Minimal Active Space: NOSCF and NOSI in Multistate Density Functional Theory

Yangyi Lu,[*,†] Ruoqi Zhao,[‡,†] Jun Zhang,[†] Meiyi Liu,[*,†] and Jiali Gao[*,†,¶]

†*Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518055, China*

‡*Institute of Theoretical Chemistry, Jilin University, Changchun, Jilin Province 130023, China*

¶*Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, United States*

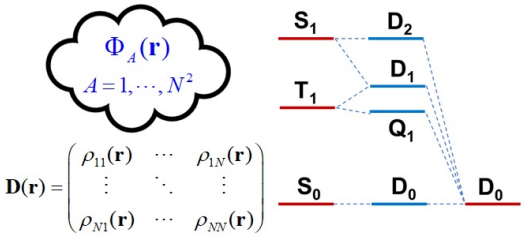
E-mail: luyy@szbl.ac.cn; liumy@szbl.ac.cn; gao@jialigao.org

Abstract

In this Perspective, we introduce a minimal active space (MAS) for the lowest N eigenstates of a molecular system in the framework of a multistate density functional theory (MSDFT), consisting of no more than N^2 nonorthgonal Slater determinants. In comparison with some methods in wave function theory in which one seeks to expand the ever increasing size of an active space to approximate the wave functions, it is possible to have an upper bound in MSDFT because the auxiliary states in a MAS are used to represent the exact N-dimensional matrix density $\mathbf{D}(\mathbf{r})$. In analogy to Kohn-Sham DFT, we partition the total Hamiltonian matrix functional $\mathcal{H}[\mathbf{D}]$ into an orbital-dependent part, including multistate kinetic energy $\mathbf{T}_{ms}$ and Coulomb-exchange energy $\mathbf{E}_{\mathrm{Hx}}$ plus an external potential energy $\int \mathrm{d}r\, v(r)\mathbf{D}(r)$, and a correlation matrix density functional $\mathcal{E}_c[\mathbf{D}]$. The latter accounts for the part of correlation energy not explicitly included in the minimal active space. However, a major difference from Kohn-Sham DFT is that state interactions are necessary to represent the N-matrix density $\mathbf{D}(\mathbf{r})$ in MSDFT, rather than a non-interacting reference state for the

scalar ground-state density $\rho_o(\mathbf{r})$. Two computational approaches are highlighted. We first derive a set of non-orthogonal multistate self-consistent-field (NOSCF) equations for the variational optimization of $\mathcal{H}[\mathbf{D}]$. We introduce the multistate correlation potential, as the functional derivative of $\mathcal{E}_c[\mathbf{D}]$, which includes both correlation effects within the MAS and that from the correlation matrix functional. Alternatively, we describe a non-orthogonal state interaction (NOSI) procedure, in which the determinant functions are optimized separately. Both computational methods are useful for determining the exact eigenstate energies and for constructing variational diabatic states, provided that the universal correlation matrix functional is known. It is hoped that this discussion would stimulate developments of approximate multistate density functionals both for the ground and excited states.

TOC Graphic


$\Phi_A(\mathbf{r})$
$A = 1, \cdots, N^2$
$\mathbf{D}(\mathbf{r}) = \begin{pmatrix} \rho_{11}(\mathbf{r}) & \cdots & \rho_{1N}(\mathbf{r}) \\ \vdots & \ddots & \vdots \\ \rho_{N1}(\mathbf{r}) & \cdots & \rho_{NN}(\mathbf{r}) \end{pmatrix}$
S1
D2
D1
T1
Q1
S0
D0
D0

1 Introduction

Kohn-Sham density functional theory (KS-DFT)[1] based on the theorems of Hohernberg and Kohn is a ground-state method.[2] Thanks to its computational efficiency and accuracy by including dynamic electron correlation, KS-DFT has been a method of choice for large systems.[3] The success is directly related to the development of many exceedingly good, approximate density functionals,[3,4] and continuing efforts are clearly needed.[5] Excited-state energies can also be determined by DFT through linear-response (LR) time-dependent approaches (TDDFT),[6] though there are well-known short-comings including the treatment of conical intersection between the ground state and the first excited state.[7] Besides developing even better approximate functionals for the ground state to overcome these difficulties, a question has been lingering for sometime:[4] what are future directions of DFT?

Recently, two of the authors, Lu and Gao (LG),[8] proved that (1) the Hamiltonian in the subspace spanned by the lowest N eigenstates is a matrix functional, $\mathcal{H}[\mathbf{D}]$, of the multistate matrix density $\mathbf{D}(r)$, and (2) the variational minimization of a multistate energy with respect to $\mathbf{D}(r)$ yields the exact energies and densities of the N eigenstates.[8] These two theorems lay a foundation for a rigorous multistate DFT (MSDFT) that treats the ground and excited states on an equal footing. The matrix density $\mathbf{D}(r)$ of the N lowest eigenstates is a matrix function of the coordinate $\mathbf{r}$, in which the diagonal elements correspond to the densities of all eigenstates and the off-diagonal terms are the transition densities between the corresponding eigenstates. Importantly, it was shown that the N-dimensional matrix density $\mathbf{D}(r)$ can be represented by no more than N^2 nonorthogonal Slater determinants.[8]

In wave function theory (WFT), complete active space self-consistent-field (CASSCF) methods are widely used in electronic structure calculations.[9,10] It is a special case of multiconfiguration SCF (MC-SCF) approaches, which can be systematically improved by expanding the size of the CAS. Although CASSCF is almost perfect for treating strongly cor-

related systems, its accuracy depends on post-SCF correction for dynamic correlation.[11] A difficulty in the latter step, aside from its computational cost, is uneven rates of convergence for different electronic states and its dependency on the choice of the active space.[10,12] Consequently, such a black-box approach still needs to be validated via trial and error in practice.[13]

Is there an alternative, a smaller active space (than CAS) or even a minimal active space (MAS) for the excited states of interest? The answer to the question is positive in excited-state DFT.[8,14] Then, the effort would be shifted from finding proper orbitals and number of electrons for a given problem to defining representative states in the active space based on physical properties. In fact, the idea of a MAS is no stranger to electronic structure calculations. Perhaps, the best-known example is Kohn-Sham density functional theory (KS-DFT), which employs a single Slater determinant to represent the exact density of the ground state.[1,15] Consequently, a non-interacting reference for the real system is possible, guaranteed by the Hohenberg-Kohn theorems.[1–3,15,16] Furthermore, LR-TDDFT provides a good balance between accuracy and cost for excited states, making use of this non-interacting reference state.[6,17] [18,19] Nevertheless, it has long been a goal of theorists to develop a time-independent density functional theory beyond the ground state.[20–25] The aim of this Perspective is to introduce a minimal active space (MAS) for a given number of N lowest-energy states in the framework of multistate density functional theory along with two computational methods: nonorthogonal self-consistent-field (NOSCF) equations and nonorthogonal state interaction (NOSI) for optimization of the auxiliary states. It is hoped that this discussion shall stimulate further developments of time-independent density functional theories to treat both the ground and excited states.

In what follows, we first discuss the two fundamental theorems of multistate DFT[8] and its relationship with previous works and wave function theory. Then, we present MAS for a given number of N states of interest, and the definition of the matrix functionals for the kinetic, Coulomb and exchange, and correlation energy operators. Approxi-

mations to construct the matrix elements of the Hamiltonian matrix functional with the use of approximate functionals developed for KS-DFT are discussed. This is followed by presentation of a multistate self-consistent-field (MS-SCF) method to optimize both the orbitals and configuration coefficients, and a nonorthogonal state interaction (NOSI) approach in which orbitals are separately optimized. The first method provides a formally exact solution for the eigenstates and vectors in the N-dimensional subspace, given that the universal correlation matrix functional was known. NOSI is a computationally efficient alternative to the full MS-SCF theory.[8,14,26,27] Finally, the two computational approaches are illustrated, respectively, by determining the potential energy surfaces about the conical intersection point between the ground and the first excited state in the photodissociation of NH_3 and a spintronic sensor involving coupling interactions between the triplet state of 1-chloro-9,10-anthraquinone and a model nitroxide free radical.

2 Multistate density functional theory

Suppose we are interested in the lowest N eigenstates of a molecular system, which defines the subspace $\mathbb{V}_{\min} \in \mathbb{H}$. The system is described by the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}_{\text{ext}}$, consisting of n_e electrons under the multiplicative external potential $\hat{V}_{\text{ext}} = \sum_{j=1}^{n_e} v(r_j)$ due to nuclear charges. Here, $\hat{H}_0 = \hat{T} + \hat{W}$ includes the standard electronic kinetic energy and Coulomb operators.

2.1 Hamiltonian matrix functional and variational principle

The original density functional theory of Hohenberg and Kohn (HK) is a ground-state theory.[2] For excited states, Theophilou introduced an analogous, subspace DFT,[21] establishing a one-to-one relationship between an N-dimensional subspace and the subspace density. However, the Theophilou subspace theory does not offer directly the energies and vectors of the individual states. Recently, LG proved two theorems for the lowest N

states of the Hamiltonian, including the ground state.

First, the Hamiltonian $\hat{H}_0$ projected on to the subspace $\mathbb{V}_{\min}$ is a universal matrix functional of the multistate matrix density $\mathbf{D}(r)$: $\mathbf{H}_0 = \mathcal{F}[\mathbf{D}]$.[8] Then, the matrix functional of the full Hamiltonian $\hat{H}$ for the system of interest is given by

$$\mathcal{H}[\mathbf{D}] = \mathcal{F}[\mathbf{D}] + \int \mathrm{d}r\, v(r)\mathbf{D}(r) \tag{1}$$

where $v(r)$ is the one-body external potential. $\mathcal{F}[\mathbf{D}]$ is universal because it does not depend on the specific external potential $v(r)$. Conceptually, the HK density functional is a scalar quantity, a (1x1)-matrix here, for which a non-interacting reference system is sufficient (KS-DFT) to represent the ground-state density $\rho_0(r)$.[1,2]

The second theorem states that the matrix density $\mathbf{D}(r)$ can be obtained variationally by minimizing the multistate energy,[8]

$$\mathbb{E}[v] = \min_{\mathbf{D}(r)}\{E_{\mathrm{MS}}[\mathbf{D}]\} \tag{2}$$

where $E_{\mathrm{MS}} = N^{-1}\sum_{I=1}^{N} E_I$, with $E_I = \sum_{A,B}^{N} C_{IA}C_{IB}H_{AB}[\mathbf{D}]$ being the energy of eigenstate I. The coefficients $\{C_{IA}\}$ are elements of the matrix $\mathbf{C}$ that diagonalizes the Hamiltonian matrix functional $\mathcal{H}[\mathbf{D}]$.

2.2 Discussion

Several remarks may be made from different perspectives.

(1) Although the minimization object of eq 2 is the same as the subspace energy $\mathbb{E}[v]$ of Theophilou,[21] the two theories are in fact different at least for three reasons. First, the fundamental variable that is optimized in eq 2 is a matrix density $\mathbf{D}(r)$, whereas the subspace optimization of Theophilou makes use of the subspace density $\rho_V(r)$, defined as $\rho_V(r) = \frac{1}{N}\sum_I^N \rho_I(r)$.[21] Theophilou showed that $\rho_V(r)$ uniquely determines the subspace

$\mathbb{V}_{\min}$ spanned by the lowest N eigenstates,[21] thus, identical to the subspace of eq 2 since $\mathbf{D}(r)$ determines $\rho_V(r)$ by $\rho_V(r) = \mathrm{tr}\,[\mathbf{S}^{-1}\mathbf{D}(r)]$.[28] Here, $\mathbf{S}$ is the corresponding overlap matrix of $\mathbf{D}(r)$ given by $S_{AB} = n^{-1}\int \mathrm{d}r\, D_{AB}(r)$. Note that $\mathbf{D}(r)$ is a matrix that consists of basis-state densities (diagonal) and transition densities (off-diagonal), but $\rho_V(r)$ is a scalar function. Although the subspace density $\rho_V(r)$ can be obtained from the matrix density $\mathbf{D}(r)$, it is not possible in the reverse direction without invoking eqs 1 and 2. Therefore, the information contents of the two theories are different.

