# A Comprehensive Analysis in Terms of Molecule-Intrinsic, Quasi-Atomic Orbitals. II. Strongly Correlated MCSCF Wave Functions

Aaron C. West, Michael W. Schmidt, Mark S. Gordon, and Klaus Ruedenberg*

Department of Chemistry and Ames Laboratory USDOE, Iowa State University, Ames, Iowa 50011, United States

ABSTRACT: A methodology is developed for the quantitative identification of the quasi-atomic orbitals that are embedded in a strongly correlated molecular wave function. The wave function is presumed to be generated from configurations in an internal orbital space whose dimension is equal to (or slightly larger) than that of the molecular minimal basis set. The quasi-atomic orbitals are found to have large overlaps with corresponding orbitals on the free atoms. They separate into bonding and nonbonding orbitals. From the bonding quasi-atomic orbitals, localized bonding and antibonding molecular orbitals are formed. The resolution of molecular density matrices in terms of these orbitals furnishes a basis for analyzing the interatomic bonding patterns in molecules and the changes in these bonding patterns along reaction paths. A new bond strength measure, the kinetic bond order, is introduced.

ELECTRONIC WAVE FUNCTION
↓
ORIENTED QUASI-ATOMIC ORBITALS
↓
FIRST ORDER DENSITY MATRIX
↓
KINETIC BOND ORDERS
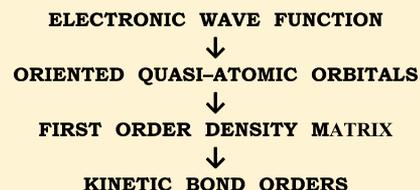
1. INTRODUCTION

The time-honored rationalizations of quantum chemical results in terms of interpretations drawn from a pool of intuitive qualitative notions are valuable and have the virtue of simplicity. But they also have limitations, and it has been recognized for some time that a more rigorous identification of the physical synergisms embedded in rigorous ab initio wave functions would be valuable. The development of rigorous methods that can achieve this goal—as distinct from methods for the computation of measurable quantities—is a fundamental part of chemical physics because it is relevant for a fruitful contact between rigorous theory and empirical chemical intuition.

The rigorous extraction of interpretative concepts from ab initio many-electron wave functions is sufficiently challenging that some computational quantum chemists regard it as unattainable, in particular as regards the concept of an atom in a molecule. This presumption shows that, although the goal seems simple, some mathematical complexity is to be expected of a formalism that generates an analysis for unraveling the various physical interactions that are inherent in molecular electronic wave functions.

In the present approach, the first step consists of identifying, for a given molecular wave function, *orbitals that have the character of chemically deformed atomic orbitals and are also capable of generating configurations from which the molecular wave function can be reconstituted*. A comprehensive general method of identifying such quasi-atomic orbitals (QUAOs) for Hartree–Fock (HF) wave functions was developed and elaborated in detail in the first paper[1] of the present series, which will be referred to as the preceding paper or as paper I. That paper also contained a review of prior work on related problems. Other closely related work is discussed below in Section 2.4.

In the present study, the approach is generalized from HF wave functions to strongly correlated wave functions. In the second section, the relation to the method of the preceding paper is discussed. In the third section, certain revisions of and additions to the method of the preceding paper are introduced. In the fourth section, the theory is generalized to several types of multiconfiguration self-consistent field (MCSCF) wave functions, in particular, full valence space MCSCF wave functions, reduced-valence-space MCSCF wave functions, and multiple active space MCSCF wave functions, which are particularly, although not only, relevant for reactions.

The qualitative and quantitative bonding information that can be deduced by this analysis is exemplified in the subsequent papers, namely, by an elucidation of the bonding structure of the urea molecule[2] and by exhibiting the changes in bonding pattern that occur along the reaction path of the dissociation of dioxetane into formaldehyde.[3]

2. RELATION TO THE ANALYSIS OF HARTREE–FOCK WAVE FUNCTIONS OF THE PRECEDING PAPER

2.1. Internal Orbital Space and Strong Correlation. Electronic structure descriptions that account for electron correlation are based on wave functions that include correlating configurations in addition to the HF self-consistent-field (SCF) approximation. It is generally recognized and elaborated in the preceding paper[1] that correlated wave functions typically consist of a strongly correlated part and a weakly correlating part (also distinguished as nondynamic and dynamic correlations.) The *strongly* correlated part is generated from a set of *internal* orbitals. The *weakly* correlating part involves, in addition, excitations into *external* orbitals.

The internal orbital space, which generates the strong correlations, must be chosen judiciously. In the first three periods of the periodic table, notably for most organic molecules, the internal space is typically spanned by orbitals

in the *full valence space*. This statement implies that full orbital *localization* in the internal orbital space yields some quasi-atomic orbital basis, that is to say, orbitals having the character of deformed minimal basis set (MBS) orbitals of the free atoms.[4−11] In this case, the dimension of the internal space is thus typically equal to the *total number of MBS orbitals*. There are molecules (e.g., BF) for which the internal space contains some additional orbitals outside the minimal basis model, for instance, a correlating orbital for certain lone pair orbitals. Such internal spaces will be considered as extended minimal basis set (MBS+) spaces.

A reason that the orbitals of the internal space contribute much more strongly to molecular wave function than the external orbitals is that the interactions between the quasi-atomic minimal basis orbitals are dominated by the one-electron kinetic energy lowerings that create bonds between atoms, and only to a lesser degree account for interelectronic correlations. The major goal of the present analysis is to generate orbitals in the internal space that are useful for identifying both the bonding interactions and the strong correlation interactions.

