A New Release of MOPAC Incorporating the INDO/S Semiempirical Model

with CI Excited States

Rebecca L. M. Gieseking

Department of Chemistry, Brandeis University
415 South Street, Waltham, Massachusetts 02453, United States

* Corresponding author: gieseking@brandeis.edu; (781)-736-2511

1
Abstract

We have incorporated the semiempirical INDO/S Hamiltonian into a new release of MOPAC2016, which has long been at the forefront of semiempirical quantum chemical methods (SEQMs). Our new code enables the calculation of excited states using the INDO/S Hamiltonian combined with a configuration interaction (CI) approach using single excitations (CIS), single and double excitations (CISD), or multiple reference determinants (MRCI) where reference determinants are generated using a complete active space (CAS) approach. The capacity to perform excited-state calculations beyond the CIS level makes INDO/CI one of the few low-cost computational methods capable of accurately modeling states with substantial double-excitation character. Solvent corrections to the ground-state and excited-state energies can be computed using the COSMO implicit solvent model, incorporating state-specific corrections to the excited states based on the solvent refractive index. We demonstrate that this code produces physically reasonable electronic structures, absorption spectra, and solvatochromic shifts at low computational costs for systems up to hundreds of atoms, and for both organic molecules and metal clusters.
1. **Introduction**

Accurate modeling of large, complex molecular systems is critical across many fields of chemistry, such as biochemical systems,\(^1\)\(^-\)\(^3\) electrochemical interfaces,\(^4\)\(^,\)\(^5\) and nanomaterials.\(^6\)\(^-\)\(^9\) Although some properties of these systems can be modeled with reasonable accuracy using classical force fields,\(^10\)\(^,\)\(^11\) quantum mechanical methods are still essential for many chemically important properties, including excited states and electron transfer. Semiempirical quantum mechanical methods (SEQMs) have experienced a resurgence in recent years because of their capacity to address these challenges at accessible computational costs.\(^1\)\(^,\)\(^12\)\(^-\)\(^20\) SEQMs use low-cost Hamiltonians simplified from Hartree-Fock. The introduction of parameters fit to experimental and high-level computational results has enabled high enough accuracy to provide understanding of chemical properties including thermochemistry,\(^13\)\(^,\)\(^21\)\(^,\)\(^22\) noncovalent interactions,\(^13\)\(^,\)\(^21\)\(^,\)\(^23\) excited states,\(^24\)\(^,\)\(^25\) \(pK_a\),\(^14\)\(^,\)\(^19\) magnetism,\(^26\) solvatochromism,\(^27\) and electrochemical formal potentials.\(^28\) We note that the Density Functional Tight Binding (DFTB) model is a semiempirical-like model based on a Hamiltonian simplified from Density Functional Theory (DFT) and has many capabilities similar to those of Hartree-Fock-based SEQMs;\(^29\)\(^-\)\(^31\) here, we focus exclusively on SEQMs derived from Hartree-Fock.

SEQMs have long contributed to understanding chemical properties. Hückel theory, which describes the \(\pi\) electrons of planar conjugated systems, provided the first description of aromaticity; the neglect of electron-electron repulsion in this theory means that the Hückel matrix must only be constructed and diagonalized once.\(^32\)\(^,\)\(^33\) Later developments in extended Hückel theory extended this model to all valence s and p orbitals, but still neglected electron-electron repulsion.\(^34\) Pople developed the Complete Neglect of Differential Overlap (CNDO) method,\(^35\)\(^,\)\(^36\) which includes electron-electron repulsion but neglects the overlap between basis
functions by setting all two-electron integrals of the form $(\mu \nu | \lambda \sigma)$ equal to zero unless $\mu = \nu$ and $\lambda = \sigma$. This approximation, known as the Zero Differential Overlap (ZDO) approximation, is used to varying extents in different SEQMs. In Pople’s Intermediate Neglect of Differential Overlap (INDO) method, the two-electron integrals of the form $(\mu \nu | \mu \nu)$ are neglected only if basis functions $\mu$ and $\nu$ are centered on different atoms (see the Theory section for more detail on the INDO approximation). Pople also extended this formalism to the Neglect of Diatomic Differential Overlap (NDDO) class of methods, which includes many modern semiempirical methods such as MNDO, AM1, PM3, PM6, and PM7. Within the NDDO approximation, integrals of the form $(\mu \nu | \lambda \sigma)$ are set to zero unless basis functions $\mu$ and $\nu$ are both centered on atom $A$ and basis functions $\lambda$ and $\sigma$ are both centered on atom $B$, where $A$ may be equal to $B$. In all of these models, the remaining integrals are parametrized based on ab initio results or experimental data.

The INDO formalism has been of particular interest because of Zerner’s parametrization called INDO/S (or equivalently, ZINDO/S), which combines the INDO framework with parameters designed to reproduce vertical excited-state energies using a configuration interaction (CI) approach with single excitations (CIS) within a limited active space. The INDO/S Hamiltonian has also contributed significantly to understanding excited states with large double-excitation character using either single and double excitations (CISD) or a multi-reference approach (MRCI); these calculations have been particularly important in understanding molecular nonlinear optical properties. The INDO framework has also been extended to transition metals with d orbitals. In recent years, we have parametrized INDO for the optical properties of Ag clusters; these parameters have given insight into features such as the quantum mechanical nature of quadrupolar plasmons, chemical effects in surface-enhance
Raman spectroscopy (SERS), and nonlinear optical properties of plasmons. Several implicit solvation methods, including PCM and COSMO, have been implemented with INDO and yield solvatochromic shifts in reasonable agreement with experimental results.

Despite the usefulness of the INDO method, many features have not been implemented in publicly available codes to date. ORCA 4.0.1 incorporates several INDO parameter sets (ZINDO/S, (Z)INDO/1, and (Z)INDO/2) and can perform CIS calculations, and includes the ability for users to modify the built-in parameters. Gaussian 16 includes INDO/1 parameters and likewise can perform CIS calculations. The Cramer group has incorporated the INDO/S Hamiltonian and real-time (RT)-INDO into NWChem, but CI-type calculations are not available. None of these codes contains features such as implicit solvation or INDO-based CISD and MRCI that make INDO applicable for a wider range of properties and systems.

MOPAC has long been a leading software package for semiempirical models, focusing to date exclusively on NDDO methods. Developed by Stewart, it incorporates the latest NDDO methods such as PM6, PM7, and variants of these methods, and excited states can be computed using a variety of CI-based approaches within a limited active space. Implicit solvent effects using the COSMO model are available for ground states and excited states. MOPAC also includes a number of features tailored to working with proteins, such as localized molecular orbital approaches that enable computations on thousands of atoms, and can compute the properties of solid-state systems within periodic boundary conditions.