Secondly, the subspace theory of Theophilou offers only the total energy of the subspace $\mathbb{E}[v]$ without the energies and densities of the individual eigenstates (eq 2).[21] On the other hand, the matrix density function in the theorems of LG provides the necessary information to construct a Hamiltonian matrix functional that can be diagonalized. Consequently, as the multistate (subspace) energy is variationally optimized (eq 2), the exact eigenenergies and vectors of all individual states are simultaneously obtained.[8]

Thirdly, because of the one-to-one relationship between $\mathbb{E}[v]$ and $\rho_V(r)$ in the Theophilou subspace theory, a noninteracting reference system with a single Slater determinant, in exactly the same way as KS-DFT for the ground state,[21] can be used to represent $\rho_V(r)$. This led to the unfortunate impression that a non-interacting system can be used to determine the eigenstate energies of an ensemble of states. Importantly, it has critical implications on the origin of the lack of individual states in subspace DFT. In contrast, the nature of the Hamiltonian matrix functional itself dictates that an interacting reference system is needed to represent the multistate matrix density $\mathbf{D}(r)$. Thus, a straightforward application of a Kohn-Sham-like approach, although sufficient for $\mathbb{E}[v]$, will be inadequate to determine individual eigenstates in an excited-state DFT.

(2) Because of the trace of a matrix is invariant to a unitary transformation, the multistate (subspace) energy can be equivalently written as[8]

$$E_{\mathrm{MS}}[\mathbf{D}] = \mathrm{tr}\left[\mathbf{S}^{-1}\mathcal{H}[\mathbf{D}]\right] \tag{3}$$

Here, we note that the subspace trace has been normalized by the dimension of the subspace N. This functional expression is conveniently used in a computational algorithm presented below. The diagonal elements of $\mathcal{H}[\mathbf{D}]$ correspond to the energies of the basis states that defines $\mathbf{D}(r)$, not necessarily the energies of the eigenstates, whether or not orthogonal or nonorthogonal basis states are used.

(3) The matrix functional of the Hamiltonian and its functional variable $\mathbf{D}(r)$ are related by linear transformation properties of basis states.[29,30] Given $\Phi'_B = \sum_{A=1}^{N} L^*_{BA}\Phi_A$, where $\{\Phi'_B; B = 1, ..., N\}$ is another set of basis states in the minimal active space, the corresponding N-matrix density $\mathbf{D}'(r)$ transforms bilinearly,[8] $\mathbf{D}'(r) = \mathbf{L}\mathbf{D}(r)\mathbf{L}^\dagger$. Then, the Hamiltonian matrix functional $\mathcal{H}[\mathbf{D}]$ follows the same bilinear transformation,[31]

$$\mathcal{H}[\mathbf{L}\mathbf{D}\mathbf{L}^\dagger] = \mathbf{L}\mathcal{H}[\mathbf{D}]\mathbf{L}^\dagger \tag{4}$$

This transformation property applies to the individual terms of $\mathcal{H}[\mathbf{D}]$, including matrices constructed below within a minimal active space, $\mathbf{T}_{ms}, \mathbf{E}_{\mathrm{Hx}}$, as well as the correlation matrix functional $\mathcal{E}_c[\mathbf{D}]$.

(4) The variational minimization of eq 2 may be recast as a generalized constrained search, originally developed for the Hohenberg-Kohn theory.[32]

$$\mathbb{E}[v] = \min_{\mathbf{D}(r)} \left\{ N^{-1} \sum_{AB}^{N} \left[(\mathbf{S}^{-1})_{AB} \langle \Phi_A | \hat{H} | \Phi_B \rangle \right] ; \{\Phi_A\} \to \mathbf{D}(r) \right\} \tag{5}$$

Thus, the optimization of the N-matrix density $\mathbf{D}(r)$ can be considered as searching for the minimal subspace energy within all possible sets of N linearly independent wave functions. In general, the wave functions are not necessarily orthogonal. The optimized wave functions all belong to the subspace $\mathbb{V}_{\mathrm{min}}$, which are linear combinations of the lowest N eigenstates.

(5) The multistate energy in eqs 2 and 3, or, identically, the subspace energy of Theophilou,[21] is defined as the average of all eigenstate energies of the subspace $\mathbb{V}_{\mathrm{min}}$. Its variational

minimization, at first glance, looks like that of the state-average CASSCF approach in WFT.[9,11] However, there is a fundamental difference between the "state average" of MS-DFT and that of WFT. The minimization of eq 2 is based on a variational principle such that the multistate energy and its individual eigenstate energies are exactly defined. Provided that the universal functional $\mathcal{F}[\mathbf{D}]$ is known, being a density functional theory,[15] the eigenstate energies do not depend on the size of the subspace, i.e., the number of states N included in the state average. In CASSCF, as is well-known, the energies of individual states in the state-average minimization depend on the number of states and the weights used.

(6) In the Gross-Oliveira-Kohn variant of subspace DFT (called ensemble DFT),[22,23] the subspace energy is written as a sum of weighted eigenstate energies,

$$\mathbb{E}_w[v] = \min_{\rho_W(r)} \{ \sum_I w_I E_I[\rho_I(r)] \} \tag{6}$$

where $\rho_W(r) = \sum_I w_I \rho_I(r)$, and the weights are normalized to unity, which must remain constant or monotonically decreasing with increasing energies of the eigenstates. The eigenstate-weighting scheme is equivalent to weighted Theophilou subspaces, whose differences indeed yield the energy changes between different subspaces.[21] A number of applications have been made to extract eigenstate energies from the derivatives of the subspace energy with respect to the state weight.[22,23,33] However, an implicit assumption in practice using eq 6 is that there is a Hohenberg-Kohn-like relationship for each individual excited state, which to our knowledge does not exist. We now know from eqs 1 and 2 that the practical application of eq 6 is only ensured when $\mathcal{H}[\mathbf{D}]$ is diagonal. In other words, state interaction is needed in subspace DFT in order to obtain the eigenstate energies.

3 Minimal active space

The fundamental variable in MSDFT is the $N \times N$ matrix density $\mathbf{D}(r)$ as a function of the spatial coordinates $\mathbf{r}$.[8] Recall that the trace of $\mathbf{D}(r)$ is the subspace density $\rho_V(r)$ that uniquely determines the subspace $\mathbb{V}_{\min}$ and the multistate energy $E_{\rm MS}$ (eq 2).[21]

In order to design a computational procedure, it is necessary to represent the exact N-matrix density $\mathbf{D}(r)$ of the real system by a set of auxiliary Slater determinants. Theorem 3 of LG[8] states that the N-matrix density $\mathbf{D}(r)$ can be sufficiently represented by no more than N^2 independent, i.e., non-orthogonal, Slater determinants.

3.1 Minimal active space for multistate DFT

As an obvious departure from Kohn-Sham DFT, the auxiliary system designed to represent $\mathbf{D}(r)$ of the subspace $\mathbb{V}_{\min}$ can no longer be non-interacting. To this end, we introduce a minimal active space, consisting of N basis states $\{\Phi_A; A = 1, \cdots, N\}$, each of which is written as a linear combination of N^2 nonorthogonal determinants,

$$\Phi_A = \sum_{\xi}^{N^2} c_\xi^A \Xi_\xi \tag{7}$$

where $\{c_\xi^A\}$ are configuration coefficients. In eq 7, Ξ_ξ is a Slater determinant of n_e one-body spin-orbitals $\{\phi_{i\sigma}^\xi\}$

$$\Xi_\xi(r_1, \cdots, r_{n_e}) = \frac{1}{\sqrt{n_e!}} \hat{A}\{\phi_{1\sigma_1}^\xi(r_1) \cdots \phi_{n_e\sigma_n}^\xi(r_{n_e})\} \tag{8}$$

where $\hat{A}$ is the anti-symmetrizer. The spin-orbital is written as a product of the spatial and spin orbitals: $\phi_{j\sigma}^\xi(r) = \phi_j^\xi(r)\sigma$, where σ represents the spin-up α or spin-down β state. The orbitals with the same spin are orthonormal $\left\langle \phi_{i\sigma}^\xi | \phi_{j\sigma'}^\xi \right\rangle = \delta_{ij}\delta_{\sigma\sigma'}$; however, orthogonality is not imposed on orbitals between different determinants. Therefore, the basis determinants in the active space are generally nonorthogonal. The spin orbitals

$\phi^{\xi}_{j\sigma}(r)$ in different determinants must differ by at least one or more terms, i.e., linearly independent, and one way of defining these targeted basis states has been described in reference 34, which is called targeted-state optimizaiton (TSO).[34] In general, we assume that all determinants in the active space, $\{\Xi_\xi\}$, have the same number of spin-up (n_α) and spin-down (n_β) electrons because the Hamiltonian matrix functional is block diagonal with different eigenvalues of $\hat{S}_z$: $\frac{1}{2}(n_\alpha - n_\beta)$.[8]

The N-matrix density $\mathbf{D}(r)$ of the minimal active space is given by its element $D_{AB}(r) = \langle\Phi_A|\hat{\rho}(r)|\Phi_B\rangle$ in terms of nonorthogonal orbitals:

$$D_{AB}(r) = \sum_{\xi\zeta}^{N^2} c^A_\xi c^B_\zeta \sum_{\sigma=\alpha,\beta} \sum_{j,k}^{n_\sigma} f^{\xi\zeta}_{jk;\sigma} \phi^{\xi}_{j\sigma}(r) \phi^{\zeta}_{k\sigma}(r) \tag{9}$$

where the coefficient $f^{\xi\zeta}_{jk}$ is given in eq A.37.

The definition of N auxiliary states $\{\Phi_A\}$ in eq 7 is aimed for representing both the diagonal densities of the basis states and the transition densities. Just as KS-DFT in which the Kohn-Sham determinant is used purely to represent the ground-state density, these auxiliary states are not the wave functions for any specific states. It is certainly possible to adopt an alternative construction to represent $\mathbf{D}(r)$ exactly, but we shall not discuss these scenarios further in this Perspective.

3.2 Kinetic, Hartree-exchange and correlation matrix functionals

Given the multistate multiconfigurational MAS defined above to represent $\mathbf{D}(r)$, the total Hamiltonian matrix functional may be written as follows.

$$\mathcal{H}[\mathbf{D}] = \mathbf{T}_{ms} + \mathbf{E}_{\mathrm{Hx}} + \mathcal{E}_c[\mathbf{D}] + \int \mathrm{d}r\, v(r)\mathbf{D}(r) \tag{10}$$

where the first two terms originate from the Hamiltonian $\mathbf{H}_0$: $< \Phi_A|\hat{H}_0|\Phi_B >= T^{AB}_{ms} + E^{AB}_{\mathrm{Hx}}$ (eq 1).

First, the multistate kinetic energy matrix functional of the MAS, $\mathbf{T}_{ms}$, is given by $T_{ms}^{AB} = \left\langle \Phi_A | \hat{T} | \Phi_B \right\rangle$ with the expression of

$$T_{ms}^{AB} = -\frac{\hbar^2}{2m} \sum_{\xi\zeta}^{N^2} c_\xi^A c_\zeta^B \sum_{\sigma=\alpha,\beta} \sum_{j,k}^{n_\sigma} f_{jk;\sigma}^{\xi\zeta} \left\langle \phi_{j;\sigma}^{\xi} | \nabla^2 | \phi_{k;\sigma}^{\zeta} \right\rangle \tag{11}$$

And, the electronic Hartree and exchange-interaction matrix $\mathbf{E}_{\mathrm{Hx}}$ is $E_{\mathrm{Hx}}^{AB} = \left\langle \Phi_A | \hat{W} | \Phi_B \right\rangle$,

$$E_{\mathrm{Hx}}^{AB} = \sum_{\xi\zeta}^{N^2} c_\xi^A c_\zeta^B \sum_{\sigma,\sigma'=\alpha,\beta} \sum_{i,k}^{n_\sigma} \sum_{j,l}^{n_{\sigma'}} f_{ij,kl;\sigma,\sigma'}^{\xi\zeta} \langle ij||kl\rangle_{\xi\zeta} \tag{12}$$

where $\langle ij||kl\rangle_{\xi\zeta}$ is the two-electron Coulomb-exchange integrals (eq A.39) with $f_{ij,kl;\sigma,\sigma'}^{\xi\zeta}$ being the coefficient (eq A.38).

Then, the correlation matrix functional is defined via eq 1

$$\mathcal{E}_c[\mathbf{D}] = \mathcal{F}[\mathbf{D}] - (\mathbf{T}_{ms} + \mathbf{E}_{\mathrm{Hx}}) \tag{13}$$

Notice that eq 10 has the same appearance as KS-DFT;[1,3] however, each term in the MSDFT Hamiltonian is an N-dimensional matrix functional, whereas those in Kohn-Sham theory are scalar quantities. The matrix functional of eq 10, one form of an implicit functional of $\mathbf{D}(r)$, consists of explicit state interactions, i.e., non-zero off-diagonal functional terms, in contrast to the single-determinant, non-interacting reference system for one state.

3.3 Spin symmetry

Here, we consider the structure of spin symmetry in MSDFT.