2.2. Strongly and Weakly Occupied Internal Orbitals. The number of occupied HF orbitals is in general less than the dimension of the internal space, i.e., the total number of MBS orbitals. Nonetheless, it is possible to extract from the HF unoccupied virtual orbital space a set of orbitals, viz., *the valence virtual orbitals (VVOs)*, that, together with the occupied HF orbitals, *span an orbital space that is a very good approximation to the internal space* (the "HF internal space"). This extraction can be accomplished by exploiting the physical fact that *the internal space can be spanned by an orbital basis that has quasi-atomic character.* Thus, Lu, Wang, Schmidt, Bytautas, Ho, and Ruedenberg[12−14] developed a method to determine VVOs and construct HF-QUAOs based on the following criteria: The space spanned by the QUAOs

  (i) is constrained to contain the space spanned by the occupied *molecular* HF orbitals of the molecule;
  (ii) deviates as little as possible from the space spanned by a set of accurate minimal basis set orbitals of the free atoms.

The method was slightly refined in the preceding paper.[1] Typically, these HF internal spaces have been found[15] to provide more effective initial approximations for MCSCF calculations in the full valence space than other previously proposed[16−20] "improved virtual orbitals".

The occupied HF orbitals and the VVOs together provide an orbital basis for the HF internal space with occupations of 2 or 0, respectively. For the internal space of a *correlated* wave function, there exists a corresponding orbital basis of *strongly and weakly occupied internal orbitals*, with respective occupations of *close* to 2 (or 1 in case of an open-shell system) on the one hand, and typically less than 0.4 and often much smaller on the other hand. These orbitals can be identified, for example, by determining natural (or canonicalized[21]) orbitals in the internal space. By virtue of the strong resemblance between the HF internal space and the internal space of a correlated wave function, the analyses of the internal spaces of HF wave functions and of correlated wave functions follow very similar patterns.

In correlated wave functions, the number of occupied orbitals is greater than in the HF case. There exist three possibilities: The number of occupied orbitals is (i) still less than, (ii) equal

to, or (iii) greater than the dimension of the internal space. In the following, the identification of the internal orbital space will be considered for all three cases. Once the internal space of a correlated wave function is determined, the orbital analysis in its internal space follows the same general pattern as that described for HF functions in the preceding paper.[1] The essential difference to the HF analysis of paper I is thus the determination of the weakly occupied internal orbitals: in the HF case, these orbitals are the VVOs (with occupations = 0); in the present correlated case, some or all weakly occupied internal orbitals are in fact optimized orbitals in the correlated wave function.

2.3. Overview of the Sequence of Orbital Constructions. Figure 1 outlines the general scheme of the analysis. The
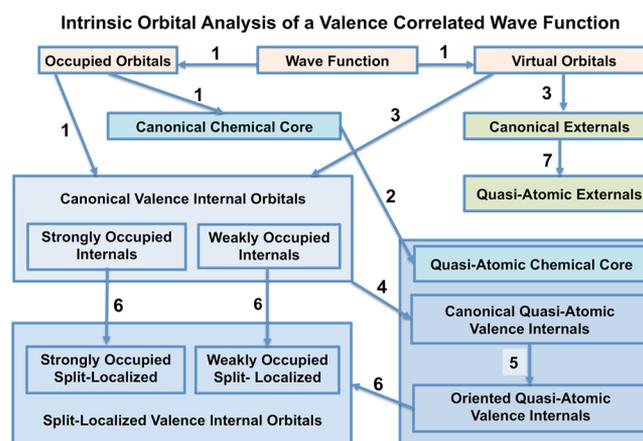
Figure 1. Overview of molecule-intrinsic orbital sets and the sequence of their determination.

numbers in the figure indicate the sequence of the steps. These steps are analogous to the steps in Figure 1 of the preceding paper[1] and correspond to the following operations.

1. The molecular orbitals are separated into the chemical core, the occupied valence orbitals, and the virtual orbitals.
2. The QUAO basis for the chemical core space is determined.
3.
   a. If the number of occupied orbitals is less than the dimension of the internal space, the required number of VVOs is extracted from the virtual MCSCF orbital space in order to obtain a full internal orbital basis, as is done in the HF case.
   b. If the number of occupied orbitals is equal to the dimension of the internal space, the preceding step is skipped.
   c. If the number of occupied valence orbitals is larger than the dimension of the internal space, the internal space is identified as a subspace of the occupied orbital space.
   d. In the following discussions, the default choice for the dimension of the internal space is taken to be the number of MBS orbitals. This choice is assumed unless explicitly stated otherwise.
4. The canonical internal orbital basis is determined, and from it, the *canonical quasi-atomic* valence MBS is generated by localization. The population-bond-order matrix is constructed.

5. On each atom, the canonical QUAOs are "hybridized" so that they are oriented and exhibit the bonding patterns. The population-bond-order matrix is updated in terms of these *oriented QUAOs*. The oriented QUAOs are also used to accomplish the next step.

6. In the internal space, the strongly occupied subspace and the weakly occupied subspace are determined. Localized molecular orbitals are then determined separately in the strongly occupied internal subspace and in the weakly occupied internal subspace. Together, they provide the split-localized basis of the valence-internal space.

7. In the valence-external space, that is, the orbital space that is orthogonal to the internal orbital space, a quasi-atomic basis is determined.

The basic techniques that are required to implement the various steps in this analysis of strongly correlated wave functions are the same as those developed in the preceding paper for the analysis of HF wave functions, subject to the modifications discussed below in Section 3. The previously developed techniques will be referenced where appropriate, but they will not be rederived.