Here, we report a new release of MOPAC2016 that incorporates the INDO/S Hamiltonian. A variety of CI-based approaches for excited states are available, including CIS, CISD, CAS, and MRCI, and the CI matrix can be controlled using a combination of a defined active space and a size-based cutoff. Implicit solvent is available using the COSMO method,
including state-specific corrections to the excited states. A default set of INDO/S parameters are included in the code, and the user can choose to modify any parameter using an external file. We demonstrate here that this code provides physically reasonable ground-state electronic structures and excited-state properties for both organic molecules and metal clusters.

2. Theory

The newly released MOPAC2016 code incorporates many of the capabilities of the CNDO/INDO code of Reimers et al.50 These modifications significantly extend the capabilities of the code beyond what was previously available in either MOPAC or CNDO/INDO. Since the INDO Hamiltonian was not previously implemented within MOPAC, all of the capabilities we discuss are new to MOPAC. The key advances over the CNDO/INDO code are:

1. We have harnessed the more advanced SCF algorithms in MOPAC to enable SCF convergence for systems much larger than were feasible within CNDO/INDO.
2. We have also implemented the COSMO implicit solvent model for INDO ground states and INDO/CI excited states. We previously implemented COSMO for INDO/CIS in unreleased code;27 the COSMO correction for INDO ground states is adapted from subroutines within MOPAC, and the excited-state COSMO corrections involve new subroutines. The extension of this COSMO correction to INDO/CISD and INDO/MRCI excited states is new to this work.

2.1. INDO/S Hamiltonian

Since the INDO/S Hamiltonian has been described in detail many times,42,47,49,60 we provide here a brief summary of the method. We focus here on closed-shell systems; the
equations for open-shell systems are analogous. The INDO/S formalism is based on INDO/1, with parameters set to reproduce spectroscopic properties.\textsuperscript{42,47} As in other SEQMs, INDO/S treats only the valence electrons explicitly within a minimal Slater basis set. The s and p basis functions consist of a single Slater function; the s and p Slater functions use the same exponent ($\zeta_s = \zeta_p$). The d basis functions consist of a linear combination of two Slater functions.

The INDO/S molecular orbitals (MOs) $\psi_i$ are linear combinations of atomic orbitals $\phi_\mu$ with LCAO-MO coefficients $c_{i\mu}$:

$$
\psi_i = \sum_{\mu=1}^{N} c_{i\mu} \phi_\mu
$$

where $N$ is the number of basis functions. Using the Zero Differential Overlap (ZDO) approximation, the overlap matrix $S$ is set to the identity matrix. Within this approximation, the secular equation is:

$$
F \mathbf{c} = \mathbf{c} \mathbf{\epsilon}
$$

where $F$ is the Fock matrix, $\mathbf{c}$ is the matrix of LCAO-MO coefficients $c_{i\mu}$, and $\mathbf{\epsilon}$ is a diagonal matrix of MO energies.

The elements of the Fock matrix are

$$
F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[ (\mu\nu|\lambda\sigma) - \frac{1}{2} (\mu\lambda|\nu\sigma) \right]
$$

where $h$ is the one-electron matrix. All two-electron integrals $(\mu\nu|\lambda\sigma)$ involving basis functions on three or four atomic centers are set to zero, as are all two-center two-electron exchange-type integrals. The two-center two-electron Coulomb-type integrals are only non-zero if they involve exactly two basis functions ($\mu_A = \nu_A$ and $\lambda_B = \sigma_B$, where the subscript indicates which atom the
basis function is centered on). These integrals are approximated using the Mataga-Nishimoto formula:

\[ (\mu_A\mu_A | \lambda_B\lambda_B) = \gamma_{AB} = \frac{1.2}{R_{AB} + F^0(\mu_A\mu_A) + F^0(\lambda_B\lambda_B)} \]

where \( R_{AB} \) is the distance between nuclei \( A \) and \( B \) and \( F^0(\mu_A\mu_A) \) are the Slater-Condon one-center two-electron integrals. One-center two-electron Coulomb and exchange integrals are approximated using the higher-order Slater-Condon factors \( F^k \) and \( G^k \).

The diagonal one-electron matrix elements \( h_{\mu_A\mu_A} \) are approximated as:

\[ h_{\mu_A\mu_A} = U_{\mu_A\mu_A} - \sum_{B \neq A} Z_B \gamma_{AB} \]

where \( U_{\mu_A\mu_A} \) is the one-center core integral, typically derived from the atomic ionization potential and \( Z_B \) is the effective charge of nucleus \( B \). One-electron matrix elements involving two different basis functions on the same atomic center are set to zero \( (h_{\mu_A\nu_A} = 0) \), and one-electron matrix elements involving basis functions on different atomic centers are approximated as:

\[ h_{\mu_A\nu_B} = \frac{1}{2} (\beta_{\mu_A} + \beta_{\nu_B}) \bar{S}_{\mu_A\nu_B} \]

where \( \beta_{\mu_A} \) is the resonance integral and \( \bar{S}_{\mu_A\nu_B} \) is a modified overlap integral of the Slater basis functions that uses geometric factors to account for the relative spatial orientations of the basis functions.

Given these approximations, the INDO/S parameters required for atoms with s, p, and d basis functions are (1) the Slater orbital exponents \( \zeta_s = \zeta_p, \zeta_{d_1}, \) and \( \zeta_{d_2} \), and weighting factors \( c_{d_1} \) and \( c_{d_2} \) for each exponent; (2) ionization potentials \( IP_s, IP_p, \) and \( IP_d \) for s, p, and d atomic
orbitals, respectively; (3) resonance integrals $\beta_s = \beta_p$ and $\beta_d$; and (4) the Slater-Condon parameters.

2.2. **INDO/CI Approaches**

One of the major advantages of INDO/CI calculations is the ability to perform not only CIS calculations but also readily incorporate double and higher excitations using a singles and doubles (CISD), complete active space (CAS), or multi-reference (MRCI) approach. The INDO/CI code in MOPAC can use several approaches to generate the list of excitations that are include in the CI matrix. We define active spaces using the notation $(m, n)$, where $m$ is the total number of molecular orbitals in the active space and $n$ is the number of doubly occupied molecular orbitals in the active space; this notation is equivalent to the keyword format used in MOPAC. If an active space is not specified, the active space for single excitations includes all molecular orbitals for the system by default.

A CAS calculation uses the same notation $(m, n)$ for the active space. The default CAS active space is $(2, 1)$, which for a closed-shell system generates three Slater determinants: the SCF ground state, a single HOMO $\rightarrow$ LUMO excitation, and the double HOMO, HOMO $\rightarrow$ LUMO, LUMO excitation. If MRCI is specified, all Slater determinants generated in the CAS step are used as reference determinants for single excitations (by default) or single and double excitations (if MRCI is paired with CISD).