Since the Hamiltonian commutes with $\hat{S}_z$, the auxiliary states $\{\Phi_A\}$ in eq 7 are linear combinations of the determinants that have the same eigenvalue of $\hat{S}_z$. Consequently, Φ_A itself is an eigenstate of $\hat{S}_z$ with an eigenvalue of S_z^A.[35] Then, the matrix density $\mathbf{D}(r)$ is

block diagonal in terms of the eigenstates of $\hat{S}_z$. That is, if $S_z^A \neq S_z^B$, $D_{AB}(r) = 0$. Hence, we also have $\mathcal{H}_{AB}[\mathbf{D}] = 0$. That is, the Hamiltonian matrix functional is block diagonal with different eigenvalues of $\hat{S}_z$. We note that all spin-complement configurations are included in the active space as the subspace contains all degenerate states. This ensures that there is no spin-symmetry breaking during the optimization. Thus, the eigenstates of the Hamiltonian matrix functional are also eigenstates of the total spin operator $\hat{S}^2$.

This is illustrated by the Hamiltonian matrix functional of the lowest five eigenstates of the diatomic molecule, LiF, which consists of two singlet states, resulting from the admixture of an ionic and a covalent configuration, plus the lowest triplet states with S_z = +1, 0, and -1. The 5-matrix density $\mathbf{D}(r)$ is block-diagonal with respect to different eigenvalues of $\hat{S}_z$, $\mathbf{D}(r) = \mathbf{D}^{(-1)}(r) \oplus \mathbf{D}^{(0)}(r) \oplus \mathbf{D}^{(+1)}(r)$, where each of $\mathbf{D}^{(\pm 1)}(r)$ has one single element as the triplet state density $\rho_T^{(\pm)}(r)$, corresponding to $S_z = \pm 1$, and $\mathbf{D}^{(0)}(r)$ is a 3-matrix density, containing two (the ground and the first excited) singlet states and one triplet state of the $S_z = 0$ projection. The 5-state Hamiltonian matrix functional $\mathcal{H}[\mathbf{D}]$ has the same block-diagonal structure

$$\mathcal{H}[\mathbf{D}] = \mathcal{H}^{(-1)}[\mathbf{D}^{(-1)}] \oplus \mathcal{H}^{(0)}[\mathbf{D}^{(0)}] \oplus \mathcal{H}^{(+1)}[\mathbf{D}^{(+1)}] \tag{14}$$

where $\mathcal{H}^{(\pm 1)}[\mathbf{D}^{(\pm 1)}] = E_T[\rho_T^{\pm}]$, is a single-state energy functional, and $\mathcal{H}^{(0)}[\mathbf{D}^{(0)}]$ represents a 3-state Hamiltonian matrix functional. As a result, the 5-state subspace energy functional (eq 2) includes three independent terms

$$E_{\mathrm{MS}}[\mathbf{D}] = \frac{1}{5}\left(E_T[\rho_T^{(+)}] + E_T[\rho_T^{(-)}] + 3E_{\mathrm{MS}}^{(0)}[\mathbf{D}^{(0)}]\right) \tag{15}$$

Minimization of eq 15 is equivalent to the minimization of the three energy functionals separately. One immediately recognizes that $E_T[\rho_T^{(\pm)}]$ are the energy functionals of KS-DFT, respectively, corresponding to a triplet determinant with either two extra spin-up or spin-down electrons. Consequently, the problem is reduced to finding the 3-state Hamil-

tonian matrix functional $\mathcal{H}^{(0)}[\mathbf{D}^{(0)}]$, giving rise to the ground and first excited singlet states and one spin-multiplet of the lowest triplet state. The latter must be energy degenerate with that of the other two blocks − a constraint to the correlation matrix functional.[36,37]

This conclusion can be straightforwardly generalized to all spin manifolds.[37]

4 Nonorthogonal self-consistent-field theory

We use the term multistate SCF (MS-SCF) to distinguish the present DFT approach from the well-known MC-SCF methods in WFT because the basis states in the active space for DFT include dynamic correlation, a dynamic-then-static ansatz.[38] However, we emphasize that these states (eq 7) are used as auxiliary functions to represent the exact matrix density $\mathbf{D}(r)$ of the N-dimensional subspace $\mathbb{V}_{\min}$.[14] In this section, we derive the self-consistent-field equations for the general case of a nonorthogonal multistate DFT, or NOSCF.

The variational minimization of the multistate energy (eq 2) is carried out using the expression in eq 3 subject to the constraints of normalization for each multiconfigurational state, $\langle\Phi_A|\Phi_A\rangle = 1$, and the orthonormal conditions for orbitals within each determinant, $\left\langle \phi^{\xi}_{j\sigma}|\phi^{\xi}_{k\sigma}\right\rangle = \delta_{jk}$. Consequently, we introduce the Lagrangian

$$\begin{aligned} L[\mathbf{D}] =& E_{\mathrm{MS}}[\mathbf{D}] - N^{-1}\sum_{A}^{N} E_A \left(\sum_{\xi\zeta}^{N^2} c^A_{\xi} S_{\xi\zeta} c^A_{\zeta} - 1\right) \\ & - \sum_{\xi}^{N^2}\sum_{\sigma=\alpha,\beta}\sum_{jk}^{n_\sigma} \epsilon^{\xi}_{jk;\sigma}\left(\left\langle \phi^{\xi}_{j\sigma}|\phi^{\xi}_{k\sigma}\right\rangle - \delta_{jk}\right) \end{aligned} \tag{16}$$

where $S_{\xi\zeta} = \langle\Xi_\xi|\Xi_\zeta\rangle$ is the overlap between two determinants, and $\{E_A\}$ and $\{\epsilon^{\xi}_{jk;\sigma}\}$ are the Lagrange multipliers that impose the constraints. The Lagrangian $L[\mathbf{D}]$ (eq 16) is minimized simultaneously with respect to variations of both the orbitals $\{\phi^{\xi}_{j\sigma}(r)\}$ and the

CI coefficients $\{c_\xi^A\}$,

$$\frac{\delta}{\delta c_\xi^A} L = 0, \quad \frac{\delta}{\delta \phi_{j\sigma}^\xi(r)} L = 0 \tag{17}$$

4.1 Functional derivatives

To begin, we introduce the functional derivatives of the correlation matrix functional $\mathcal{E}_c[\mathbf{D}]$. We emphasize that $\mathcal{E}_c[\mathbf{D}]$ is a one-to-one mapping between the matrix density $\mathbf{D}(r)$ and the correlation matrix, rather than a term-by-term mapping between the corresponding elements of the two matrices.[30] That is, in the general scenario, each matrix element, $\mathcal{E}_c^{AB}[\mathbf{D}]$, is an implicit functional of all elements of $\mathbf{D}(r)$.[8] This complication arises naturally from the definition of the correlation matrix functional $\mathcal{E}_c[\mathbf{D}]$ (eq 13), which is the part of electronic correlation that is not explicitly included in configuration interaction among the auxiliary states used to represent the exact matrix density $\mathbf{D}(r)$ of the subspace $\mathbb{V}_{\text{min}}$. In analogy to the Kohn-Sham exchange-correlation functional,[3,18] it is expected that $\mathcal{E}_c[\mathbf{D}]$ is a nonlinear functional of the matrix density $\mathbf{D}(r)$. Therefore, the functional derivative of each element $\mathcal{E}_c^{AB}[\mathbf{D}]$ with respect to the transition density, $D_{EF}(r)$, leads to a four-index potential,

$$v_{EF}^{AB}(r) = \frac{\delta}{\delta D_{EF}(r)} \mathcal{E}_c^{AB}[\mathbf{D}] \tag{18}$$

where $A, B, E, F = 1, \cdots, N$.

It follows that the variation of $\mathcal{E}_c[\mathbf{D}]$ with respect to the CI coefficients is given by

$$\frac{\delta}{\delta c_\xi^A} \mathcal{E}_c^{BE}[\mathbf{D}] = 2 \sum_F^N \sum_\zeta^{N^2} c_\zeta^F \int \mathrm{d}r \, v_{AF}^{BE}(r) D_{\xi\zeta}(r) \tag{19}$$

where $D_{\xi\zeta}(r) = \langle \Xi_\xi | \hat{\rho}(r) | \Xi_\zeta \rangle$ is the transition density between two determinants. Simi-

larly, the variation of $\mathcal{E}_c[\mathbf{D}]$ with respect to the orbital $\phi^{\xi}_{i\sigma}(r)$ reads

$$\frac{\delta}{\delta\phi^{\xi}_{i\sigma}(r)}\mathcal{E}_c^{BC}[\mathbf{D}] = \sum_{EF}^{N}\sum_{\zeta}^{N^2}\left(c^E_{\xi}c^F_{\zeta} + c^E_{\zeta}c^F_{\xi}\right)\int \mathrm{d}r'\, v^{BC}_{EF}(r')\frac{\delta}{\delta\phi^{\xi}_{i\sigma}(r)}D_{\xi\zeta}(r') \tag{20}$$

The functional derivative of the transition density is written as

$$\begin{aligned}\frac{\delta}{\delta\phi^{\xi}_{i\sigma}(r)}D_{\xi\zeta}(r') = \sum_{k}^{n_\sigma}\Bigg(& f^{\xi\zeta}_{ik;\sigma}\delta(r-r')\phi^{\zeta}_{k\sigma}(r') + \sum_{j\neq i}^{n_\sigma}\frac{\delta f^{\xi\zeta}_{jk;\sigma}}{\delta\phi^{\xi}_{i\sigma}(r)}\phi^{\xi}_{j\sigma}(r')\phi^{\zeta}_{k\sigma}(r') \\ &+ \sum_{j}^{n_{\bar\sigma}}\frac{\delta f^{\xi\zeta}_{jk;\bar\sigma}}{\delta\phi^{\xi}_{i\sigma}(r)}\phi^{\xi}_{j\bar\sigma}(r')\phi^{\zeta}_{k\bar\sigma}(r')\Bigg)\end{aligned} \tag{21}$$

where the coefficient $f^{\xi\zeta}_{jk;\sigma}$ is given in eq A.37 and $\bar\sigma$ represents the opposite spin of σ. The last two terms on the right-hand-side are due to the use of nonorthogonal orbitals in different determinants, where the derivatives of coefficients are given in eq 36 of Supporting Information.

4.2 Generalized Fock equations for configuration coefficients

Variation of the Lagrangian $L[\mathbf{D}]$ with respect to c^A_p gives rise to

$$\begin{aligned}\frac{\delta}{\delta c^A_{\xi}}L =& N^{-1}\sum_{BC}^{N}\left(\mathbf{S}^{-1}_{BC}\frac{\delta}{\delta c^A_{\xi}}\mathcal{H}_{BC}[\mathbf{D}] - \frac{\delta S_{BC}}{\delta c^A_{\xi}}\left[\mathbf{S}^{-1}\mathcal{H}[\mathbf{D}]\mathbf{S}^{-1}\right]_{BC}\right) \\ & -2N^{-1}E_A\sum_{\zeta}^{N^2}S_{\xi\zeta}c^A_{\zeta} = 0\end{aligned} \tag{22}$$

where we have used the identity $\frac{\delta}{\delta c^A_{\xi}}\mathbf{S}^{-1} = -\mathbf{S}^{-1}\left(\frac{\delta}{\delta c^A_{\xi}}\mathbf{S}\right)\mathbf{S}^{-1}$. Solution of eq 22 yields the Fock-like equation,

$$\sum_{B}^{N}\sum_{\zeta}^{N^2}{}^{c}\mathbb{F}^{AB}_{\xi\zeta}c^B_{\zeta} = E_A\sum_{\eta}^{N^2}S_{\xi\eta}c^A_{\eta} \tag{23}$$

The appearance of the overlap matrix, $\{S_{\xi\zeta}\}$, in eq 23 is due to the use of nonorthogonal determinants to construct each configuration state Φ_A (eq 7). The Lagrangian multiplier E_A acquires the physical meaning of the state energy for state Φ_A. The Fock matrix for the CI coefficients in eq 23 reads as follows.

$$ {}^{c}\mathbb{F}^{AB}_{\xi\zeta} = (\mathbf{S}^{-1})_{AB} H_{\xi\zeta} + \int \mathrm{d}r\, v_c^{AB}(r) D_{\xi\zeta}(r) - (\mathbf{S}^{-1}\mathcal{H}[\mathbf{D}]\mathbf{S}^{-1})_{AB} S_{\xi\zeta} \tag{24} $$

