.; Windus, T. L. Valence virtual orbitals: An unambiguous ab initio quantification of the LUMO concept. *J. Phys. Chem. A* **2015**, *119*, 10408−10427.
(42) Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. Maximally localized Wannier functions: Theory and applications. *Rev. Mod. Phys.* **2012**, *84*, 1419.
(43) Lee, Y.-S.; Nardelli, M. B.; Marzari, N. Band structure and quantum conductance of nanostructures from maximally localized Wannier functions: the case of functionalized carbon nanotubes. *Phys. Rev. Lett.* **2005**, *95*, 076804.
(44) Gresch, D.; Wu, Q.; Winkler, G. W.; Häuselmann, R.; Troyer, M.; Soluyanov, A. A. Automated construction of symmetrized Wannier-like tight-binding models from ab initio calculations. *Physical Review Materials* **2018**, *2*, 103805.
(45) Calzolari, A.; Marzari, N.; Souza, I.; Buongiorno Nardelli, M. Ab initio transport properties of nanostructures from maximally localized Wannier functions. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2004**, *69*, 035108.
(46) Franchini, C.; Kováčik, R.; Marsman, M.; Sathyanarayana Murthy, S.; He, J.; Ederer, C.; Kresse, G. Maximally localized Wannier functions in LaMnO3 within PBE+ U, hybrid functionals and partially self-consistent GW: an efficient route to construct ab initio tight-binding parameters for eg perovskites. *J. Phys.: Condens. Matter* **2012**, *24*, 235602.
(47) Jung, J.; MacDonald, A. H. Tight-binding model for graphene π-bands from maximally localized Wannier functions. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *87*, 195450.
(48) Qian, X.; Li, J.; Qi, L.; Wang, C.-Z.; Chan, T.-L.; Yao, Y.-X.; Ho, K.-M.; Yip, S. Quasiatomic orbitals for ab initio tight-binding analysis. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *78*, 245112.
(49) Qian, X.; Li, J.; Yip, S. Calculating phase-coherent quantum transport in nanoelectronics with ab initio quasiatomic orbital basis set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2010**, *82*, 195442.
(50) D'Amico, P.; Agapito, L.; Catellani, A.; Ruini, A.; Curtarolo, S.; Fornari, M.; Nardelli, M. B.; Calzolari, A. Accurate ab initio tight-binding Hamiltonians: effective tools for electronic transport and optical spectroscopy from first principles. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *94*, 165166.
(51) Agapito, L. A.; Fornari, M.; Ceresoli, D.; Ferretti, A.; Curtarolo, S.; Nardelli, M. B. Accurate tight-binding Hamiltonians for two-dimensional and layered materials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *93*, 125137.
(52) Yue, S.-Y.; Qin, G.; Zhang, X.; Sheng, X.; Su, G.; Hu, M. Thermal transport in novel carbon allotropes with s p 2 or s p 3 hybridization: An ab initio study. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *95*, 085207.
(53) Popov, I. V.; Slavin, V. V.; Tchougréeff, A. L.; Dronskowski, R. Deductive molecular mechanics of four-coordinated carbon allotropes. *Phys. Chem. Chem. Phys.* **2019**, *21*, 18138−18148.

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817

(54) Wiberg, K. B. Bent bonds in organic compounds. *Acc. Chem. Res.* **1996**, *29*, 229−234.
(55) Hedin, L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **1965**, *139*, A796.
(56) Hedin, L.; Lundqvist, S. *Solid State Physics*; Elsevier: 1970; Vol. *23*, pp 1−181..
(57) Hybertsen, M. S.; Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *34*, 5390.
(58) Rudenko, A. N.; Katsnelson, M. I. Quasiparticle band structure and tight-binding model for single-and bilayer black phosphorus. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *89*, 201408.
(59) Aguilera, I.; Friedrich, C.; Blügel, S. Many-body corrected tight-binding Hamiltonians for an accurate quasiparticle description of topological insulators of the Bi 2 Se 3 family. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2019**, *100*, 155147.
(60) Yu, J.; Katsnelson, M. I.; Yuan, S. Tunable electronic and magneto-optical properties of monolayer arsenene: From G W 0 approximation to large-scale tight-binding propagation simulations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2018**, *98*, 115117.
(61) Grüneis, A.; Attaccalite, C.; Wirtz, L.; Shiozawa, H.; Saito, R.; Pichler, T.; Rubio, A. Tight-binding description of the quasiparticle dispersion of graphite and few-layer graphene. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *78*, 205425.
(62) Cho, Y.; Berkelbach, T. C. Optical properties of layered hybrid organic−inorganic halide perovskites: A tight-binding GW-BSE study. *J. Phys. Chem. Lett.* **2019**, *10*, 6189−6196.
(63) Sawamura, A.; Otsuka, J.; Kato, T.; Kotani, T. Nearest-neighbor sp3s* tight-binding parameters based on the hybrid quasi-particle self-consistent GW method verified by modeling of type-II superlattices. *J. Appl. Phys.* **2017**, *121*, 235704.
(64) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.
(65) Perdew, J. P.; Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1981**, *23*, 5048.
(66) Deslippe, J.; Samsonidze, G.; Strubbe, D. A.; Jain, M.; Cohen, M. L.; Louie, S. G. BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. *Comput. Phys. Commun.* **2012**, *183*, 1269−1289.
(67) Yang, L.; Park, C.-H.; Son, Y.-W.; Cohen, M. L.; Louie, S. G. Quasiparticle energies and band gaps in graphene nanoribbons. *Phys. Rev. Lett.* **2007**, *99*, 186801.
(68) Hossain, M.; Bhattacharjee, J. Transferability of self-energy correction in tight-binding basis constructed from first principles. *J. Chem. Phys.* **2020**, *153*, 144103.
(69) Raty, J.-Y.; Galli, G. Optical properties and structure of nanodiamonds. *J. Electroanal. Chem.* **2005**, *584*, 9−12.
(70) Raty, J.-Y.; Galli, G.; Bostedt, C.; Van Buuren, T. W.; Terminello, L. J. Quantum confinement and fullerenelike surface reconstructions in nanodiamonds. *Phys. Rev. Lett.* **2003**, *90*, 037401.
(71) Drummond, N.; Williamson, A.; Needs, R.; Galli, G. Electron emission from diamondoids: a diffusion quantum Monte Carlo study. *Phys. Rev. Lett.* **2005**, *95*, 096801.
(72) Sasagawa, T.; Shen, Z.-x. A route to tunable direct band-gap diamond devices: Electronic structures of nanodiamond crystals. *J. Appl. Phys.* **2008**, *104*, 073704.

https://doi.org/10.1021/acs.jpca.1c00320
J. Phys. Chem. A 2021, 125, 6805−6817