In a CISD calculation, the active space for double excitations can be set using the same $(m, n)$ notation as for single excitations; by default, CISD uses a $(10, 5)$ active space. To generate double excitations, all single excitations within the CIS active space are used as starting points for generating a second excitation. Thus, only the second excitation must fall within the active
space for double excitations. This is equivalent to using the CIS single excitations as reference determinants for excitations within the smaller double-excitation active space.

The INDO/CI code also requires setting a maximum size of the CI matrix; if the number of electron configurations generated within the requested active space is greater than the maximum CI matrix size, configurations are selected on the basis of lowest energy. All determinants are spin-adapted before they are added to the CI matrix. The CI matrix is then diagonalized to obtain the excited states.

2.3. Implementation of COSMO with INDO/CI

Our implementation of the COSMO solvent model for INDO/S Hamiltonian follows the methodology laid out by Klamt; we extend here our previously reported implementation to not only CIS excited states but also CISD and MRCI states. In the Conductor-like Screening Model (COSMO), the solvent surrounding a molecule is modeled as a continuum with a dielectric constant $\epsilon$. The cavity is constructed by using a union of atom-centered spheres and dividing the surface into discrete triangles that define the positions of point charges. The solvation energy is computed using two matrices that capture interactions between charges in the system. Matrix $B$ contains all interactions between the solvent surface charges $\mu$ and molecular charges $i$:

$$B_{i\mu} = \frac{1}{\|r_i - r_\mu\|}$$

Similarly, matrix $A$ contains all interactions between pairs of surface charges $\mu$ and $\nu$:

$$A_{\mu\nu} = \frac{1}{\|r_\mu - r_\nu\|}$$
where the vectors \( \mathbf{r}_i, \mathbf{r}_\mu, \) and \( \mathbf{r}_\nu \) indicate the positions of the indicated charges. For a perfect conductor, the surface charges \( q \) that minimize the total energy are:

\[
q = -A^{-1}BQ
\]

The screening energy is scaled by a factor of \( f(\epsilon) = \frac{\epsilon^{-1}}{\epsilon + z} \) relative to a perfect conductor. The total screening energy \( \Delta E \) for the solute charge distribution \( Q \) is thus:

\[
\Delta E = f(\epsilon) \left[ QBq + \frac{1}{2} QAq \right] = -\frac{1}{2} f(\epsilon) QBA^{-1}BQ = \frac{1}{2} f(\epsilon) QBq
\]

Within the MOPAC code, the total ground-state surface charges are partitioned into surface charges induced by the nuclei and surface charges induced by the electrons:

\[
q_{\text{tot}} = q_{\text{nuc}} + q_{\text{elec}}
\]

Since the one-electron matrix is computed after the nuclear positions are set but prior to the self-consistent field (SCF) calculation, its elements are initially corrected for interactions with \( q_{\text{nuc}} \). In each SCF cycle, the elements of the Fock matrix are computed using the \( q_{\text{nuc}} \)-corrected one-electron matrix elements and then corrected for interactions with \( q_{\text{elec}} \) based on the current electron density; the Fock matrix incorporating corrections for both \( q_{\text{nuc}} \) and \( q_{\text{elec}} \) is diagonalized to provide a new set of MO coefficients and energies. Since the Fock matrix includes corrections for \( q_{\text{tot}} \) from the combination of these two steps, the MO energies obtained through this process likewise include corrections for \( q_{\text{tot}} \). Corrections for \( q_{\text{elec}} \) can be added to the one-electron matrix after SCF convergence is achieved if they are needed.
The state-specific COSMO corrections to the CI energies follow the procedure of Klamt. On the time scale of absorption, only the electronic structure of the solvent relaxes quickly enough to respond to the change in molecular electron density; the geometric and orientational relaxation that dominate the dielectric constants of polar solvents occur too slowly to contribute. Thus, to obtain accurate absorption spectra, the dynamic relaxation of the solvent is scaled by 

\[ f(n^2) = \frac{n^2-1}{n^2+1} \]

where \( n \) is the solvent refractive index. Within the CI procedure, the COSMO-corrected energy of the transition \( T \) is:

\[
E_{T,solv} = E_{T,0} + f(\epsilon)\Delta_T D P_0 - \frac{1}{2} f(n^2)\Delta_T D \Delta_T
\]

where \( \Delta_T \) is the change in charge distribution when transition \( T \) occurs and \( E_{T,0} \) is the gas-phase transition energy from the SCF ground state. The term proportional to \( f(\epsilon) \) is the energy correction related to the ground-state surface charges. When the CI matrix elements are computed using the one-electron matrix elements or MO energies that are corrected for \( q_{tot} \), this term is incorporated automatically into the transition energies.

The term proportional to \( f(n^2) \) is the energy correction related to the polarization of the solvent upon excitation on optical time scales. The corrections are computed via a two-step process. First, while the CI matrix is being constructed, the diagonal elements of the CI matrix are corrected based on the change in electron density upon excitation of the system from the SCF ground state to the pure single, double, etc. excitation. As in Klamt’s original procedure, the off-diagonal elements of the CI matrix are not modified. The solvent-corrected CI matrix is then diagonalized to obtain the CI states.

In a second step, the CI states are corrected for the change in electron density in the excited state. Since the CI states are linear combinations of excitations, the change in electron
density upon excitation to the excited state is likewise a linear combination of the changes in electron density of the component excitations. For most excited states, the change in electron density is thus smaller than the changes in the electron density of the individual component excitations, so the initial solvent-corrected energies obtained upon diagonalization of the CI matrix are too low. To correct the CI state energies, the diagonal CI matrix elements excluding the $f(n^2)$ corrections are computed, and the energy of each CI state is computed based on this uncorrected CI matrix. The change in electron density associated with the excited state is computed and used to compute the state-specific solvent correction proportional to $f(n^2)$.

Although Klamt’s original results showed that a single-step incorporation of the state-specific corrections was adequate and the two-step procedure was unnecessary, his results focused on small molecules with small to moderate solvatochromic shifts. In our implementation, we are interested in ensuring that the results are reliable for (1) excited states with substantial charge-transfer (CT) character that have large solvatochromic shifts and (2) excited states in large molecules that are linear combinations of many excitations. In this scenario, the initial correction step ensures that the relative energies of the CT excitations with large solvatochromic shifts are better aligned with the energies of local excitations with small solvatochromic shifts to give appropriate mixing when the CI matrix is diagonalized, and the second step ensures correct energies of the CI states. We note that since the state-specific corrections are perturbative corrections relative to the SCF ground state, the excited-state energies computed using this approach are only expected to be accurate for systems where the ground state is adequately described by a single reference determinant.
3. Computational Methods

The geometric structures used in this work came from several sources. The polyenes were optimized at the B3LYP/cc-pVDZ level\textsuperscript{64,65} using Gaussian 16;\textsuperscript{57} minima were confirmed by the absence of imaginary frequencies. Previously published geometric structures were used for all other molecules: P3B2 and f-coronene at the B3LYP/6-31G(d) level,\textsuperscript{18} donor-acceptor substituted polyenes at the ωB97XD/6-31G* level using the COSMO solvent model in toluene and ethanol,\textsuperscript{27} and Ag\textsubscript{20} at the BP86/DZ level.\textsuperscript{17}

Excited states were computed using the INDO/S Hamiltonian as implemented in the MOPAC2016 code. For all excited-state calculations (CIS, CISD, and MRCI), unless otherwise specified the active space for single excitations included all molecular orbitals within the minimal INDO/S basis, the CI matrix included a maximum of 2000 excitations selected on the basis of energy, and the lowest 500 excited states were printed. For all CISD and MRCI calculations, the (2,1) active space was used, indicating that the active space included two total molecular orbitals, of which one was doubly occupied.