The first and second terms of the Fock matrix ${}^{c}\mathbb{F}^{AB}_{\xi\zeta}$ come from the variation of the Hamiltonian matrix functional $\mathcal{H}[\mathbf{D}]$ with respect to c_p^A and the last term is due to variation of the overlap matrix $\mathbf{S}$. We have defined in eq 24 the direct interaction matrix element between two determinants,

$$ H_{\xi\zeta} = \left\langle \Xi_\xi \middle| \hat{H} \middle| \Xi_\zeta \right\rangle \tag{25} $$

4.3 Multistate correlation matrix potential

The *multistate correlation potential* between states A and B, $v_c^{AB}(r)$, in the second term of the Fock matrix (eq 24) is defined by the weighted average of the four-index potentials (eq 18),

$$ v_c^{AB}(r) = \sum_{EF}^{N} (\mathbf{S}^{-1})_{EF} v_{AB}^{EF}(r) \tag{26} $$

The summation extends to interactions including all determinants of states Φ_A and Φ_B. Using eq 18, we further find that

$$ v_c^{AB}(r) = \sum_{EF}^{N} (\mathbf{S}^{-1})_{EF} \frac{\delta}{\delta D_{AB}(r)} \mathcal{E}_c^{EF}[\mathbf{D}] \tag{27} $$

As a result, this allows us to define the *multistate correlation matrix potential*, $\mathcal{V}_c[\mathbf{D}]$, as the matrix with elements of $\{v_c^{AB}(r)\}$. Under the basis transformation with matrix $\mathbf{L}$ within

the minimal active space and using the expression of eq 27, it is straightforward to show that $\mathcal{V}_c[\mathbf{D}]$ is also transformed bilinearly,

$$\mathcal{V}_c[\mathbf{L}\mathbf{D}\mathbf{L}^\dagger] = \mathbf{L}\mathcal{V}_c[\mathbf{D}]\mathbf{L}^\dagger \tag{28}$$

Therefore, the multi-state correlation matrix potential $\mathcal{V}_c[\mathbf{D}]$ is induced by the outer dynamic correlation ($\mathcal{E}_c[\mathbf{D}]$) and is a property of the exact subspace $\mathbb{V}_{\min}$. The correlation potential also highlights the origin of density-driven and state-driven correlation functionals in ensemble DFT,[39,40] which can be understood as an artefact of ignoring state interactions, i.e., lacking the property of the above transformation. The relationship of eq 28 unifies the two classes of correlation functions on an equal footing. Naturally, we expect $\mathcal{V}_c[\mathbf{D}]$ also plays a crucial role in the generalized Fock equation for orbitals.

4.4 Generalized Fock equation for orbitals

Similarly, variation of the Lagrangian $L[\mathbf{D}]$ with respect to $\phi^\xi_{j\sigma}(r)$ leads to the generalized Fock equation for orbitals:

$$\sum_{AB}^{N}\sum_{\zeta}^{N^2}\sum_{k}^{n_\sigma}\left(c^A_\xi c^B_\zeta + c^A_\zeta c^B_\xi\right)\ {}^{o}\mathbb{F}^{AB;\sigma}_{\xi\zeta;jk}\phi^\zeta_{k\sigma}(r) = e^\xi_{j\sigma}\phi^\xi_{j\sigma}(r) \tag{29}$$

where $e^\xi_{j\sigma}$ has the meaning of the orbital energy of $\phi^\xi_{j\sigma}(r)$ and is the j-th eigenvalue of the matrix $\{\epsilon^\xi_{ik;\sigma}\}$ with $i,k = 1,\ldots,n_\sigma$ in eq 16. Analogous to the Fock matrix for the configuration coefficients (eq 24), the Fock matrix for orbitals ${}^{o}\mathbb{F}^{AB;\sigma}_{\xi\zeta;jk}$ consists of contributions from four different sources, the Hamiltonian matrix $\mathbf{H}_\Phi$ constructed in the minimal active space with $H^{AB}_\Phi = \left\langle \Phi_A|\hat{H}|\Phi_B \right\rangle$, the overlap matrix, the normalization condition (eq 16),

and the correlation matrix functional:

$$
\begin{aligned}
{}^{o}\mathbb{F}^{AB;\sigma}_{\xi\zeta;jk} =& (\mathbf{S}^{-1})_{AB} \left({}^{o}\mathbf{F}^{\sigma}_{\xi\zeta} \right)_{jk} - \left(\mathbf{S}^{-1} \mathcal{H}[\mathbf{D}] \mathbf{S}^{-1} \right)_{AB} f^{\xi\zeta}_{jk;\sigma} - \delta_{AB} E_A f^{\xi\zeta}_{ik;\sigma} \\
&+ \left[f^{\xi\zeta}_{jk;\sigma} v_c^{AB}(r) + \left(\sum_{i\neq j; l\neq k}^{n_\sigma} f^{\xi\zeta}_{ij,lk;\sigma,\sigma} \left\langle \phi^{\xi}_{i\sigma} | v_c^{AB} | \phi^{\zeta}_{l\sigma} \right\rangle + \sum_{il}^{n_{\bar{\sigma}}} f^{\xi\zeta}_{ij,lk;\bar{\sigma},\sigma} \left\langle \phi^{\xi}_{i\bar{\sigma}} | v_c^{AB} | \phi^{\zeta}_{l\bar{\sigma}} \right\rangle \right) \right]
\end{aligned}
\tag{30}
$$

Since nonorthogonal orbitals are used in constructing different determinants, the functional derivatives of various coupling constants with respect to orbitals should be addressed in the Fock equation, resulting in more complexity compared to the Fock equations based on orthogonal orbitals. The Fock matrix of the first term on the right-hand-side of eq 30, ${}^{o}\mathbf{F}^{\xi\zeta;\sigma}$, is given by the functional derivative of the Hamiltonian matrix $\mathbf{H}_\Phi$,

$$
\sum_{AB}^{N} \sum_{\zeta}^{N^2} \sum_{k}^{n_\sigma} \left(c^A_\xi c^B_\zeta + c^A_\zeta c^B_\xi \right) \left(\mathbf{S}^{-1} \right)_{AB} \left({}^{o}\mathbf{F}^{\sigma}_{\xi\zeta} \right)_{jk} \phi^{\zeta}_{k\sigma}(r) = \sum_{AB}^{N} \left(\mathbf{S}^{-1} \right)_{AB} \frac{\delta}{\delta \phi^{\xi}_{j\sigma}(r)} H_\Phi^{AB} \tag{31}
$$

The derivation of the expression for ${}^{o}\mathbf{F}^{\sigma}_{\xi\zeta}$ is given in the Supplementary Information.

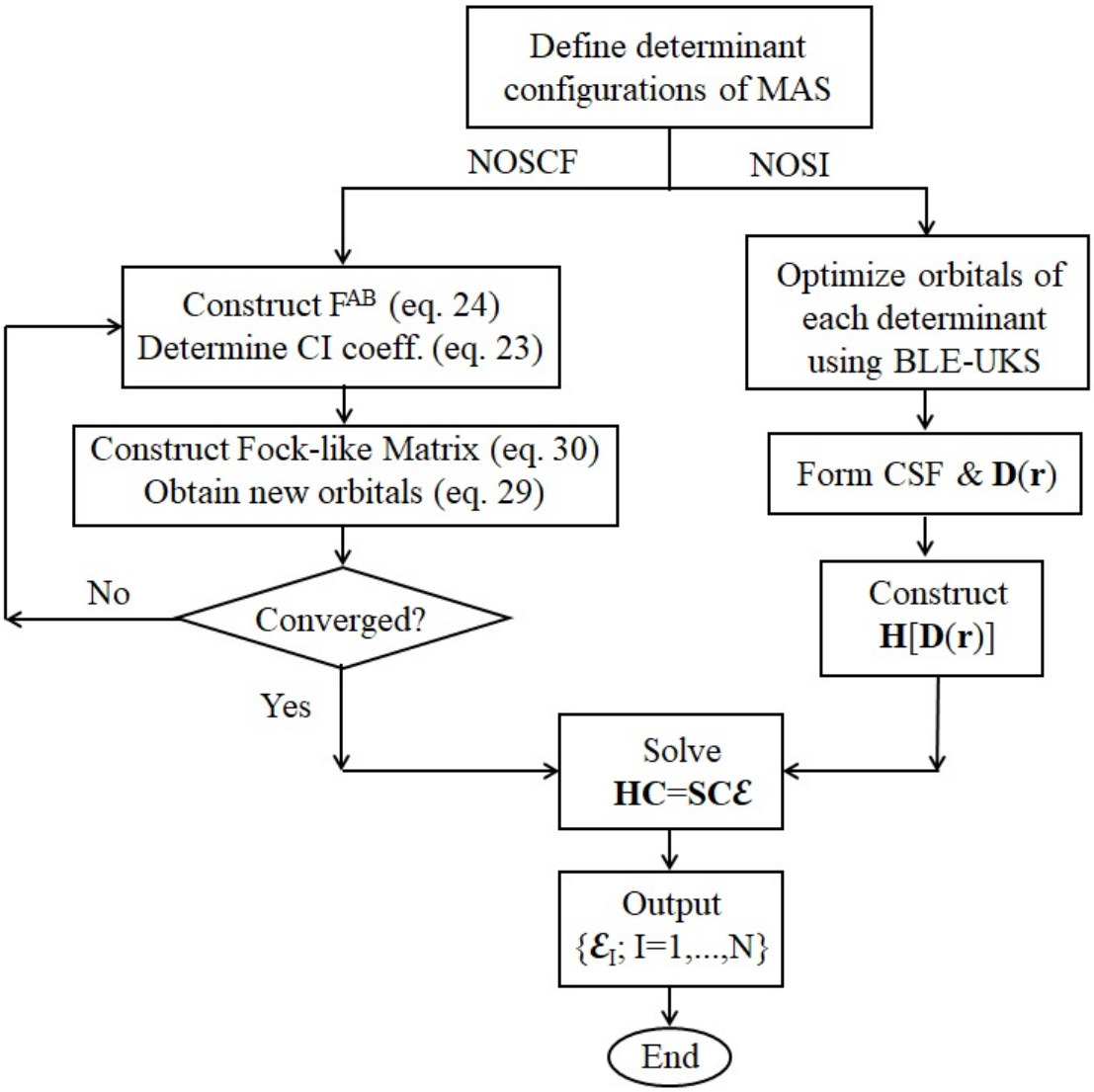
4.5 Eigenstate energies and densities

With the solutions of eqs 29 and 23, the optimized configuration states $\{\Phi_A\}$ can be used to construct the Hamiltonian matrix functional $\mathcal{H}[\mathbf{D}]$ (eq 1). The solution of the generalized secular equation of the Hamiltonian matrix functional,

$$
\mathcal{H}[\mathbf{D}]\mathbf{C} = \mathbf{S}\mathbf{C}\mathcal{E}_0 \tag{32}
$$

gives the N-lowest adiabatic energies $\mathcal{E}_0 = \mathrm{diag}\,(E_1, \cdots, E_N)$ and state vectors that give the exact densities $\{\rho_I(r)\}$ of the individual eigenstates.

In summary, a flowchart of the NOSCF computational procedure and that of the NOSI method below is given in Scheme 1.



Scheme 1: Flowchart illustrating key computational steps for the non-orthogonal self-consistent-field (NOSCF) and non-orthogonal state interaction (NOSI) methods. Acronyms are listed at the end of the article.