2.4. Notes on Related Work. Some of the above-mentioned methods antedate the preceding paper.[1] Thus, as already noted, the QUAOs used here are only slight modifications of the QUAMBOs that were introduced[12−14,22] in 2004. However, the localization method formulated in the preceding paper[23] to generate the split-localized orbitals by reference to the QUAOs goes back to a method[24] of 1982, seven years before the Pipek−Mezey localization,[25] which is somewhat related in spirit.

Interesting contributions have recently been made by Knizia.[26] He has independently devised an algorithm consisting of a sequence of projections and matrix inversions involving only the occupied HF orbitals. This algorithm generates a set of orbitals that Knizia calls "intrinsic atomic orbitals". By establishing the formal mathematical equivalence as well as by quantitative evaluations, Janowski[27] has subsequently shown that these intrinsic atomic orbitals span the same space as the QUAOs of ref 1 on the same atom. Knizia furthermore developed[26] an adaptation of the Pipek and Mezey[25] localization method to QUAOs, which made this method very effective in the context of large basis sets. This approach[26] is independent of the earlier localization method[1,24] referred to in the preceding paragraph, which is used in the present study. The Pipek−Mezey−Knizia scheme maximizes an expression in terms of total atomic populations, which are invariant under intra-atomic orbital mixing. In the method used here, the individual *oriented* QUAOs enter explicitly, so that the optimization is sensitive to intra-atomic orbital mixing.

3. MODIFICATIONS OF AND ADDITIONS TO THE PRECEDING PAPER

In light of our experiences with the methods outlined in the preceding paper[1] for HF wave functions, several modifications and additions were made in that analysis, which are also used in the approach discussed here.

3.1. Orthogonalization of Quasi-Atomic Orbitals in the Internal Space. The first modification pertains to the construction of the internal QUAOs in Step 4 of Figure 1. These QUAOs are obtained by two sequential operations.[28] First, preliminary QUAOs are determined separately for each atom by singular value decompositions. Since the resulting QUAOs of different atoms are not orthogonal to each other, they are then orthogonalized. In the previously described procedure, this is accomplished by using $S^{-1/2}$, that is, by symmetric orthogonalization using the overlap matrix, which preserves the character of the nonorthogonal QUAOs as much as possible. However, it seems more desirable to base this orthogonalization on a criterion that would *directly maximize the quasi-atomic character* of the orthogonal QUAOs. To that end, a new orthogonalization was formulated. It is described in the Appendix.

3.2. A New Weighted Orthogonalization of External Quasi-Atomic Orbitals. The second modification pertains to the orthogonalization of the external orbitals in Step 7 of Figure 1.[29] To avoid contamination of compact orbitals by diffuse orbitals, the interatomic orthogonalization of these orbitals was performed with the *weighted* Carlson−Keller algorithm.[30] This algorithm was found to break down for too wide a spread in the weights. A new stable weighted orthogonalization method was developed and has been published.[31]

3.3. Kinetic Bond Orders. The interatomic off-diagonal elements $p_{Aa,Bb}$ of the density matrix for QUAOs[32] provide measures of constructive QUAO interferences, which generate charge accumulations in bonds and, hence, of electron delocalization between the QUAOs.[33,34] [$|Aa\rangle$ denotes orbital a on atom A.] The $p_{Aa,Bb}$ are therefore related to covalent bonding and are denoted as bond orders.[35,36] If the phases of the QUAOs can be appropriately chosen, positive bond orders can be taken as indicators of covalent bonding interactions and negative bond orders as implying antibonding interactions. In larger molecules, the interpretation is, however, less straightforward for several reasons. It is difficult to control the phases of the QUAOs and, in some cases, constructive interference (bonding) is associated with negative bond orders. Another problem is that, in fact, the actual bonding interactions are given by the *products* of bond orders and certain energy integrals between the QUAOs. These energy integrals often decay much faster than the corresponding bond orders as the distance between the QUAOs increases.

To compensate for these shortcomings of the "population bond orders" $p_{Aa,Bb}$ between the QUAOs $|Aa\rangle$ and $|Bb\rangle$, the $p_{Aa,Bb}$ will be complemented by also considering the kinetic interference energies, which are defined as

$$t_{Aa,Bb} = p_{Aa,Bb} \left\langle Aa \left| -\frac{1}{2}\nabla^2 \right| Bb \right\rangle = p_{Aa,Bb} \langle \nabla Aa | \nabla Bb \rangle / 2 \tag{1}$$

In these expressions the accidental phases of the QUAOs manifestly cancel. In agreement with the general analysis of covalent binding,[33] the $t_{Aa,Bb}$ have consistently been found to be negative in all cases where the interactions between the QUAOs were manifestly bonding, regardless of the phases of the QUAOs. Moreover, these quantities $t_{Aa,Bb}$ also incorporate the distance dependence of the physical interactions.

Since the kinetic interference energies have been shown to be responsible for covalent bonding,[33] it is likely that the $t_{Aa,Bb}$ reflect bond strengths and can serve a similar purpose as "resonance energies". There are two caveats: (i) The $t_{Aa,Bb}$ represent energy lowerings between *orthogonalized* QUAOs, between which interference energies are markedly larger in magnitude than the corresponding energies would be between nonorthogonal QUAOs; (ii) The interference contributions of the *potential* energy, which are typically antibonding,[33] are not taken into account. In accordance with these considerations,

the $t_{Aa,Bb}$ values that have been found in organic molecules so far have consistently been an order of magnitude larger than the empirically inferred chemical bond energies. Notwithstanding these considerations, it is found in the detailed analyses documented in the two subsequent papers[2,3] that *nearly all kinetic bond energies, the strong ones as well as the weak ones, are negative, that is, bonding.* Provisionally, we therefore consider the empirically adjusted kinetic interference energies

$$k_{Aa,Bb} = 0.1 \times t_{Aa,Bb} = p_{Aa,Bb} \left\langle Aa \left| -\frac{1}{2}\nabla^2 \right| Bb \right\rangle / 10 \qquad (2)$$

as reasonable measures for the covalent bonding implications of the first order density matrix. These adjusted kinetic bond energies will be denoted as *kinetic bond orders* (KBOs). Future work will address the problem of resonance integrals in greater depth. The various kinds of known *noncovalent* interactions may call for different bonding measures.