4. Results

4.1. INDO Electronic Structure

We first demonstrate that the INDO/S Hamiltonian yields physically reasonable ground-state electronic structures by comparing the molecular orbital (MO) energies and wavefunctions to the corresponding features at the B3LYP/cc-pVDZ level. For \textit{trans}-1,3-butadiene, we focus on the four frontier MOs, which at both the B3LYP/cc-pVDZ and INDO/S levels are π MOs (Figure 1a-c). The MO wavefunctions are consistent at the two levels of theory: the lowest-energy π MO is fully bonding, and each successively higher-energy MO has one additional node
along the long molecular axis. Since the meaning of MO energies within a DFT framework is somewhat different from the meaning within the Hartree-Fock-derived SEQM framework,\textsuperscript{66} it is unsurprising that the HOMO-LUMO gap is 3.4 eV larger at the INDO/S level than at the B3LYP/cc-pVDZ level. As we will see later, the first excited-state energies at the INDO/MRCI(2,1) and TD-B3LYP/cc-pVDZ levels are within 0.1 eV of each other, which shows that the difference in HOMO-LUMO gaps is due to these differences in the definitions of the MO energies and not to any deficiencies in the INDO/S Hamiltonian. Since the first excited-state has essentially HOMO $\rightarrow$ LUMO character, we expect it to be reasonably well represented at both levels of theory.
Figure 1. Geometric structures of (a) trans-1,3-butadiene, (d) tetrahedral Ag\textsubscript{20}, and (g) P3B2. Molecular orbital diagrams of (b,c) trans-1,3-butadiene, (e,f) tetrahedral Ag\textsubscript{20}, and (h,i) P3B2 at the (b,e,h) B3LYP/cc-pVDZ and (c,f,i) INDO/S levels of theory.

Similarly, the INDO/S MO diagram for the tetrahedral Ag\textsubscript{20} cluster is in good agreement with the B3LYP/cc-pVDZ-PP results (Figure 1d-f). Ag\textsubscript{20} has been used extensively as a
computational model system for plasmonic noble metal clusters.\textsuperscript{17,67–69} Since a neutral Ag atom has a [Kr]4d\textsuperscript{10}5s\textsuperscript{1} electron configuration, the Ag\textsubscript{20} cluster has a fully filled 4d valence band and 20 electrons in the conduction (sp) band. As has been previously observed, the conduction band MOs are superatomic orbitals, MOs distributed across the full metal cluster with shapes that resemble those of atomic orbitals.\textsuperscript{70} Since the tetrahedral shape of Ag\textsubscript{20} is fairly far from spherical, the atomic-orbital-like shapes of the superatomic orbitals are noticeably distorted, and the degeneracy is reduced. Both levels of theory predict a doubly degenerate HOMO slightly higher in energy than a triply degenerate set of MOs; all five of these occupied MOs have d-like superatomic shapes. The triply degenerate LUMO is part of a set of f-like superatomic MOs. We have previously shown that INDO/CIS predicts an absorption spectrum for Ag\textsubscript{20} consistent with TDDFT results.\textsuperscript{17,51}

This consistency between INDO/S and B3LYP/cc-pVDZ applies not only to relatively small molecules but also to much larger π-conjugated molecules. For P3B2, previously used as a prototypical large molecule for RT-TDDFT and RT-INDO studies,\textsuperscript{18,71} the nodal patterns of the frontier molecular orbitals are identical at the INDO/S and B3LYP/cc-pVDZ levels (Figure 1g-i). At both levels of theory, the HOMO is delocalized across the full molecule and has slightly more electron density on the central porphyrin ring than on the terminal porphyrin rings, whereas the LUMO is localized primarily on the central porphyrin ring and has very little electron density on the terminal porphyrin rings. These results for several prototypical systems show that the INDO/S model can produce accurate electronic structures for a wide variety of molecular systems.
4.2. \textit{INDO/CIS, INDO/CISD, and INDO/MRCI for Small Molecules}

To show the applicability of the various INDO/CI approaches available in MOPAC2016, we examine the excited states of linear polyenes (Figure 2a; \( n = 1-5 \)); the smallest of these molecules is \textit{trans}-1,3-butadiene, discussed in the previous section. In these molecules, the first excited state with strong one-photon absorption has \( B_u \) symmetry (\( 1B_u \) state) and dominantly HOMO \( \rightarrow \) LUMO character, and the first excited state with strong two-photon absorption has \( A_g \) symmetry (\( 2A_g \) state) and is a linear combination of HOMO – 1 \( \rightarrow \) LUMO, HOMO \( \rightarrow \) LUMO + 1, and the double excitation HOMO, HOMO \( \rightarrow \) LUMO, LUMO. For many \( \pi \)-conjugated molecules, the absorbing excited states all have dominant singly excited character, and INDO/CIS is adequate to represent the absorption spectra. The large doubly excited character of the polyene \( 2A_g \) state makes these molecules a good test of the INDO/CISD and INDO/MRCI methods in particular.

For the smallest polyenes in the series, the \( 1B_u \) state is the first excited state. At a length of 8 carbon atoms (\( n = 3 \)), there is a crossover between these two states such that the \( 2A_g \) state becomes the first excited state experimentally.\(^{72}\) Calculations at the CASPT2/TZVP level reproduce the experimental crossover in excited-state energies (Figure 2b).\(^{73}\)
Figure 2. (a) Chemical structure of the all-trans polyenes containing 4-12 carbon atoms \((n = 1-5)\). Energies of the \(1\text{Bu}\) and \(2\text{Ag}\) excited states of all-trans polyenes (b) at the CASPT2/TZVP level (data from ref. 73), (c) at the TD-B3LYP/cc-pVDZ level, and using the INDO Hamiltonian with (d) CIS, (e) CISD with a (2,1) active space for double excitations, and (f) MRCI with reference determinants generated using CAS with a (2,1) active space.