5 Nonorthogonal state interaction

In wave function theory, a multiconfigurational (MC) wave function can be either optimized self-consistently (MC-SCF) or determined by configuration interaction (CI). In MS-DFT, the same two routes can be taken. In the latter case (Scheme 1), each determinant of the MAS is variationally optimized first, followed by a single step of diagonalization, and this approach is called nonorthogonal state interaction (NOSI). NOSI differs from NOCI in WFT in that dynamic correlation is included in the first place.[14,38] Although the configuration coefficients are not fully optimized as in MS-SCF (Scheme 1), an important benefit of the NOSI approach is to produce variationally optimized diabatic configuration-states (VDC).[41] The latter is more useful and relevant to energy decomposition analysis,[42,43] and to studying charge transfer reactions and excited-state energy transfer processe[44–46] than the consistently produced basis states (CDC) from MS-SCF.[41]

5.1 Diagonal elements of the Hamiltonian matrix functional

In NOSI, we use the determinant states $\{\Phi_A\}$ directly, which are defined by a constrained KS-DFT for a particular set of non-aufbau configurations.[34,47] This may be formally considered as applying a constraining potential that forces the determinant to keep in a specific pattern of orbital occupation.[48] We argue that the use of an approximate density functional developed for KS-DFT to determine the energies of these states, corresponding to the diagonal elements of the Hamiltonian matrix functional $\mathcal{H}[\mathbf{D}]$, is appropriate because a single determinant is used in constrained ground-state optimization.[14,26,49]

$$H_{\Phi}^{AA} = E_{KS}[\rho_A(r)] \tag{33}$$

where $E_{KS}[\rho_A(r)]$ is the KS-DFT energy obtained using the density from a single-determinant state Φ_A. Each of these non-aufbau states can be variationally optimized either by the block-local excitation (BLE) method – a form of local Δ-SCF calculation[47] – or by the

targeted-state optimization (TSO) approach.[34] In this regard, we point out that the constrained DFT optimizations in the BLE and TSO methods are performed in the orbital space,[14,26,27] which is different from that by spatial partition of the electron density.[50]

In MS-SCF calculations, individual basis states, $\{\Phi_A\}$, are expressed as a linear combination of N^2 determinants (eq 7), but Kohn-Sham exchange-correlation functional is applied at the level of each individual determinant state, rather than a multi-configurational density.[51–55] Therefore, in principle, approximate density functionals for KS-DFT can also be used. However, it might be necessary to adjust the correlation energy in such an MS-SCF optimization.[51] We have examined the Fermi-level scaled approach proposed by Savin and coworkers with encouraging results in preliminary tests.[56] More thorough validation is certainly needed in future studies, and perhaps, a multistate approximation to the correlation matrix functional is really desired and a focus in the future.

5.2 Transition density functional

As in KS-DFT, the exact multistate correlation matrix functional $\mathcal{E}_c[\mathbf{D}]$ is unknown. Unlike KS-DFT for which many approximate functionals have been developed in the past six decades, exhibiting excellent performance in present-day applications, there is currently no approximate matrix functional for subspace correlation among individual states. In fact, the very concept of a transition density functional (TDF) for the off-diagonal elements of $\mathcal{E}_c[\mathbf{D}]$ is new, accounting for the dynamic correlation in state interactions.[14]

Equation 13 shows that $\mathcal{E}_c[\mathbf{D}]$ is a property of the subspace $\mathbb{V}_{\min}$ itself, expressed in terms of the full matrix density $\mathbf{D}(r)$ and the corresponding set of basis states. Recall that the matrix functional relationship is not a simple element-by-element mapping between $\mathcal{E}_c^{AB}$ and $D_{AB}(r)$. That is, the total correlation in eq 13 is unique as far as the N-dimensional subspace $\mathbb{V}_{\min}$ is concerned, but its matrix functional values are dependent on the basis state transformation relationship of eq 4, also specifically shown for the correlation potential in eq 28.

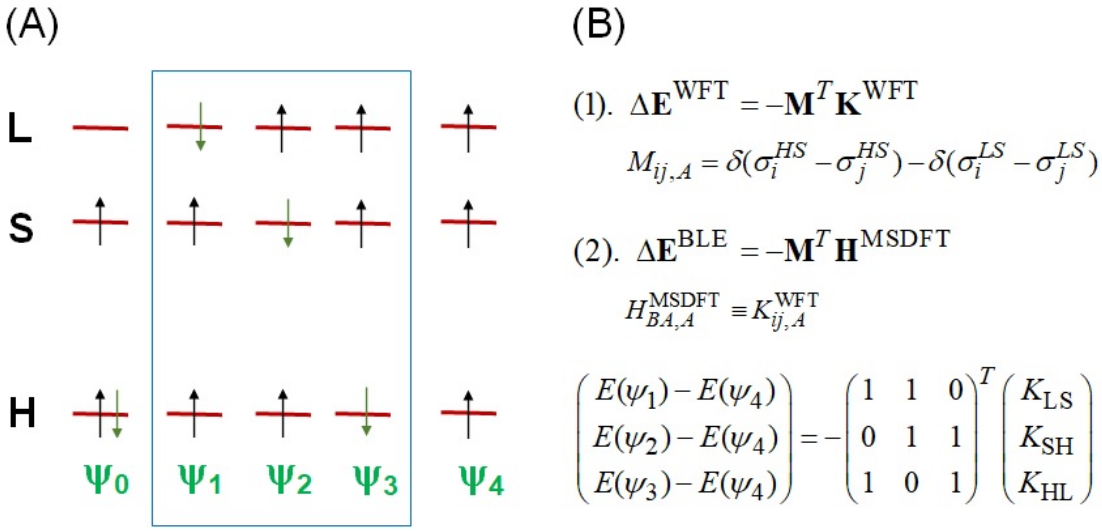


Figure 1: Transition density functional for spin-coupling interactions in the $M_S = 1/2$ manifold of quartet and doublet states. (A). Schematic illustration of the ground and **H** $\rightarrow$ **L** excitation ($M_S > 0$ components are shown). (B). Conditions for spin multiplet degeneracy of the $M_S = 1/2$ and $M_S = 3/2$ states. In wave function theory (1), the matrix element K_{ij} is the exchange integral for orbitals i and j ($i, j =$ **L,S,H**), δ is the Kronecker delta function, and σ_i is the spin coordinate of an electron in orbital i where the superscripts indicate the low-spin and high-spin configurations. For multistate density functional theory (2), the wave functions are the block-localized excited (BLE) Kohn-Sham determinants ($A = 1, \cdots, 4$) and the orbital indices in (1) for the exchange integrals are changed to the corresponding BLE-determinants for the off-diagonal matrix element of the Hamiltonian matrix functional, corresponding to the spin flip, $H_{BA}^{MSDFT} = K_{ij}^{WFT}$. ΔE is an energy difference between one of the low-spin configurations ($\psi_1 - \psi_3$) and that of the high-spin configuration (ψ_4).

Although the functional form of the TDF is unknown, in special situations such as electron spin-pairing interactions, the TDF correlation energy can be fully determined consistently with the energy of a high-spin state that can be obtained,[36,37,57] in principle exactly, using KS-DFT.[58] This is achieved by enforcing the degeneracy condition of spin multiplets of the same total spin S.[59] Let $\mathbf{K}^{MSDFT}$ be a column vector with the off-diagonal matrix-element H_{Φ}^{AB} placed in sequence and $\Delta\mathbf{E}^{BLE}$ be a column vector of the KS-DFT energy difference between the high-spin state $\Psi^{BLE}(S+1)$ and the mixed-spin microstates $\{\Psi_A^{BLE}(S)\}$, arranged in order consistent with the matrix $\mathbf{M}$ to yield the effective TDF terms $\mathbf{K}^{MSDFT}$ in MSDFT (see Figure 1).[37] Then, we have the relationship that ensures all spin multiplets of spin S be energy-degenerate.[37]

$$\mathbf{K}^{MSDFT} = -(\mathbf{U}\sigma^{-1}\mathbf{V}^{T})\Delta\mathbf{E}^{BLE} \tag{34}$$

where $\mathbf{U}\sigma^{-1}\mathbf{V}^{T}$ is a singular value deomposition of the matrix $\mathbf{M}$ that defines non-zero elements of the exchange integrals between the determinant $\Psi^{BLE}(S+1)$ of the high-spin configuration (all-spin up) and all the microstate determinants $\{\Psi_A^{BLE}(S)\}$ with spin S.[37] Thus, provided KS-DFT is adequate with a single determinante for the highest spin state (all-spin up of unpaired electrons), which is,[58] the TDF energies (off-diagonal terms of $\mathcal{H}_{\Phi}[\mathbf{D}]$) are consistently defined if the same approximate density functional is used to obtain the energies of the diagonal terms (eq 33). If nonorthogonal determinants are used, a Lowdin transformation may be applied to ensure that the parameters in eq 34 are consistent with state orthogonality.[37] For further details, see the original references.[36,37]

For example, the TDF correlation energy for the spin coupling interactions between two unpaired electrons, resulting in a singlet state and the $M_S = 0$ multiplet of the triplet state, can be determined by enforcing the energy degeneracy condition between the multiconfiguration $M_S = 0$ state and the $M_S = 1$ triplet states. The latter ($E_{KS}(|1,1>)$ can be

adequately determined by KS-DFT with a single determinant.[36]

$$\mathcal{E}_c^{AB}[\Psi_{\downarrow\uparrow}, \Psi_{\uparrow\downarrow}] = \mathcal{E}_c^{\mathrm{KS}}[\rho(\Psi_{\uparrow\uparrow})] - \mathcal{E}_c^{\mathrm{KS}}[\rho(\Psi_{\uparrow\downarrow})] \tag{35}$$

where $\Phi_{\downarrow\uparrow}$ and $\Phi_{\uparrow\downarrow}$ are determinants A and B with different spin combinations of the two spin-coupled electrons, and $\mathcal{E}_c^{\mathrm{KS}}[\rho(\Phi_{\uparrow\uparrow})]$ and $\mathcal{E}_c^{\mathrm{KS}}[\rho(\Phi_{\uparrow\downarrow})]$ are energies using a KS-DFT correlation functional for the high-spin (triplet) determinant and one of the spin-mixed determinant. Here, an extra calculation is needed to determine the high-spin energy. Equation 35 can be obtained by inversion of the $\mathbf{M}$ matrix directly without the need for singular value decomposition (though only KS-correlation energies are included in the elements of $\Delta\mathbf{E}^{BLE}$).[37]

In other cases, we have found that the approximation using an overlap-weighted Kohn-Sham correlation energy of the two interacting determinants to yield reasonable results in many applications.[14,27]

$$\mathcal{E}_c^{AB}[\mathbf{D}] = \frac{1}{2} S_{AB}(E_c^{\mathrm{KS}}[\rho_A] + E_c^{\mathrm{KS}}[\rho_B]) \tag{36}$$

In eq 36, $E_c^{\mathrm{KS}}[\rho_A]$ and $E_c^{\mathrm{KS}}[\rho_B]$ are the correlation energies for states A and B, determined by the KS-DFT correlation functional that is used to model the diagonal terms. In fact, eq 36 is the leading term in the matrix correlation function under the local density approximation.[60]

6 Illustrative examples

In this section, we illustrate two examples that are not easily treated, if not impossible, by using conventional KS-DFT and linear response TDDFT: (1) a conical intersection between the ground and the first excited state by MS-SCF, and (2) the spin-coupling interactions between an excited chromophore and a free radical species by NOSI.

Table 1: Summary of Computational Details for the Conical Intersection of Ammonia Photodissociation (NH_3) and the Anthraquinone and Nitroxide (AQ-NO) complex.

System	CSF[1]	Determinants[2]	Orbitals[3]	CI-Opt[4]	Method[5]
NH_3	6	9	Orth.	Yes	ROKS
AQ-NO	4	5	Non-orth.	No	gs: KS-DFT es: BLE/UKS

1. Configuration-State Functions; 2. Number of determinants included; 3. Type of orbitals: orthogonal (orth.) or non-orthogonal (non-orth.); 4. CI coefficients optimization; 5. ROKS: restricted open-shell Kohn-Sham DFT; gs: ground-state optimized by KS-DFT; es: non-Aufbau determinant optimized by block-localized excitation with unrestricted KS-DFT (BLE-UKS).

6.1 Photodissociation of ammonia

The photochemical process, $NH_3 \rightarrow NH_2$ + H, is a classic example that has been extensively studied (see references in refs. 61 and 62).[61,62] At the crossing point between the $NH_2(^2B_1)$ and $NH_2(^2A_1)$ states along the bond dissociation coordinate in the molecular plane, bending about the molecular plane lifts the energy degeneracy, resulting in a conical intersection with a double cone feature in the two potential energy surfaces for the ground state and the first excited state. Previously, we have used NOSI to determine the potential energy curves along the N$-$H stretch coordinate, employing a set of twelve valence bond configurations.[14] Here, to illustrate the MS-SCF method in MSDFT calculations, we reconstructed the potential energy surfaces using delocalized molecular orbitals in an active space consisting of 4 electrons and three orbitals, rather than the localized valence bond states previously.[14] Since we have not fully implemented the nonorthogonal optimization algorithm, the present calculations are performed employing a common set of orthogonal orbitals, which are adequate in view of the small size of the orbital space. Here, a total of six spin-adapted configurations (nine determinants) are included (Table 1), and this example shows that MS-SCF can be conveniently adapted into any code with MC-SCF (CASSCF) capability.