3.4. Corrigenda to Paper I. Equations (A22a) and (A22b) in Paper I contain two misprints. In eq (22a), the term $2P_{1112}$ should read $2P_{1122}$. In the expression for P_c in eq (A22b), the term $(-6P_{1112})$ should read $(-6P_{1122})$.

4. ANALYSIS OF CORRELATED WAVE FUNCTIONS IN AN INTERNAL ORBITAL BASIS

4.1. Full Valence Internal MCSCF Wave Function (FORS). The conceptually simplest case is that of an MCSCF optimized wave function in the *full* valence space, i.e., a complete-active-space self-consistent-field (CASSCF) wave function with the number of active orbitals being exactly *equal to* the number of MBS orbitals. As in our original work on this subject,[4] we use the term full optimized reaction space (FORS) model for these specific CAS(N, MBS) wave functions, where N is the number of electrons. This type of wave function is the most general optimal wave function based on the use of all orbitals in the full internal space and only these orbitals.

The optimized orbitals of the FORS wave function, the FORS molecular orbitals (MOs), yield a full basis for the internal space so that no VVOs are needed. The natural orbitals (NOs) in the valence space provide a canonical basis in the FORS internal space, which orders the internal FORS NO basis according to the strengths of their occupations.

The determination of the QUAOs is accomplished by means of singular value decompositions.[37] The core QUAOs (Step 2 of Figure 1) are obtained exactly as in the HF case.[38] The determination of VVOs,[39] i.e., Step 3 of Figure 1, is not needed, as mentioned in the preceding paragraph, and is therefore skipped. In Step 4, the canonical FORS valence QUAOs are determined by a singular value adaptation[40] followed by an orthogonalization that is modified as described above in Section 3.1. Finally, in Step 5, the oriented QUAOs are determined by the orientation algorithm.[41] The representation of the first-order density matrix in terms of the oriented QUAOs exhibits the population distributions and the bond orders. The FORS wave function can be expressed in terms of QUAOs.

In Step 6, the split-localized orbital basis of the internal space is constructed.[42] To this end, the internal FORS NOs must be divided into a strongly occupied group and a weakly occupied group on the basis of their occupation numbers. This division must be made with judicious consideration of the bonding situation, in particular, when one is following changes in

electronic structure and bonding along a reaction path. The following criterion has been found effective:

(i) FORS NOs with occupations larger than 0.9 are taken to be strongly occupied.

(ii) For all orbitals with lower occupations, the largest *gap between two consecutive occupation numbers* is determined. The orbitals above this gap are then included in the strongly occupied subspace. The orbitals below this gap form the weakly occupied subspace. After the strongly and the weakly occupied FORS spaces have been identified, the split-localized orbitals are determined by maximizing the expression

$$\sum_n \sum_A \sum_a \langle Aa|\phi_n\rangle^4 \qquad (3)$$

separately in these two spaces.[43,44] The sum $\sum_A \sum_a$ in eq 3 runs over all oriented QUAOs $|Aa\rangle$ on all atoms, and the sum $\sum_n$ runs, respectively, either over all strongly occupied or over all weakly occupied split-localized orbitals ϕ_n. The FORS wave function can also be expressed in terms of split-localized orbitals.

As mentioned earlier, at the equilibrium geometry, the split-localized orbitals typically represent localized bonding and antibonding orbitals. In systems where the HF wave function is a reasonable approximation, the strongly occupied FORS orbital space is very close to the HF-occupied orbital space. The weakly occupied FORS orbital space has a somewhat weaker, but still strong, overlap with HF-VVO orbital space, as was exemplified in Table 2 of the preceding paper.[1] Because of this correspondence, the occupied HF orbitals and the HF-VVOs provide excellent first approximations for any iterative MCSCF procedure that determines the FORS orbitals.

The virtual FORS orbitals furnish a basis for the external space. Canonical and quasi-atomic external orbitals are obtained by the algorithm[45] described in Paper I, modified as discussed above in Section 3.2.

4.2. Wave Functions Based on Multiple Active Subspaces in a Full Internal Orbital Basis. *4.2.1. Wave Functions in Reduced Full Valence Spaces.* For full valence spaces, the lengths of the configuration interaction (CI) expansions increase very rapidly with the size of the molecular system so that full valence space calculations cannot be used even for molecules of moderate size. The search for reduced spaces that generate effective strong correlations is therefore an active research topic. In many instances, typically along reaction paths and notably for transition metals compounds, useful descriptions of the relevant electronic structures and energy changes have been successfully obtained through MCSCF wave functions that are formed in relevant subspaces of the entire full valence space of the molecular system. Such reduced full MCSCF calculations are widely used.