Because of the large doubly excited character of the \(2\text{Ag}\) state, single-excitation methods predict \(2\text{Ag}\) energies that are too large and do not reproduce the expected crossover with the \(1\text{Bu}\) state. This failure of single-excitation methods is seen at both the TD-B3LYP/cc-pVDZ and INDO/CIS levels. At the INDO/CIS level, the \(2\text{Ag}\) state is more than 1 eV higher in energy than
the 1Bu state across the range of molecular lengths studied here. At the TD-B3LYP/cc-pVDZ level, the 1Bu-2Ag gap decreases from 1.2 eV to 0.5 eV as molecular length increases.

In contrast, calculations at the INDO/CISD and INDO/MRCI level with (2, 1) active spaces yield 2Ag and 1Bu excited-state energies within 0.35 eV of each other at all molecular lengths. INDO/MRCI shows a crossover in excited-state ordering at a length of 10 carbon atoms, only one double-bond longer than seen experimentally. Considering the low computational cost of INDO/MRCI, the agreement with experiment and with CASPT2/TZVP is quite good. The 1Bu-2Ag energy difference at the INDO/MRCI level is within 0.12 eV of the CASPT2/TZVP results for all molecules where comparisons can be made, and the errors in excited-state energies decrease from 0.4 eV for \( n = 1 \) to < 0.1 eV for \( n = 3 \). For all molecules, the 2Ag state has at least 22% double HOMO, HOMO \( \rightarrow \) LUMO, LUMO excitation character, accounting for the substantial difference in excited-state energy.

Although INDO/CISD correctly predicts that the 1Bu and 2Ag states are quite close in energy, it incorrectly predicts that the 2Ag state is lower in energy than the 1Bu state by 0.03-0.12 eV at all molecular lengths and does not reproduce the crossover in energies. The only difference between the INDO/MRCI and INDO/CISD calculations with our choice of active spaces is that INDO/MRCI includes some triple excitations, generated as single excitations relative to a doubly excited HOMO, HOMO \( \rightarrow \) LUMO, LUMO reference determinant. These triple excitations have small contributions to the 1Bu state, stabilizing it and improving the agreement with experimental and high-level computational trends. Although computations involving multiply excited configurations require more care in selecting an appropriate CI approach to ensure accuracy for the properties of interest, these results highlight the crucial role of double excitations to capture dark states and the ability of INDO/CI to model these states at low computational cost.
4.3. **INDO/CIS for Large Molecules**

To show the applicability of this code to the absorption spectra of large molecules, we have computed the INDO/CIS absorption spectra of two prototypical large conjugated molecules that have previously been used in benchmarking RT-TDDFT\textsuperscript{71} and RT-INDO:18 P3B2 (130 atoms, 406 basis functions) and f-coronene (162 atoms, 522 basis functions), shown in Figure 3a-b. These molecules allow us to demonstrate the convergence of the absorption spectra with respect to size of the active space and of the CI matrix. Active spaces use the notation \((m, n)\), where \(m\) is the total number of molecular orbitals in the active space and \(n\) is the number of doubly occupied molecular orbitals in the active space. Thus, our smallest active space \((20, 10)\) includes 10 occupied MOs and 10 virtual MOs; the total number of single excitations within this active space is \(10 \times 10 = 100\). For all calculations, the lowest 2000 excited states are included in the absorption spectrum, or all excited states if there are fewer than 2000 single excitations in the CI matrix.
Figure 3. Chemical structures of (a) P3B2 and (b) f-coronene. Absorption spectra of (c,e) P3B2 and (d,f) f-coronene computed at the INDO/CIS level. The size of the CI matrix is set (c,d) via choice of active space and (e,f) by generating all single excitations and selecting the $n$ lowest energy single excitations to include in the CI matrix.

For both P3B2 and f-coronene, the INDO/CIS calculations with the largest active space (all molecular orbitals) and CI matrix (7000 configurations) are in fairly good agreement with previous LR-TD-DFT and RT-TD-DFT results. For P3B2, the first strong absorption peak is at 2.69 eV at the INDO/CIS level, slightly lower than the LR-B3LYP/6-31G(d) value of 3.1 eV.\textsuperscript{71} This difference is comparable to the mean absolute error of 0.5 eV seen in previous benchmarking of INDO/CIS relative to CASPT2 reference excited-state energies for small
organic molecules. For f-coronene, the INDO/CIS absorption peaks are similarly a few tenths of an eV lower in energy that the LR-B3LYP/6-31G(d) values.

We compare the INDO/CIS absorption spectra including all molecular orbitals in the active space to four smaller active spaces: (20, 10), (40, 20), (80, 40), and (160, 80), shown in Figure 3c-d. For all of these active spaces, all excitations within the active space were included in the CI matrix. The first few excited states converge relatively quickly with the size of the active space. For P3B2, the first strong absorption peak (correspond to the ninth excited state) decreases in energy by 0.14 eV between the smallest and largest active spaces, and its oscillator strength decreases by 6%. In f-coronene, the first absorption peak (corresponding to the third or fourth excited state) is somewhat more affected by the size of the active space, particularly in terms of oscillator strength: the energy decreases by 0.21 eV and the oscillator strength by 37% between the smallest and largest active spaces. However, an increase in active space to (40, 20) shifts this peak to within 0.1 eV and 15% oscillator strength of the full active space limit. The higher-energy excited states require larger active spaces to obtain absorption spectra that are in qualitative agreement with a large active space. The absorption spectrum with a (40, 20) active space qualitatively agrees with the full active space up to around 5 eV for P3B2 and 4 eV for f-coronene, and the (160, 80) active space is required to obtain qualitative agreement up to 7-8 eV. This is unsurprising given the large density of states for these molecules: when all MOs are included in the active space, P3B2 and f-coronene have 126 and 237 excited states below 5 eV and 771 and 1387 excited states below 8 eV, respectively. The (40, 20) active space only includes 400 single excitations, so it is expected that far fewer than 400 excited states will have properties that are already converged with respect to active space. This suggests that to reliably
obtain excited-state properties for $n$ excited states, an active space including at least $4n$
excitations must be used.

We also examine the effects of the size of the CI matrix (Figure 3e-f). In these
calculations, all molecular orbitals were included in the active space, for a total of 41,109 (P3B2)
or 68,085 (f-coronene) single excitations possible within the INDO/S basis. Of these excitations,
the $m$ lowest-energy single excitations are included in the CI matrix, where $m$ ranges from 100 to
7000. Overall, the trends are quite similar to the changes in active space: capping the CI matrix
to the $m$ lowest-energy single excitations yields absorption spectra of relatively similar quality to
using all excitations within an active space that includes $m$ single excitations. There are minor
differences in the convergence of individual peaks, but the overall trends are quite similar. These
calculations show the capability of computing absorption spectra of moderately large molecules
up to thousands of excited states at low computational costs.