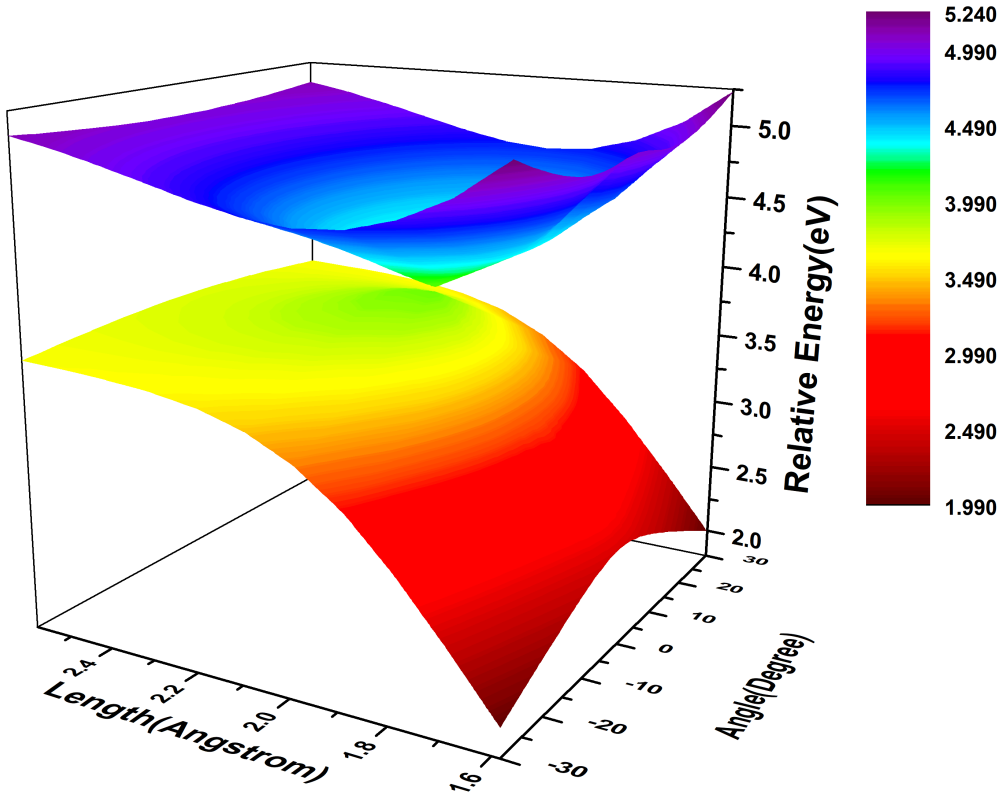


Figure 2: Potential energy surfaces of the ground and first excited states as a function of the N−H bond-dissociation coordinate (in Å) and its bending angle about the plane of the NH_2 group (in degrees). The PBE approximate functional is used along with the cc-pVTZ basis set. An active space consisting of four electrons and three orbitals is used with a total of five configuration-state functions in multistate self-consistent-field (MS-SCF) optimizations. Energies are given in eV.

Figure 1 displays the potential energy surfaces for the ground and first excited states in the bond-dissociation coordinate along one of the N−H bond and its bending angle from the plane of the NH_2 group. The conical intersection point occurs at a distance of 2.001 Å with an energy of 5.00 eV above that of ammonia, which may be compared with values of 1.990 Å and 5.16 eV in the full-dimension potential energy surface.[61] The topology of the PES about the conical intersection point is correctly produced from MS-SCF optimizations.

Figure 1 shows that the use of a commons set of orbitals in MS-SCF optimization is adequate for the ammonia bond dissociation process in the S_1 state. There are many benefits to use orthogonal orbitals in MS-SCF calculations over that with nonorthogonal orbitals since the computational efforts and costs will be significantly reduced. Indeed, preliminary results indicate that it is possible to define a finite active space to fully represent the exact matrix density $\mathbf{D}(r)$ with orthonormal orbitals, but the number of determinant configurations will be more than that by using nonorthogonal determinants.

6.2 Spin coupling of anthraquinone and nitroxide free radical

Management of spin state in spintronic materials can be useful to affect display efficiency of organic light-emitting diode, among other applications.[63,64] For example, an aromatic compound such as benzene and pyridine can be directly excited to the triplet state in the presence of a free radical species such as oxygen and NO,[65,66] whereas the triplet state of closed-shell porphyrins can become highly luminescent with an unpaired metal d-electron.[67] In a recent biomolecular application of the spintronic coupling between 9,10-anthraquinone and a stable nitroxide radical (Figure 2A),[68] it was found that its spin state can be controlled by visible light through DNA binding. The S_1 and T_1 excited states of 9,10-anthraquinone can strongly couple with the nearby nitroxide radical (D_0) in the absence of DNA, resulting in a quartet (Q_1) state and a doublet state (Figure 3B)[68] termed as tripdoublet (D_1), plus another high-energy singdoublet (not shown).[67] Intercalation

into a double stranded DNA structure causes the nitroxide free-radical fragment from forming a folded complex with the anthraquinone unit, rendering disappearance of the quartet state from EPR experiments to give a distantly coupled T_1 and D_0 combination.[68]

Employing a minimal active space shown in Figure 3B, we determined the spin coupling interactions among the T_1 and S_1 excited states of 1-chloroanthraquinone (**1**) and the D_0 ground state of the model nitroxide, **2**, (Figure 3C), involving three electrons in three orbitals. The BLE method was used to optimize the orbitals of each individual microstates plus the spin all-up quartet state ($|3/2, 3/2>$). The latter was used to enforce the energy degeneracy of the multiconfigurational $|3/2, 1/2>$ multiplet of the Q_1 state (see Figure 1 and text).[37,57] In addition, we included a charge-transfer configuration, $\mathbf{A}^{-\cdot}(D_0)\mathbf{N}^{+}(S_0)$. A total of five microstates are included in the NOSI calculation (Figure 3B).

We first optimized the bimolecular complex between 1-chloro-9,10-anthraquinone (**1**) and the nitroxide free radical **2** in the doublet ground state, which has a distance of about 3.5 Å between the nitrogen atom of **2** and the C4 atom of **1** (Figure 3C, the optimized structure has roughly a C_s symmetry). The computed binding energy is -9.0 kcal/mol from M06-2X/cc-pVDZ optimization, whereas NOSI@M06-2X/cc-pVDZ single-point energy calculations lowers the binding energy by 1.2 kcal/mol (0.05 eV) to -10.2 kcal/mol, showing a small amount of multiconfiguraiton character in the doublet ground-state complex. The singlet and triplet excited states due to HOMO $\rightarrow$ LUMO excitation of **1** is strongly spin-coupled with the doublet state of **2**.[57,67,68] Consequently, the T_1 state is split into a quartet (Q_1) state and a tripdoublet with an energy gap of 22.1 kcal/mol (0.96 eV). The vertical excitation energy of the spin-coupled quartet state (Q_1) is 67.1 kcal/mol (2.91 eV), slightly greater than the vertical T_1 excitation energy of **1** (66.2 kcal/mol or 2.87 eV); the small difference is largely due to ground-state effect. For comparison, the T_1(**1**) energy is in reasonable agreement with an experimental value of 62.4 kcal/mol (2.75 eV) for the triplet state of chloroanthraquinone.[69] The tripdoublet-state energy, however, is significantly increased to 89.2 kcal/mol (3.87 eV), largely responsible for the computed

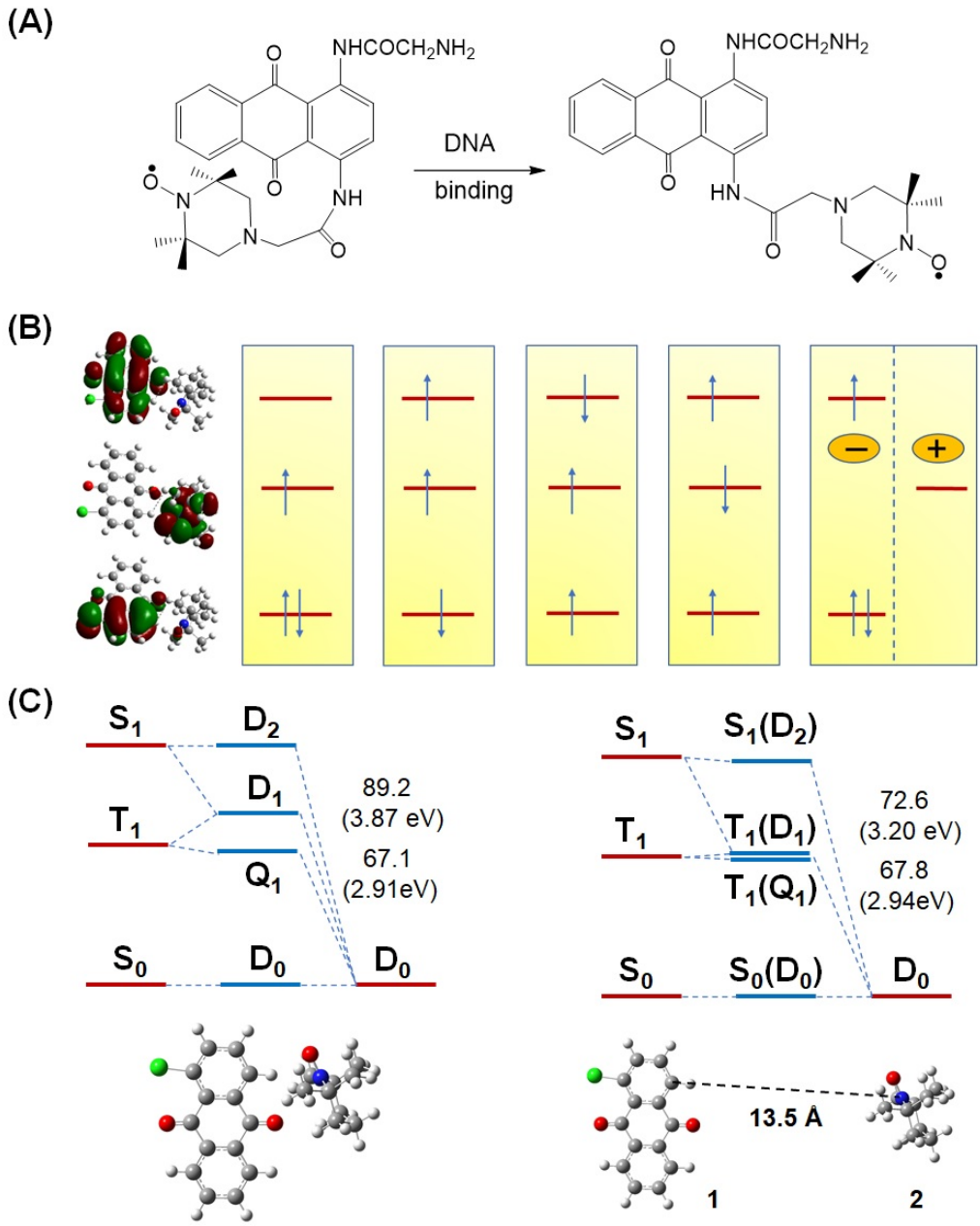


Figure 3: Illustration of (A) the change in molecular conformation of an anthraquinone-nitroxide compound upon binding to a segment of double-strand DNA, (B) orbitals and microstates used in nonorthogonal state interaction (NOSI) calculations for 1-chloroanthraquinone (**1**) and an organic nitroxide model (**2**), and (C) strengths of spin coupling in the bimolecular complex and at 13.5 Å separation between the C4 carbon of **1** and the nitrogen atom of **2**. The vertical line in the charge-transfer structure of (B) indicates fragment block localization. Energies are given in kcal/mol or electronvolts in parentheses relative to the doublet ground state in each structure.

quartet-doublet splitting gap. As the two molecular fragments are pulled apart by 10 Å to 13.5 Å between the two distinguished atoms above, the energy splitting is reduced to 5.8 kcal/mol, still significant in view of the large intermolecular distance. As in the complex structure, the energy of the Q_1 state for the separated structure is nearly unchanged, but the tripdoublet is now lowed to 72.6 kcal/mol (3.20 eV).

We found that the $\mathbf{2} \rightarrow \mathbf{1}$ charge-transfer state becomes lower in energy than the tripdoublet at 84.6 kcal/mol (3.67 eV) in the molecular complex, which is the lowest spin-coupled doublet state (local singlet for **1**). The energy of the charge-transfer state is increased to 119.2 kcal/mol (5.17 eV) as the two molecular fragments are separated to 13.5 Å. Surprisingly, in both cases, the charge-transfer state has little effects on the energies of the covalent excited states. The experimental work did not report absorption or emission spectra of the spintronic system in solution; the authors focused on analysis of EPR spectra to demonstrate the formation of a quartet state in the absence of DNA, and the loss of its signal when a segment of DNA was introduced.[68] These features are reflected in Figure 3.