The GAMESS program suite[46,47] contains a very general MCSCF module that offers the options for calculations with a wide range of reduced full configuration space choices. This module allows for the optimization of wave functions in *occupation restricted multiple active spaces* (ORMAS),[48] which are constructed as follows: The orthogonal molecular orbitals of the system are divided into a number of ORMAS groups (OGs). For each OG a number of orbitals are chosen. The electron occupation of each OG can then vary between a minimum and a maximum number of electrons. The number of OGs, the number of orbitals in each OG, and the occupation

minima and maxima of the OGs can all be arbitrarily chosen. All configurations satisfying these constraints are taken into account in the CI calculation, and all orbitals are optimized.

If the *molecular orbital set* from which an ORMAS wave function is constructed is in fact the *full* basis of the internal space (i.e., if their number is equal to the total number of MBS orbitals), then the optimized orbitals of this wave function yield again a full basis for the internal space. As was the case for the FORS wave function in the preceding Section 4.1, there is no need to determine VVOs. The virtual orbitals of the ORMAS calculation span the external space in which canonical orbitals and QUAOs are determined as described for the FORS wave functions in Section 4.1.

4.2.2. Determination of Quasi-Atomic Orbitals. There exist, however, two ways of forming the valence QUAOs in the internal space, namely, whether MOs from different ORMAS groups are allowed to mix or not in the construction of the QUAOs:

Option 1: Unrestrained Quasi-Atomic Orbitals. In this case, the core QUAOs, the canonical and the oriented valence QUAOs, as well as the external orbitals, are determined in exactly the same manner as described in the preceding Section 4.1 for FORS wave functions.

Option 2: ORMAS-Restrained Quasi-Atomic Orbitals. In general, the QUAOs that result from the just-described unrestrained procedure are linear combinations that encompass MOs from *different* OGs. For some purposes however, it is desired to construct the QUAOs in such a way that they belong to individual OGs. For instance, this construction is required if one wants to express the wave function in terms of QUAOs under preservation of its original ORMAS format. The procedure to determine such QUAOs differs from that in the preceding Section 4.1 as follows.

The critical first step in determining QUAOs consists of singular value decompositions of the overlap matrices $\langle VIn | A^*a \rangle$ between the valence-internal *molecular* orbitals $|VIn\rangle$ and the accurate atomic minimal basis set orbitals $|A^*a\rangle$ of each individual atom.[49] In the case of the FORS wave functions of Section 4.1 (and in the procedure for unrestrained QUAOs), the molecular orbitals $|VIn\rangle$ in each overlap matrix $\langle VIn|A^*a \rangle$ include *all* internal molecular orbitals. In the ORMAS-restrained procedure, the MOs $|VIn\rangle$ are restricted to the molecular orbitals of *one* OG. Thus, for each of these overlap matrices, the corresponding SVD is performed. Thereby, a set of (positive) SVD eigenvalues is obtained *for each atom in each OG*. From the collection of all these SVD eigenvalues, the largest M values are chosen, where M is the *total* number of MBS orbitals *in the molecule*. This procedure has been found to assign the QUAOs appropriately to the atoms in each OG. The determination of the oriented QUAOs is then also performed separately in each OG on each atom.

4.2.3. Determination of Split-Localized Molecular Orbitals. Regardless of the choice of the QUAOs, the construction of the split-localized orbitals for ORMAS-type wave functions is slightly different from the construction in the case of the FORS wave functions discussed in the preceding Section 4.1. In that case, *every* split-localized orbital ϕ_n is a linear combination of *all* FORS MOs in the full internal space. Consequently, the sum $\sum_n$ in eq 3 runs, respectively, either over *all* strongly occupied or over *all* weakly occupied split-localized orbitals ϕ_n. Using this procedure in the ORMAS case would generate split-localized orbitals that contain components from *different* ORMAS groups so that they could not be used to reconstitute

the ORMAS wave function *in the original format*. Split-localized orbitals of this kind are not considered at present.

To preserve the original ORMAS wave function format in terms of the split-localized orbitals, the split-localized orbitals are therefore *generated separately for each ORMAS group*. Hence, the sum $\sum_n$ in eq 3 goes only over the (either strongly or weakly occupied) split-localized orbitals from one OG at a time so that each split-localized orbital becomes a linear combination of the MOs from that OG only. The sum $\sum_A \sum_a$, however, still goes over *all* oriented QUAOs $|Aa\rangle$ on *all* atoms. In terms of these split-localized orbitals, the ORMAS wave function can then be formulated in the original ORMAS format.

4.3. Wave Functions Based on Multiple Active Spaces Constructed from Partial Internal Orbital Bases.
Many problems of concrete chemical interest in larger molecules involve reactions within a limited region of the entire system. For such cases, wave functions are useful in which a full configuration space is generated that uses only the local orbitals in the reactive region of interest. In this context, it is typically useful to divide the molecular valence orbitals into active and inactive orbitals. For instance, the former may be involved in a reaction, while the latter are "spectators". In such a case, the valence part of the electronic wave function consists of a multiconfigurational active part to describe the reactive part of the system and an inactive spectator part of a set of closed-shell doubly occupied orbitals. The active part can have the form of a CAS-type or an ORMAS-type wave function. The set of all inactive orbitals can be taken as one OG with the number of electrons exactly equal to twice the number of orbitals.