4.4. Solvatochromism using the COSMO implicit solvation model

Another important feature for excited-state calculations is the ability to model
solvatochromic effects. We have implemented the COSMO model of implicit solvation with
state-specific energy corrections in combination with the full suite of INDO/CI approaches
implemented in MOPAC, 27 as detailed in the Theory section. To show that the
INDO/CIS/COSMO method yields reasonable accuracy for solvatochromic shifts, we focus on a
set of 24 donor-acceptor-substituted polyenes for which experimental absorption spectra are
available in both toluene and ethanol (Figure 4a). 74–76 The solvatochromic shifts in this series of
molecules depends strongly on the details of the molecular structure. In donor-acceptor-
substituted polyenes, the absorption energy reaches a minimum near the cyanine limit when
approximately half an electron has transferred from the donor to the acceptor.\textsuperscript{77,78} As the solvent dielectric constant increases, the extent of ground-state CT from the donor to the acceptor increases. For molecules with relatively weak donors and acceptors, when the solvent dielectric constant increases, the increase in ground-state CT shifts the molecule closer to the cyanine limit and decreases the first excited-state energy, defined as a positive solvatochromic shift. In contrast, for molecules with strong enough donors and acceptors that the ground-state CT is greater than 0.5, an increase in ground-state CT shifts the molecule away from the cyanine limit and increases the first excited state energy, defined as a negative solvatochromic shift.

Across this series of molecules, INDO/CIS/COSMO predicts solvatochromic shifts that are consistently slightly more positive than the experimental values, but the deviation is relatively constant across the full series; the mean absolute error (MAE) is 0.07 eV.\textsuperscript{27} Comparison of the computed and experimental shifts yields $R^2 = 0.74$ (Figure 4b). This accuracy is comparable to the solvatochromic shifts computed using DFT at the $\omega$B97XD/6-31G* level, which yields an MAE of 0.05 eV and a slightly smaller $R^2$ of 0.63.\textsuperscript{27} At both levels of theory the deviations are largest for the molecules with the largest positive and negative experimental solvatochromic shifts. For many of these molecules, the experimental absorption peaks show changes in vibronic structure with solvent, so the shift in the first excited-state energy may be smaller than the shift in the absorption peak. Interestingly, both INDO/CIS and $\omega$B97XD/6-31G* yield solvatochromic shifts that are more accurate than the values computed at the B3LYP/6-31G* (MAE = 0.08 eV; $R^2 = 0.12$) and CIS/6-31G* (MAE = 0.06 eV; $R^2 = 0.59$) levels of theory (Figure 4c).
Figure 4. (a) Chemical structures of donor-acceptor substituted polyenes studied in this work. Comparison of solvatochromic shifts between toluene and ethanol solvents computed at the (b) \( \omega \text{B97XD}/6-31\text{G}^* \), INDO/CIS, (c) B3LYP/6-31\text{G}^*, and CIS/6-31\text{G}^* levels to experimental values. Data is reproduced from ref. 27.

The MOPAC2016 code also includes state-specific COSMO corrections for INDO/CISD and INDO/MRCI. To show the differences between these CI approaches, we study the behavior
of the molecule shown in Figure 5a as a function of dielectric constant; this molecule is selected from the set of donor-acceptor-substituted polyenes in Figure 4. As expected, as the dielectric constant and \( f(\epsilon) \) increase, the ground-state dipole moment increases (Figure 5b). The dipole moments at the INDO/CISD and INDO/MRCI levels are consistently 0.6-0.8 D smaller than the dipole moment at the INDO/CIS level; this is due to mixing of around 7% double excitation character into the ground state.

Figure 5. (a) Chemical structure of the prototypical donor-acceptor substituted polyene studied here. Change in (b) the computed dipole moment and (c) the first (solid line) and second (dotted line) excited-state energies of the molecule shown in panel (a) as a function of dielectric constant. \( f(\epsilon) = (\epsilon - 1)/(\epsilon + 1/2) \) ranges from \( \epsilon = 2 \) to \( \epsilon = 80 \), and \( n^2 = 2 \) for all solvents. All INDO/CISD and INDO/MRCI calculations were performed using a (2,1) active space.
Because this molecule has relatively small ground-state CT character from the donor to the acceptor, the first two excited-state energies decrease with increasing dielectric constant. Across the range of dielectric constants, the first two excited-state energies are lower at the INDO/MRCI and INDO/CISD levels than at the INDO/CIS level. The first excited-state energy in toluene \( f(\varepsilon) = 0.4 \) at the INDO/MRCI and INDO/CISD levels are closer to the experimental absorption peak (2.01 eV) than the INDO/CIS energy. The INDO/CIS solvatochromic shift for this molecule (0.14 eV) is slightly smaller than the experimental shift (0.22 eV), whereas the INDO/MRCI and INDO/CISD solvatochromic shifts are somewhat larger (0.31 and 0.30 eV, respectively). These results show that the COSMO implicit solvation in combination with various INDO/CI approaches can yield physically reasonable solvatochromic behavior.

5. **Conclusions**

We have released a new version of MOPAC2016 that incorporates the INDO/S method with a suite of CI-based excited-state capabilities and COSMO implicit solvation. This code can compute reasonably accurate ground-state and excited-state properties of both organic molecules and metal clusters up to hundreds of atoms at low computational cost, and can compute up to several thousand excited states of these systems. For a prototypical set of polyenes that have low-lying excited states with large double-excitation character, we show that INDO/MRCI yields results in good agreement with CASPT2 and experimental results, including the crossover in ordering of the excited-state energies. For a set of molecules with solvatochromic behavior, we show that the INDO/CIS solvatochromic shifts have accuracy comparable to those using the
ωB97XD functional and significantly better than those using B3LYP. We expect that these substantial new capabilities of MOPAC2016 will enable the application of INDO/S to a broad range of new systems.

Acknowledgments

The author thanks Dr. James Stewart for assistance with the MOPAC code and Prof. Jeffrey Reimers for assistance with the CNDO/INDO code. This work was supported by start-up funding at Brandeis University.

References

(1) Christensen, A. S.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. Chem. Rev. 2016, 116 (9), 5301–5337. https://doi.org/10.1021/acs.chemrev.5b00584.

(2) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. Chem. Rev. 2016, 116 (14), 7898–7936. https://doi.org/10.1021/acs.chemrev.6b00163.

(3) Sponer, J.; Bussi, G.; Krepl, M.; Banas, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jurečka, P.; et al. RNA Structural Dynamics as Captured by Molecular Simulations: A Comprehensive Overview. Chem. Rev. 2018, 118 (8), 4177–4338. https://doi.org/10.1021/acs.chemrev.7b00427.

(4) Urban, A.; Seo, D. H.; Ceder, G. Computational Understanding of Li-Ion Batteries. npj Comput. Mater. 2016, 2 (October 2015). https://doi.org/10.1038/npjcompumats.2016.2.

(5) Bedrov, D.; Piquemal, J.-P.; Borodin, O.; MacKerell, A. D.; Roux, B.; Schröder, C.
Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields. *Chem. Rev.* **2019**, *119*, 7940–7995.

https://doi.org/10.1021/acs.chemrev.8b00763.