7 Summary

Whereas density functional theory has become a method of choice for electronic structure calculations of large molecules in ground state, excitation energies are obtained in the framework of linear-response time-dependent approaches. In this Perspective, we discuss the fundamental theory, computational methods, and optimization techniques of a time-independent DFT for the lowest N states, and highlight two applications to illustrate the capability of such a multistate density functional theory.

We first present the theorems on the existence of a Hamiltonian matrix functional and the variational optimization of the multistate energy for a subspace spanned by the N states. Importantly, the minimization of the multistate energy leads to the solutions

for the exact energies and densities of the N-lowest eigenstates, provided the universal Hamiltonian matrix functional is known. The method enables the ground and excited states to be treated on an equal footing.

Based on the third theorem of reference 8 for representation of the multistate matrix density $\mathbf{D}(r)$,[8] we introduce a minimal active space for the N-dimensional subspace, consisting of no more than N^2 nonorthogonal Slater determinants. This is possible because the "active space", unlike that in WFT, is for the purpose of representing $\mathbf{D}(r)$ rather than searching for the exact wave functions of the N-lowest eigenstates. The latter is only possible in the limit of full configuration interaction, while the number of variational variables for representing the N-matrix density is finite and limited by the dimension of the subspace. Given a multistate MAS, we can define the matrix functionals for the kinetic energy, Coulomb-exchange energy and the external potential energy, which leads to the introduction of correlation matrix functional. Clearly, development of approximate matrix functionals suitable for multistate calculations is essential.

Two computational approaches are discussed. First, a set of nonorthogonal self-consistent-field equations (NOSCF) are presented to optimize the orbitals and configuration coefficients for a set of N auxiliary states constructed from N^2 Slater determinants. Alternatively, in nonorthogonal state interaction (NOSI), each determinant is separately optimized, followed by a single step of diagonalization. In both cases, state interaction is essential. While NOSCF yields the exact solution, NOSI provides variational diabatic states that can be used in simulations of electron transfer and excited-state energy transfer processes.

We also present two illustrative cases, one on the conical intersection between the ground and the first excited state in the photodissociation of NH_3 and the other on spin coupling between a triplet excited state and a ground-state radical. Both examples are not easily treated with standard time-dependent DFT methods for lacking, respectively, of proper dimension in state interaction and of double excitation. The results are promising.

Of course, there are several challenges and limitations that need to be overcome to gain broad applications.

1. An obvious choice to approximate the diagonal elements of the correlation matrix functional is to directly use one of the approximate density functionals developed for KS-DFT. However, its scope of applicability beyond the ground state needs to be carefully examined in view of the possibility of introducing double counting of correlation.[14,51]

2. There is currently no explicit functional form to approximate the transition density functional since the very concept is new. It is of great interest to develop a complete and consistent correlation matrix functional to treat both the ground and excited states.

3. The development of efficient optimization methods, especially with the use of nonorthogonal orbitals, is needed. The future development of MSDFT models can certainly benefit from the recent spurt of studies on optimization techniques in NOCI-based methods, for instance, the construction of reference determinants and SCF optimization of nonorthogonal orbitals.[41,44,70–77]

4. It will be useful to implement analytical gradient techniques and to develop dynamics simulation methods to treat nonadiabatic processes.

With the proof of the fundamental theorems for an excited-state density functional theory, we hope that this Perspective could stimulate further developments of multistate DFT for applications.

List of Acronyms

AQ-NO: anthraquinone-nitroxide

BLE: block-local excitation

CASSCF: complete-active-space self-consistent-field

CDC: consistent diabatic state

CI: configuration interaction

CSF: configuration-state function

$\Delta-$SCF: delta-self-consistent-field

DFT: density functional theory

HK: Hohenberg and Kohn

HOMO: highest occupied molecular orbital

KS-DFT: Kohn-Sham density functional theory

LG: Lu and Gao

LR: linear response

LUMO: lowest unoccupied molecular orbital

MAS: minimal active space

MC-SCF: multi-configurational self-consistent-field

MSDFT: multistate density functional theory

MS-SCF: multistate self-consistent-field

NOSCF: non-orthogonal self-consistent-field

NOSI: non-orthogonal state interaction

ROKS: restricted open-shell Kohn-Sham

SCF: self-consistent-field

TDF: transition density functional

TDDFT: time-dependent density functional theory

UKS: unrestricted Kohn-Sham

VDC: variational diabatic state

WFT: wave function theory

Appendices

A.1 Expressions of coefficients

The coefficient $f_{jk;\sigma}^{Ap,Bq}$ is given by the coupling between two determinants with $(n-1)$ orbitals, the determinant Ξ_X less of orbital $\phi_{j\sigma}^X$ and Ξ_Y less of orbital $\phi_{k\sigma}^Y$ (with $X = Ap, Y = Bq$),

$$f_{jk;\sigma}^{XY} = \frac{(-1)^{j+k}}{(n-1)!} \int ||\phi_{1\sigma_1}^X \cdots \hat{\phi}_{j\sigma}^X \cdots \phi_{n\sigma_n}^X|| ||\phi_{1\sigma_1}^Y \cdots \hat{\phi}_{k\sigma}^Y \cdots \phi_{n\sigma_n}^Y|| \tag{A.37}$$

Similarly, the coefficient $f_{jk;\sigma}^{Ap,Bq}$ is given by the coupling between two determinants with $(n-2)$ orbitals, the determinant Ξ_X less of orbitals $\phi_{i\sigma}^X, \phi_{j\sigma'}^X$ and Ξ_Y less of orbitals $\phi_{k\sigma}^Y, \phi_{l\sigma'}^Y$ (with $X = Ap, Y = Bq$),

$$f_{ij,kl;\sigma,\sigma'}^{XY} = \frac{(-1)^{i+j+k+l}}{(n-2)!} \int ||\phi_{1\sigma_1}^X \cdots \hat{\phi}_{i\sigma}^X \cdots \hat{\phi}_{j\sigma'}^X \cdots \phi_{n\sigma_n}^X|| ||\phi_{1\sigma_1}^Y \cdots \hat{\phi}_{k\sigma}^Y \cdots \hat{\phi}_{l\sigma'}^Y \cdots \phi_{n\sigma_n}^Y|| \tag{A.38}$$

The two-electron Coulomb-exchange integral $\langle ij||kl\rangle_{Ap,Bq}$ reads: if $\sigma = \sigma'$,

$$\langle ij||kl\rangle_{Ap,Bq} = \frac{e^2}{4\pi\epsilon_0} \int \mathrm{d}r_1 \mathrm{d}r_2 \, \frac{\phi_{i\sigma}^X(r_1)\phi_{j\sigma}^X(r_2)}{|r_1 - r_2|} \left(\phi_{k\sigma}^Y(r_1)\phi_{l\sigma}^Y(r_2) - \phi_{l\sigma}^Y(r_1)\phi_{k\sigma}^Y(r_2)\right) \tag{A.39}$$

and if $\sigma \neq \sigma'$,

$$\langle ij||kl\rangle_{Ap,Bq} = \frac{e^2}{4\pi\epsilon_0} \int \mathrm{d}r_1 \mathrm{d}r_2 \, \frac{\phi_{i\sigma}^X(r_1)\phi_{j\sigma'}^X(r_2)}{|r_1 - r_2|} \phi_{k\sigma}^Y(r_1)\phi_{l\sigma'}^Y(r_2) \tag{A.40}$$

Supporting Information Available

Two documents containing proofs and derivation of equations (11 pages), and additional computational details (9 pages). The Supporting Information is available free of charge at https://pubs.acs.org/doi/.

Author Information

Corresponding Authors.

Yangyi Lu - Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518055, China; Orcid, https://orcid.org/0000-0002-1602-1661; Email: luyy@szbl.ac.cn

Meiyi Liu - Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518055, China; Email: liumy@szbl.ac.cn

Jiali Gao - Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518055, China; Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, United States; Orcid, https://orcid.org/0000-0003-0106-7154; Email: gao@jialigao.org

Authors.

Ruoqi Zhao - Institute of Theoretical Chemistry, Jilin University, Changchun, Jilin Province 130023, China; Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518055, China; Email: zhaorq1214@mails.jlu.edu.cn

Jun Zhang - Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen 518055, China; Email: zhangjun@szbl.ac.cn

Notes

The authors declare no competing financial interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgement

This work has been partially supported by Shenzhen Municipal Science and Technology Innovation Commission (KQTD2017-0330155106581), the Key-Area Research and Development Program of Guangdong Province (Grant 2020B0101350001), and Shenzhen Bay Laboratory (SZBL). YL is a Qihang Scholar of the SZBL. Work performed at Minnesota was supported by the National Institutes of Health (GM046736). Discussion with Professor L. M. Thompson has been most helpful.

References

(1) Kohn, W.; Sham, L. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(2) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864–B871.

(3) Becke, A. D. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.* **2014**, *140*, 1–18.

(4) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. T. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289320.

(5) Teale, A. M.; Helgaker, T.; Savin, A.; Adamo, C.; Aradi, B.; Arbuznikov, A. V.; Ayers, P. W.; Baerends, E. J.; Barone, V.; Calaminici, P.; Cancès, E.; Carter, E. A.; Chattaraj, P. K.; Chermette, H.; Ciofini, I.; Crawford, D.; Proft, F. D.; Dobson, J.; Draxl, C.; Frauenheim, T.; Fromager, E.; Fuentealba, P.; Gagliardi, L.; Galli, G.; Gao, J.; Geerlings, P.; Gidopoulos, N.; Gill, P. W.; Gori-Giorgi, P.; Görling, A.; Gould, T.; Grimme, S.; Gritsenko, O.; Jensen, H. J. A.; Johnson, E. R.; Jones, R. O.; Kaupp, M.; Köster, A. M.; Kronik, L.; Krylov, A. I.; Kvaal, S.; Laestadius, A.; Levy, M.; Lewin, M.;

Liu, S.; Loos, P.-F.; Maitra, N. T.; Neese, F.; Perdew, J. P.; Pernal, K.; Pernot, P.; Piecuch, P.; Rebolini, E.; Reining, L.; Romaniello, P.; Ruzsinszky, A.; Salahub, D. R.; Scheffler, M.; Schwerdtfeger, P.; Staroverov, P. N. V.; Sun, J.; Tellgren, E.; Tozer, D. J.; Trickey, S. B.; Ullrich, C. A.; Vela, A.; Vignale, G.; Wesolowski, T. A.; Xu, X.; Yang, W. DFT Exchange: Sharing Perspectives on the Workhorse of Quantum Chemistry and Materials Science. *Phys. Chem. Chem. Phys.* **2022**, 1–85.

(6) Marques, M. A. L.; Ullrich, C. A.; Nogueira, F.; Rubio, A.; Burke, K.; Gross, E. K. U. *Time-dependent density Functional theory*; The Lecture Notes in Physics; Springer, 2006; p 603.

(7) Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. Conical intersections and double excitations in time-dependent density functional theory. *Mol. Phys.* **2006**, *104*, 1039–1051.

(8) Lu, Y.; Gao, J. Multistate Density Functional Theory of Excited States. *J. Phys. Chem. Lett.* **2022**, *13*, 7762–7769.

(9) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. A. Complete Active Space SCF Method (CASSCF) Using a Density Matrix Formulated Super-CI Approach. *Chem. Phys.* **1980**, *48*, 157–173.

(10) Lischka, H.; Nachtigallova, D.; Aquino, A. J. A.; Szalay, P. G.; Plasser, F.; Machado, F. B. C.; Barbatti, M. Multireference Approaches for Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7293–7361.

(11) Andersson, K.; Áke Malmqvist, P.; Roos, B. O.; Sadlej, A. J.; Wolinsk, K. Second-Order Perturbation Theory with a CASSCF Reference Function. *J. Phys. Chem.* **1980**, *94*, 5483–5488.

(12) Soichi Shirai, Y. K.; Yanai, T. Computational Evidence of Inversion of L-1(a) and L-1(b)-Derived Excited States in Naphthalene Excimer Formation from ab Initio Mul-

tireference Theory with Large Active Space: DMRG-CASPT2 Study. *J. Chem. Phys.* **2016**, *12*, 2366–2372.

(13) Veryazov, V.; Malmqvist, P. A.; Roos, B. O. How to Select Active Space for Multiconfigurational Quantum Chemistry? *Int. J. Quantum Chem.* **2011**, *111*, 3329–3338.

(14) Gao, J.; Grofe, A.; Ren, H.; Bao, P. Beyond Kohn-Sham Approximation: Hybrid Multistate Wave Function and Density Functional Theory. *J. Phys. Chem. Lett.* **2016**, *7*, 5143–5149.

(15) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press, New York, 1989; p 352.

(16) Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **2015**, *87*, 897–923.

(17) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.