The number of occupied molecular orbitals in such a wave function, though larger than in the HF case, is typically less than the total number of MBS orbitals. To generate a basis for the full internal space, one manifestly must find those valence virtual orbitals (VVOs) that complement the closed-shell spectator part of the wave function. These VVOs are determined by a procedure that is analogous to that developed for HF wave functions.[50] To this end, the singular value decomposition is performed for the overlap matrix $\langle Vn|A^{\#}a\rangle$ where the $|Vn\rangle$ run over *all virtual (unoccupied) orbitals of the MCSCF calculation* and the $|A^{\#}a\rangle$ run through all *orthogonalized* Accurate [free] Atom Minimal Basis Set (AAMBS) orbitals[51] on all atoms. The VVOs are then chosen as those singular vectors *in the virtual MCSCF space* that correspond to the N_{VVO} largest singular values, where N_{VVO} is now the *difference between the total number of minimal basis set orbitals and the number of occupied MOs in the MCSCF wave function*. These singular values are typically close to unity, whereas the remaining nonzero singular values are significantly smaller. The occupied MCSCF orbitals together with the VVOs obtained in this way provide then a basis that spans the internal orbital space.

For the calculation of the QUAOs, there exist again two options, as was the case in Section 4.2.2. The *unrestrained* QUAOs are obtained by allowing the mixture of *all* MOs of the internal space (i.e., all occupied MCSCF orbitals *plus all VVOs*) in the construction of *all* QUAOs. This option yields the most localized QUAOs.

On the other hand, if one wants to be able to express the wave function in terms of QUAOs under preservation of the original ORMAS format, then *restrained* QUAOs are required. They are now obtained by determining the QUAOs separately for each ORMAS group, as described above in Section 4.2.2, *except for the ORMAS group that contains all inactive occupied valence orbitals*. Expressing the occupied inactive MOs in terms

of QUAOs would manifestly destroy the closed-shell format of this part of the ORMAS wave function. These spectator MOs can however be expressed in terms of the occupied localized bond orbitals and lone-pair orbitals, which result from the transformation to split-localized orbitals. The procedure for determining the split-localized MOs is the same as that described above in Sections 4.1 and 4.2.3. That is also the case for the determination of the externals QUAOs.

5. SUMMARY

The analysis of HF wave functions in terms of QUAOs, which was developed in the first paper on this subject,[1] has been generalized for the analysis of strongly correlated wave functions that are generated from configurations in an optimized internal orbital space. The dimension of the internal orbital space is typically perceived to be equal to that of the conceptual MBS of the entire molecule, although the dimension may slightly exceed the MBS space (see remarks in Section 2.1). The wave function may encompass all internal configurations (FORS/CAS), or it may comprise part of the internal space in form of an ORMAS-type partitioning. Several improvements and generalizations regarding the methodology of the preceding paper[1] have also been developed. All aspects of the analysis described in this paper have been implemented, tested, and included as a module in the GAMESS suite of molecular programs.[46,47] The accompanying studies of urea[2] and the dioxetane dissociation[3] demonstrate the capabilities of the described methodology.

■ APPENDIX. ORTHOGONAL QUASI-ATOMIC ORBITALS WITH MAXIMAL PROJECTIONS ON THE ORBITAL SPACES OF THE FREE ATOMS

Let $|\nu\rangle$ be an orthonormal orbital set that spans the internal orbital space. They are chosen to be the QUAOs obtained by the symmetric orthogonalization of the nonorthogonal SVD-QUAOs, as described in Section V.B.1 of the preceding paper. They are therefore already associated with specific atoms. The desired new set of orthogonal QUAOs, $|Aa\rangle$, is obtained from the $|\nu\rangle$ by an orthogonal transformation $\mathbf{U}$:

$$|Aa\rangle = \sum_\nu |\nu\rangle U_{\nu,Aa}; \qquad U^T U = I \tag{A.0}$$

The orthogonal transformation $\mathbf{U}$ shall be determined by the requirement that each $|Aa\rangle$ has a maximal projection on the small space spanned by the AAMBS orbitals $|A^*\alpha\rangle$ on the respective atom A. Here, the orbitals $|A^*\alpha\rangle$ are the "Accurate Atomic Minimal Basis Set orbitals" on atom A that were introduced in Section III.B.2 of paper I.

Since the orbitals $|A^*\alpha\rangle$ on any one atom are orthonormal, the square of the projection of the QUAO $|Aa\rangle$ onto the space spanned by the orbitals $|A^*\alpha\rangle$ of atom A is given by

$$\sum_\alpha \langle A^*\alpha|Aa\rangle^2 = \sum_\mu \sum_\nu U_{\mu,Aa}\left[\sum_\alpha \langle \mu|A^*\alpha\rangle\langle A^*\alpha|\nu\rangle\right]U_{\nu,Aa} \tag{A.1}$$

where the sum $\sum_\alpha$ goes over all AAMBS orbitals on atom A, whereas the sums $\sum_\mu \sum_\nu$ go over *all* basis orbitals of the internal space.

The sum of these projection squares for all QUAOs $|Aa\rangle$ that belong to atom A is then

$$P_A = \sum_\mu \sum_\nu \left(Q_{\mu\nu}{}^A\right)\left(U_{\mu\nu}{}^A\right) \tag{A.2}$$

$$Q_{\mu\nu}{}^A = \sum_\alpha \langle \mu|A^*\alpha\rangle\langle A^*\alpha|\nu\rangle \tag{A.3}$$

$$U_{\mu\nu}{}^A = \sum_a (U_{\mu,Aa})(U_{\nu,Aa}) \tag{A.4}$$

where the sum $\sum_a$ goes over all new orbitals $|Aa\rangle$ to be associated with atom A. The criterion for the determination of the orthogonal matrix $U_{\nu,Aa}$ is then taken as

$$P = \sum_A P_A = \text{maximal} \tag{A.5}$$

The optimization is performed iteratively by a sequence of Jacobi rotations. It is manifest from eq A.4 that the quantity $U_{\mu\nu}{}^A$ is invariant under any orthogonal transformation that mixes only orbitals on the same atom. Consequently, only those Jacobi rotations must be considered that involve orbitals from two different atoms. Hence, for a given Jacobi rotation, the sums $\sum_\mu$ and $\sum_\nu$ in eq A.2 run only over these two atoms, and the same is true for the sum $\sum_A$ in eq A.5.