(6) Odoh, S. O.; Cramer, C. J.; Truhlar, D. G.; Gagliardi, L. Quantum-Chemical Characterization of the Properties and Reactivities of Metal-Organic Frameworks. *Chem. Rev.* **2015**, *115*(12), 6051–6111. https://doi.org/10.1021/cr500551h.

(7) Kilina, S.; Kilin, D.; Tretiak, S. Light-Driven and Phonon-Assisted Dynamics in Organic and Semiconductor Nanostructures. *Chem. Rev.* **2015**, *115*(12), 5929–5978.

https://doi.org/10.1021/acs.chemrev.5b00012.

(8) Fernando, A.; Weerawardene, K. L. D. M. D. M.; Karimova, N. V.; Aikens, C. M. Quantum Mechanical Studies of Large Metal, Metal Oxide, and Metal Chalcogenide Nanoparticles and Clusters. *Chem. Rev.* **2015**, *115*(12), 6112–6216.

https://doi.org/10.1021/cr500506r.

(9) Kupgan, G.; Abbott, L. J.; Hart, K. E.; Colina, C. M. Modeling Amorphous Microporous Polymers for CO2 Capture and Separations. *Chem. Rev.* **2018**, *118*(11), 5488–5538.

https://doi.org/10.1021/acs.chemrev.7b00691.

(10) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annu. Rev. Biophys.* **2019**, *48*(1), 371–394. https://doi.org/10.1146/annurev-biophys-070317-033349.

(11) Nerenberg, P. S.; Head-Gordon, T. New Developments in Force Fields for Biomolecular Simulations. *Curr. Opin. Struct. Biol.* **2018**, *49*, 129–138.

https://doi.org/10.1016/j.sbi.2018.02.002.
(12) Thiel, W. Semiempirical Quantum-Chemical Methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*(2), 145–157. https://doi.org/10.1002/wcms.1161.

(13) Daniel, C. R. A.; Rodrigues, N. M.; Da Costa, N. B.; Freire, R. O. Are Quantum Chemistry Semiempirical Methods Effective to Predict Solid State Structure and Adsorption in Metal Organic Frameworks? *J. Phys. Chem. C* **2015**, *119*(41), 23398–23406. https://doi.org/10.1021/acs.jpcc.5b05599.

(14) Kromann, J. C.; Larsen, F.; Moustafa, H.; Jensen, J. H. Prediction of PKa Values Using the PM6 Semiempirical Method. *PeerJ* **2016**, *4*(2013), e2335. https://doi.org/10.7717/peerj.2335.

(15) Dral, P. O.; Wu, X.; Spörkel, L.; Koslowski, A.; Weber, W.; Steiger, R.; Scholten, M.; Thiel, W. Semiempirical Quantum-Chemical Orthogonalization-Corrected Methods: Theory, Implementation, and Parameters. *J. Chem. Theory Comput.* **2016**, *12*(3), 1082–1096. https://doi.org/10.1021/acs.jctc.5b01046.

(16) Tuna, D.; Lu, Y.; Koslowski, A.; Thiel, W. Semiempirical Quantum-Chemical Orthogonalization-Corrected Methods: Benchmarks of Electronically Excited States. *J. Chem. Theory Comput.* **2016**, *12*(9), 4400–4422. https://doi.org/10.1021/acs.jctc.6b00403.

(17) Gieseking, R. L.; Ratner, M. A.; Schatz, G. C. Semiempirical Modeling of Ag Nanoclusters: New Parameters for Optical Property Studies Enable Determination of Double Excitation Contributions to Plasmonic Excitation. *J. Phys. Chem. A* **2016**, *120*, 4542–4549. https://doi.org/10.1021/acs.jpca.6b04520.

(18) Ghosh, S.; Andersen, A.; Gagliardi, L.; Cramer, C. J.; Govind, N. Modeling Optical Spectra of Large Organic Systems Using Real-Time Propagation of Semiempirical
Effective Hamiltonians. *J. Chem. Theory Comput.* **2017**, *13*(9), 4410–4420. 
https://doi.org/10.1021/acs.jctc.7b00618.

(19) Jensen, J. H.; Swain, C. J.; Olsen, L. Prediction of PK a Values for Druglike Molecules Using Semiempirical Quantum Chemical Methods. *J. Phys. Chem. A* **2017**, *121*(3), 699–707. https://doi.org/10.1021/acs.jpca.6b10990.

(20) Husch, T.; Reiher, M. Comprehensive Analysis of the Neglect of Diatomic Differential Overlap Approximation. *J. Chem. Theory Comput.* **2018**, *14*(10), 5169–5179. 
https://doi.org/10.1021/acs.jctc.8b00601.

(21) Korth, M.; Thiel, W. Benchmarking Semiempirical Methods for Thermochemistry, Kinetics, and Noncovalent Interactions: OMx Methods Are Almost as Accurate and Robust as DFT-GGA Methods for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7*, 2929–2936. https://doi.org/10.1021/ct200434a.

(22) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-Optimization of Parameters. *J. Mol. Model.* **2013**, *19*, 1–32. https://doi.org/10.1007/s00894-012-1667-x.

(23) Hostaš, J.; Řezáč, J.; Hobza, P. On the Performance of the Semiempirical Quantum Mechanical PM6 and PM7 Methods for Noncovalent Interactions. *Chem. Phys. Lett.* **2013**, *568–569*, 161–166. https://doi.org/10.1016/j.cplett.2013.02.069.

(24) Silva-junior, M. R.; Thiel, W. Benchmark of Electronically Excited States for Semiempirical Methods: MNDO, AM1, PM3, OM1, OM2, OM3, INDO/S, and INDO/S2. *J. Chem. Theory Comput.* **2010**, *6*(5), 1546–1564. https://doi.org/10.1021/ct100030j.

(25) Voityuk, A. A. INDO/X: A New Semiempirical Method for Excited States of Organic and Biological Molecules. *J. Chem. Theory Comput.* **2014**, *10*(11), 4950–4958.
(26) Saito, T.; Kitagawa, Y.; Kawakami, T.; Yamanaka, S.; Okumura, M.; Takano, Y. Assessment of Semi-Empirical Molecular Orbital Calculations for Describing Magnetic Interactions. *Polyhedron* **2017**, *136*, 52–57.

(27) Gieseking, R. L.; Ratner, M. A.; Schatz, G. C. Implementation of INDO/SCI with COSMO Implicit Solvation and Benchmarking for Solvatochromic Shifts. *J. Phys. Chem. A* **2016**, *120*, 9878–9885. https://doi.org/10.1021/acs.jpca.6b10487.

(28) Gieseking, R. L. M.; Ratner, M. A.; Schatz, G. C. Benchmarking Semiempirical Methods To Compute Electrochemical Formal Potentials. *J. Phys. Chem. A* **2018**, *122*, 6809–6818. https://doi.org/10.1021/acs.jpca.8b05143.