(18) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(19) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(20) Mermin, N. D. Thermal Properties of the Inhomogeneous Electron Gas. *Phys. Rev.* **1965**, *137*, 1441–1443.

(21) Theophilou, A. K. The energy density functional formalism for excited states. *J. Phys. C Solid State Phys.* **1979**, *12*, 5419–5430.

(22) Gross, E. K. U.; Oliveira, L. N.; Kohn, W. Rayleigh-Ritz variational principle for ensembles of fractionally occupied states. *Phys. Rev. A* **1988**, *37*, 2805–2808.

(23) Gross, E. K. U.; Oliveira, L. N.; Kohn, W. Density-functional theory for ensembles of fractionally occupied states. I. Basic formalism. *Phys. Rev. A* **1988**, *37*, 2809–2820.

(24) Görling, A. Density-functional theory for excited states. *Phys. Rev. A - At. Mol. Opt. Phys.* **1996**, *54*, 3912–3915.

(25) Ayers, P. W.; Levy, M.; Nagy, A. Time-independent density-functional theory for excited states of Coulomb systems. *Phys. Rev. A - At. Mol. Opt. Phys.* **2012**, *85*, 1–7.

(26) Cembran, A.; Song, L.; Mo, Y.; Gao, J. Block-Localized Density Functional Theory (BLDFT), Diabatic Coupling, and Their Use in Valence Bond Theory for Representing Reactive Potential Energy Surfaces. *J. Chem. Theory Comput.* **2009**, *5*, 2702–2716.

(27) Ren, H.; Provorse, M. R.; Bao, P.; Qu, Z.; Gao, J. Multistate Density Functional Theory for Effective Diabatic Electronic Coupling. *J. Phys. Chem. Lett.* **2016**, *7*, 2286–2293.

(28) McWeeny, R. Some recent advances in density matrix theory. *Rev. Mod. Phys.* **1960**, *32*, 335–369.

(29) Landau, L. D.; Lifshitz, E. M. *Quantum Mechanics Non-relativistic theory*; Course of Theoretical Physics; Pergamon Press, 1977; Vol. 3; p 679.

(30) Hall, B. C. *Lie groups, Lie algebras and representations: An elementary introduction*; Graduate Texts in Mathematics; Springer, 2015; p 453.

(31) Higham, N. J. *Functions of matrices theory and computation*; Society for Industrial and Applied Mathematics, Philadelphia, 2008; p pp. 445.

(32) Levy, M. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 6062–6065.

(33) Theophilou, A. K.; Papaconstantinou, P. G. Local spin-density approximation for spin eigenspaces and its application to the excited states of atoms. *Phys. Rev. A* **2000**, *61*, 1–10.

(34) Grofe, A.; Zhao, R.; Wildman, A.; Stetina, T. F.; Li, X.; Bao, P.; Gao, J. Generalization of Block-Localized Wave Function for Constrained Optimization of Excited Determinants. *J. Chem. Theory Comput.* **2021**, *17*, 277–289.

(35) Jacob, C. R.; Reiher, M. Spin in density-functional theory. *Int. J. Quant. Chem.* **2012**, *112*, 3661–3684.

(36) Grofe, A.; Chen, X.; Liu, W.; Gao, J. Spin-multiplet components and energy splittings by multistate density functional theory. *J. Phys. Chem. Lett.* **2017**, *8*, 4838–4845.

(37) Zhao, R.; Grofe, A.; Wang, Z.; Bao, P.; Chen, X.; Liu, W.; Gao, J. Dynamic-then-Static Approach for Core Excitations of Open-Shell Molecules. *J. Phys. Chem. Lett.* **2021**, *12*, 7409–7417.

(38) Liu, W.; Hoffmann, M. R. SDS: the 'static–dynamic–static' framework for strongly correlated electrons. *Theor. Chem. Acc.* **2014**, *133*, 1481.

(39) Gould, T.; Pittalis, S. Density-Driven Correlations in Many-Electron Ensembles: Theory and Application for Excited States. *Phys. Rev. Lett.* **2019**, *123*, 16401.

(40) Fromager, E. Individual Correlations in Ensemble Density Functional Theory: State-And Density-Driven Decompositions without Additional Kohn-Sham Systems. *Phys. Rev. Lett.* **2020**, *124*, 243001.

(41) Song, L.; Gao, J. On the Construction of Diabatic and Adiabatic Potential Energy Surfaces Based on Ab Initio Valence Bond Theory. *J. Phys. Chem. A* **2008**, *112*, 12925–12935.

(42) Mo, Y.; Gao, J. Energy decomposition analysis of intermolecular interactions using a block-localized wave function approach. *J. Chem. Phys.* **2000**, *112*, 5530–5538.

(43) Mo, Y.; Bao, P.; Gao, J. Energy decomposition analysis based on a block-localized wavefunction and multistate density functional theory. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6760–6775.

(44) Mo, Y.; Gao, J. Ab Initio QM/MM Simulations with a Molecular Orbital-Valence Bond (MOVB) Method: Application to an SN2 Reaction in Water. *J. Comput. Chem.* **2000**, *21*, 1458–1469.

(45) Cembran, A.; Provorse, M. R.; Wang, C.; Wu, W.; Gao, J. The third dimension of a more O'Ferrall-Jencks diagram for hydrogen atom transfer in the isoelectronic hydrogen exchange reactions of (PhX) 2H • with X = O, NH, and CH2. *J. Chem. Theory Comput.* **2012**, *8*, 4347–4358.

(46) Li, H.; Wang, Y.; Ye, M.; Li, S.; Li, D.; Ren, H.; Wang, M.; Du, L.; Li, H.; Veglia, G.; Gao, J.; Weng, Y. Dynamical and allosteric regulation of photoprotection in light harvesting complex II. *Science China Chem.* **2020**, *63*, 1121–1133.

(47) Bao, P.; Hettich, C. P.; Shi, Q.; Gao, J. Block-Localized Excitation for Excimer Complex and Diabatic Coupling. *J. Chem. Theory Comput.* **2021**, *17*, 240–254.

(48) Dederichs, P. H.; Blugel, S.; Zeller, R.; Akai, H. Ground states of constrained systems: Application to Cerium impurities. *Phys. Rev. Lett.* **1984**, *53*, 2512–2515.

(49) Liu, M.; Chen, X.; Grofe, A.; Gao, J. Diabatic States at Construction (DAC) through Generalized Singular Value Decomposition. *J. Phys. Chem. Lett.* **2018**, *9*, 6038–6046.

(50) Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained density functional theory. *Chem. Rev.* **2012**, *112*, 321–370.

(51) Becke, A. D.; Savin, A.; Stoll, H. Extension of the local-spin-density exchangecorrelation approximation to multiplet stat. *Theor. Chim. Acta* **1997**, *91*, 147–156.

(52) Savin, A. A combined density functional and configuration-interaction method. *Int. J. Quantum Chem.* **1988**, *S22*, 59–69.

(53) Moscardo, F.; Sanfabian, E. Density-functional formalism and the 2-body problem. *Phys. Rev. A* **1991**, *44*, 1549–1553.

(54) Grafenstein, J.; Cremer, D. The self-interaction error and the description of nondynamic electron correlation in density functional theory. *Theor. Chem. Acc.* **2009**, *123*, 171–182.

(55) Ukai, T.; Nakata, K.; Yamanaka, S.; Takada, T.; Yamaguchi, K. A CAS-DFT study of fundamental degenerate and nearly degenerate systems. *Mol. Phys.* **2007**, *105*, 2667–2679.

(56) Burkhard Miehlich, H. S.; Savin, A. A correlation-energy density functional for multideterminantal wavefunctions. *Mol. Phys.* **1997**, *91*, 527–536.

(57) Yang, L.; Grofe, A.; Reimers, J.; Gao, J. Multistate density functional theory applied with 3 unpaired electrons in 3 orbitals: The singdoublet and tripdoublet states of the ethylene cation. *Chem. Phys. Lett.* **2019**, *736*, 136803.

(58) Gunnarsson, O.; Lundqvist, B. I. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Phys. Rev. B* **1976**, *13*, 4274.

(59) Ziegler, T.; Rauk, A.; Baerends, E. J. Calculation of Multiplet Energies by Hartree-Fock-Slater Method. *Theor. Chim. Acta* **1977**, *43*, 261–271.

(60) Lu, Y.; Gao, J. Multi-State Density Functional Theory for Ground and Excited States. *ChemRxiv,* **2021**, 10.26434/chemrxiv–2021–h2bmc.

(61) Li, Z. H.; Valero, R.; Truhlar, D. G. Improved direct diabatization and coupled potential energy surfaces for the photodissociation of ammonia. *Theor. Chem. Acc.* **2007**, *118*, 9–24.

(62) Xie, C.; Ma, J.; Zhu, X.; Zhang, D. H.; Yarkony, D. R.; Xie, D.; ; Guo, H. Full-Dimensional Quantum State-to-State Nonadiabatic Dynamics for Photodissociation of Ammonia in its A-Band. *J. Phys. Chem. Lett.* **2014**, *5*, 10551060.

(63) Li, F.; Gillett, A. J.; Gu, Q.; Ding, J.; Chen, Z.; Hele, T. J. H.; Myers, W. K.; Friend, R. H.; Evans, E. W. Singlet and triplet to doublet energy transfer: improving organic light-emitting diodes with radicals. *Nat. Commun.* **2022**, *13*, 2744.

(64) Maylander, M.; Nolden, O.; Franz, M.; Chen, S.; Bancroft, L.; Qiu, Y.; Wasielewski, M. R.; Gilch, P.; Richert, S. Accessing the triplet state of perylenediimide by radical-enhanced intersystem crossing. *Chem. Sci.* **2022**, *13*, 6732–6743.

(65) Evans, F. D. Magnetic Perturbation of the Lowest Triplet States of Aromatic Molecules by Dissolved Oxygen. *Nature* **1956**, *178*, 534–535.

(66) Evans, F. D. Nagnetic Perturbation of Singlet-Triplet Transitions. Part II. *J. Chem. Soc.* **1957**, 3885–3888.

(67) Ake, R. L.; Gouterman, M. Porphyrins XIV. Theory for the Luminescent State in VO, Co, Cu Complexes. *Theoret. Chim. Acta (Berl.)* **1969**, *15*, 20–42.

(68) Bortolus, M.; Ribaudo, G.; Toffoletti, A.; Carbonera, D.; Zagotto, G. Photo-induced spin switching in a modified anthraquinone modulated by DNA binding. *Photochem. Photobio. Sci.* **2019**, *18*, 2199–2207.

(69) Lin, Z.-P.; Aue, W. A. Triplet-state energies and substituent effects of excited aroyl compounds in the gas phase. *Spectrochimica Acta Part A* **1999**, *56*, 111–117.

(70) Sundstrom, E. J.; Head-Gordon, M. Non-orthogonal configuration interaction for the calculation of multielectron excited states. *J. Chem. Phys.* **2014**, *140*, 114103.

(71) Mayhall, N. J.; Horn, P. R.; Sundstrom, E. J.; Head-Gordon, M. Spin–flip non-orthogonal configuration interaction: a variational and almost black-box method for describing strongly correlated molecules. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22694–22705.

(72) Olsen, J. Novel methods for configuration interaction and orbital optimization for wave functions containing non-orthogonal orbitals with applications to the chromium dimer and trimer. *J. Chem. Phys.* **2015**, *143*, 114102.

(73) Oosterbaan, K. J.; White, A. F.; Head-Gordon, M. Non-orthogonal configuration interaction with single substitutions for the calculation of core-excited states. *J. Chem. Phys.* **2018**, *149*, 044116.

(74) Kathir, R. K.; de Graaf, C.; Broer, R.; Havenith, R. W. A. Reduced Common Molecular Orbital Basis for Nonorthogonal Configuration Interaction. *J. Chem. Theory Comput.* **2020**, *16*, 2941–2951.

(75) Straatsma, T. P.; Broer, R.; Faraji, S.; Havenith, R. W. A.; Suarez, L. E.; Kathir, R. K.; Wibowo, M.; Graaf, C. D. GronOR: Massively parallel and GPU-accelerated non-orthogonal configuration interaction for large molecular systems. *J. Chem. Phys.* **2020**, *152*, 064111.

(76) Jiménez-Hoyos, C. A. Approaching the full configuration interaction ground state from an arbitrary wavefunction with gradient descent and quasi-Newton algorithms. *J. Chem. Phys.* **2020**, *153*, 234113.

(77) Mahler, A. D.; Thompson, L. M. Orbital optimization in nonorthogonal multiconfigurational self-consistent field applied to the study of conical intersections and avoided crossings. *J. Chem. Phys.* **2021**, *154*, 244101.