Let one orbital be on atom A and the other on an atom B. Denote the *known* initial orbitals of a particular Jacobi iteration as $|\mu\rangle = |Aa'\rangle$ and $|\nu\rangle = |Bb'\rangle$ and the *final* orbitals as $|Aa\rangle$ and $|Bb\rangle$, corresponding to the notation in eq A.4. Then, the Jacobi rotation is given by

$$|Aa\rangle = |Aa'\rangle U_{Aa',Aa} + |Bb'\rangle U_{Bb',Aa}$$
$$= |Aa'\rangle\cos\theta + |Bb'\rangle\sin\theta$$

$$|Bb\rangle = -|Aa'\rangle U_{Aa',Aa} + |Bb'\rangle U_{Bb',Aa}$$
$$= -|Aa'\rangle\sin\theta + |Bb'\rangle\cos\theta \tag{A.6}$$

Insertion of this rotation in place of $\mathbf{U}$ into eqs A.2 to A.5 yields, after some transformations:

$$P = C + (D^2 + F^2)^{1/2}\cos(\gamma - 2\theta) \tag{A.7}$$

with γ being defined by

$$\cos\gamma = D/(D^2 + F^2)^{1/2} \tag{A.8a}$$

$$\sin\gamma = F/(D^2 + F^2)^{1/2} \tag{A.8b}$$

where

$$C = (Q_{Aa',Aa'}{}^A + Q_{Bb',Bb'}{}^A + Q_{Aa',Aa'}{}^B + Q_{Bb',Bb'}{}^B)/2 \tag{A.9}$$

$$D = (Q_{Aa',Aa'}{}^A - Q_{Bb',Bb'}{}^A + Q_{Bb',Bb'}{}^B - Q_{Aa',Aa'}{}^B)/2 \tag{A.10}$$

$$F = (Q_{Aa',Bb'}{}^A - Q_{Aa',Bb'}{}^B) \tag{A.11}$$

The projection sum P of eq A.7 is maximal for $\theta = \gamma/2 + n\pi$, where n is an integer that should be chosen such that $-\pi/2 < \theta < \pi/2$.

If the molecular wave function is not given in the full valence space but is of the ORMAS type, it is sometimes expedient to construct restrained QUAOs, as discussed in Section 4.2.2 and in the last paragraph of Section 4.3. In this case, the summations over μ and ν in all equations are restricted to indices within the same ORMAS group.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 515-294-5253. E-mail: ruedenberg@iastate.edu.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) 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 hereafter referred to as Paper I or the preceding paper..

(2) West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. A Comprehensive Analysis in Terms of Molecule-Intrinsic, Quasi-Atomic Orbitals. III. The Covalent Bonding Structure of Urea. *J. Phys. Chem. A* **2015**, 150915114548004, first subsequent paper.10.1021/acs.jpca.5b03400

(3) West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. A Comprehensive Analysis in Terms of Molecule-intrinsic, Quasi-atomic Orbitals. IV. Bond Breaking and Bond Forming along the Dissociative Reaction Path of Dioxetane. *J. Phys. Chem. A* **2015**, 150915124335001, second subsequent paper.10.1021/acs.jpca.5b03402

(4) Ruedenberg, K.; Sundberg, K. MCSCF Studies of Chemical Reactions: Natural Reaction Orbitals and Localized Reaction Orbitals. In *Quantum Science*; Calais, J. L., Goscinski, O., Linderberg, J., Öhrn, Y., Eds.; Plenum Press: New York, 1976; pp 505–115.

(5) Ruedenberg, K. Determination of Orbitals and Selection of Configurations Through the Method of the Full Optimized Reaction Space (FORS). In *Post-Hartree-Fock: Configuration Interaction*; Lawrence Berkeley Laboratory: U. of California, Report LBL 8233, 1979; pp 46–64.

(6) Cheung, L. M.; Sundberg, K. R.; Ruedenberg, K. Electronic Rearrangements during Chemical Reactions. II. Planar Dissociation of Ethylene. *Int. J. Quantum Chem.* **1979**, *16*, 1103–39.

(7) Ruedenberg, K. MCSCF Optimization through Iterative CI Calculations in the Single Excitation Space, and MCSCF Wave Functions in the Full Reaction Space. In *Recent Developments and Applications of Multi-Configuration Hartree-Fock Methods*; Dupuis, M., Ed.; Lawrence Berkeley Laboratory: U. of California, Report LBL 12157, 1981; pp 51–57.

(8) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. Are Atoms Intrinsic to Molecular Electronic Wavefunctions? I. The FORS Model. *Chem. Phys.* **1982**, *71*, 41–49.

(9) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M. Are Atoms Intrinsic to Molecular Electronic Wavefunctions? II. Analysis of FORS Orbitals. *Chem. Phys.* **1982**, *71*, 51–64.

(10) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M.; Elbert, S. T. Are Atoms Intrinsic to Molecular Electronic Wavefunctions? III. Analysis of FORS Configurations. *Chem. Phys.* **1982**, *71*, 65–78.

(11) Subotnik, J. E.; Dutoi, A. D.; Head-Gordon, M. Fast Localized Orthonormal Virtual Orbitals which Depend Smoothly on Nuclear Coordinates. *J. Chem. Phys.* **2005**, *123*, 114108.

(12) Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. Molecule Intrinsic Minimal Basis Sets. I. Exact Resolution of ab initio Optimized Molecular Orbitals in Terms of Deformed Atomic Minimal-basis Orbitals. *J. Chem. Phys.* **2004**, *120*, 2629–2637.

(13) Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. Molecule Intrinsic Minimal Basis Sets. II. Bonding Analyses for Si_4H_6 and Si_2 to Si_{10}. *J. Chem. Phys.* **2004**, *120*, 2638–2651.

(14) Lu, W. C.; Wang, C. Z.; Chan, T. L.; Ruedenberg, K.; Ho, K. M. Representation of Electronic Structures in Crystals in Terms of Highly Localized Quasiatomic Minimal Basis Orbitals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2004**, *70*, 041101(R).

(15) Schmidt, M. W. Unpublished results. See also Table 3 of reference 1.

(16) Hunt, W. J.; Goddard, W. A. Excited States of H_2O using Improved Virtual Orbitals. *Chem. Phys. Lett.* **1969**, *3*, 414–418.

(17) Das, G.; Wahl, A. C. New Techniques for the Computation of Multiconfiguration Self-Consistent Field Wavefunctions. *J. Chem. Phys.* **1972**, *56*, 1769–1775.

(18) Whitten, J. L. Remarks on the Description of Excited Electronic States by Configuration Interaction Theory and a Study of the $^1(\pi\text{-}>\pi^*)$ State of H_2CO. *J. Chem. Phys.* **1972**, *56*, 5458–5466.

(19) Bauschlicher, C. W. The Construction of Modified Virtual Orbitals (MVOs) which are Suited for Configuration Interaction Calculations. *J. Chem. Phys.* **1980**, *72*, 880–885.

(20) Feller, D.; Davidson, E. R. An Approximation to Frozen Natural Orbitals through the use of the Hartree-Fock Exchange Potential. *J. Chem. Phys.* **1981**, *74*, 3977–3979.

(21) See Section III.C.3 of Paper I.

(22) The difference between the algorithms in refs 12–14 and 1 is that the AAMBS orbitals used in the first two paragraphs of Section V.A.1 of ref 1 are orthogonalized between atoms, whereas this orthogonalization had been omitted in refs 12–14. In ref 27, it has been suggested that the nonorthogonal AAMBS' orbitals are desirable in some contexts.

(23) See Section V.C.2 of Paper I.

(24) Equation (5.9) in ref 9.

(25) Pipek, J.; Mezey, P. Z. A Fast Intrinsic Localization Procedure Applicable for ab initio and Semiempirical Linear Combination of Atomic Orbital Wavefunctions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

(26) Knizia, G. Intrinsic Atomic Orbitals: an Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.

(27) Janowski, T. Near Equivalence of Intrinsic Atomic Orbitals and Quasiatomic Orbitals. *J. Chem. Theory Comput.* **2014**, *10*, 3085–3091.

(28) See Section V.B.1 of paper I.

(29) See Section VI.C of Paper I.

(30) Carlson, B. C.; Keller, J. M. Orthogonalization Procedures and the Localization of Wannier Functions. *Phys. Rev.* **1957**, *105*, 102–3.

(31) West, A. C. Weighted Orthogonalization of Atomic Orbitals: a Stable Alternative to the Carlson-Keller Method. *Comput. Theor. Chem.* **2014**, *1045*, 73–77.

(32) See Section V.B.2 of Paper I.

(33) Schmidt, M. W.; Ivanic, J.; Ruedenberg, K. Covalent Bonds are Created by the Drive of Electron Waves to Lower their Kinetic Energy through Expansion. *J. Chem. Phys.* **2014**, *140*, 204104.

(34) Schmidt, M. W.; Ivanic, J.; Ruedenberg, K. The Physical Origin of Covalent Bonding. In *The Chemical Bond, Fundamental Aspects of Chemical Bonding*; Frenking, G., Shaik, S., Eds.; Wiley-VCH: Weinheim, Germany, 2014; pp 1–67.

(35) Coulson, C. A.; Longuet-Higgins, H. C. The Electronic Structure of Conjugated Systems. I. General Theory. *Proc. R. Soc. London, Ser. A* **1947**, *191*, 39–60.

(36) Chirgwin, B. H. Summation Convention and the Density Matrix in Quantum Theory. *Phys. Rev.* **1957**, *107*, 1013–1025.

(37) See Appendix 1 of Paper I.

(38) See Section IV of Paper I.

(39) See Section V.A of Paper I.

(40) See Section V.B.1 of Paper I.

(41) See Section V.B.2 and 3 of Paper I.

(42) The details of the procedure are described in in Section V.C of Paper I.

(43) See Section V.C.2 and Appendix 2 of Paper I.

(44) As mention above at the end of Section 2.4, this localization method goes back to the paper cited above in ref 9.

(45) See Section VI of Paper I.

(46) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; et al. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993**, *14*, 1347−1363.

(47) Gordon, M. S.; Schmidt, M. W. In *Theory and Applications of Computational Chemistry, the first forty years;* Dykstra, C. E., Frenking, G., Kim, K.S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005; pp 1167−1189.

(48) Ivanic, J. Direct Configuration Interaction and Multiconfigurational Self-Consistent-Field method for Multiple Active Spaces with Variable Occupations. I. Method and II. Application to oxoMn(salen) and N_2O_4. *J. Chem. Phys.* **2003**, *119*, 9364−9376 and 9377−9385.

(49) See Section V.B.1 of Paper I.

(50) See Section V.A.1 of Paper I.

(51) See Sections III.B.2 and V.A.1 of Paper I.