(29) Frauenheim, T.; Seifert, G.; Elstner, M.; Niehaus, T.; Köhler, C.; Amkreutz, M.; Sternberg, M.; Hajnal, Z.; Di Carlo, A.; Suhai, S. Atomistic Simulations of Complex Materials: Ground-State and Excited-State Properties. *J. Phys. Condens. Matter* **2002**, *14*(11), 3015–3047. https://doi.org/10.1088/0953-8984/14/11/313.

(30) Spiegelman, F.; Tarrat, N.; Cuny, J.; Dontot, L.; Posenitskiy, E.; Martí, C.; Simon, A.; Rapacioli, M. Density-Functional Tight-Binding: Basic Concepts and Applications to Molecules and Clusters. *Adv. Phys. X* **2020**, *5*(1).

https://doi.org/10.1080/23746149.2019.1710252.

(31) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*(4), 931–948. https://doi.org/10.1021/ct100684s.

(32) Huckel, E. Quantentheoretische Beiträge Zum Benzolproblem. II. Quantentheorie Der Indezierten Polaritaten. *Z. Phys.* **1931**, *72*, 310–337. https://doi.org/10.1007/BF01341953.
(33) Huckel, E. Quantentheoretische Beitrage Zum Problem Der Aromatischen Und
Ungesattigten Verbindungen. III. Z. Phys. 1932, 76, 628–648.

(34) Hoffmann, R. An Extended Hückel Theory. I. Hydrocarbons. J. Chem. Phys. 1963, 39,
1397–1412.

(35) Pople, J. A.; Santry, D. P.; Segal, G. A. Approximate Self-Consistent Molecular Orbital
Theory. I. Invariant Procedures. J. Chem. Phys. 1965, 43, S129.

(36) Pople, J. A.; Segal, G. A. Approximate Self-Consistent Molecular Orbital Theory. II.
Calculations with Complete Neglect of Differential Overlap. J. Chem. Phys. 1965, 43,
S136.

(37) Pople, J. A.; Beveridge, D. L.; Dobosh, P. A. Approximate Self-Consistent Molecular-
Orbital Theory. V. Intermediate Neglect of Differential Overlap. J. Chem. Phys. 1967, 47,
2026–2033. https://doi.org/10.1063/1.1712233.

(38) Dewar, M. J. S.; Thiel, W. Ground States of Molecules. 38. The MNDO Method.
Approximations and Parameters. J. Am. Chem. Soc. 1977, 99, 4899–4907.
https://doi.org/10.1021/ja00457a004.

(39) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and Use of
Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum
Mechanical Molecular Model. J. Am. Chem. Soc. 1985, 107, 3902–3909.
https://doi.org/10.1021/ja00299a024.

(40) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods I. Method. J.
Comput. Chem. 1989, 10, 221–264. https://doi.org/10.1002/jcc.540100209.

(41) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification
of NDDO Approximations and Application to 70 Elements. J. Mol. Model. 2007, 13,
1173–1213. https://doi.org/10.1007/s00894-007-0233-4.

(42) Ridley, J.; Zerner, M. An Intermediate Neglect of Differential Overlap Technique for Spectroscopy: Pyrrole and the Azines. *Theor. Chim. Acta* **1973**, *32*, 111–134. https://doi.org/10.1007/BF00528484.

(43) Pierce, B. M. A Theoretical Analysis of Third-Order Nonlinear Optical Properties of Linear Polyenes and Benzene. *J. Chem. Phys.* **1989**, *91*, 791–811. https://doi.org/10.1063/1.457132.

(44) Geskin, V. M.; Lambert, C.; Brédas, J. L. Origin of High Second- and Third-Order Nonlinear Optical Response in Ammonio/Borato Diphenylpolyene Zwitterions: The Remarkable Role of Polarized Aromatic Groups. *J. Am. Chem. Soc.* **2003**, *125*, 15651–15658. https://doi.org/10.1021/ja035862p.

(45) Li, J.; Feng, J.; Sun, J. Quantum Chemical Calculations on the Spectra and Nonlinear Third-Order Optical Susceptibility of C60. *Chem. Phys. Lett.* **1993**, *203* (5), 560–564. https://doi.org/10.1016/0009-2614(93)85310-K.

(46) Gieseking, R. L.; Ensley, T. R.; Hu, H.; Hagan, D. J.; Risko, C.; Van Stryland, E. W.; Brédas, J.-L. Nonlinear Optical Properties of X(C 6 H 5 ) 4 (X = B – , C, N + , P + ): A New Class of Molecules with a Negative Third-Order Polarizability. *J. Am. Chem. Soc.* **2015**, *137*, 9635–9642. https://doi.org/10.1021/jacs.5b04377.

(47) Zerner, M. C.; Loew, G. H.; Kirchner, R. F.; Mueller-Westerhoff, U. T. An Intermediate Neglect of Differential Overlap Technique for Spectroscopy of Transition-Metal Complexes. Ferrocene. *J. Am. Chem. Soc.* **1980**, *102*, 589–599. https://doi.org/10.1021/jacs.5b00522a025.

(48) Anderson, W. P.; Cundari, T. R.; Zerner, M. C. An Intermediate Neglect of Differential
Overlap Model for Second-Row Transition Metal Species. *Int. J. Quantum Chem.* 1991, 39, 31–45.

(49) Bacon, A. D.; Zerner, M. C. An Intermediate Neglect of Differential Overlap Theory for Transition Metal Complexes: Fe, Co and Cu Chlorides. *Theor. Chim. Acta* 1979, 53, 21–54. https://doi.org/10.1007/BF00547605.

(50) Shapley, W. A.; Reimers, J. R.; Hush, N. S. INDO/S Parameters for Gold. *Int. J. Quantum Chem.* 2002, 90, 424–438. https://doi.org/10.1002/qua.10058.

(51) Gieseking, R. L.; Ratner, M. A.; Schatz, G. C. Theoretical Modeling of Voltage Effects and the Chemical Mechanism in Surface-Enhanced Raman Scattering. *Faraday Discuss.* 2017, 205, 149–171. https://doi.org/10.1039/c7fd00122c.

(52) Gieseking, R. L.; Ratner, M. A.; Schatz, G. C. Quantum Mechanical Identification of Quadrupolar Plasmonic Excited States in Silver Nanorods. *J. Phys. Chem. A* 2016, 120, 9324–9329. https://doi.org/10.1021/acs.jpca.6b09649.

(53) Gieseking, R. L. M.; Lee, J.; Tallarida, N.; Apkarian, V. A.; Schatz, G. C. Bias-Dependent Chemical Enhancement and Nonclassical Stark Effect in Tip-Enhanced Raman Spectromicroscopy of CO-Terminated Ag Tips. *J. Phys. Chem. Lett.* 2018, 9, 3074–3080. https://doi.org/10.1021/acs.jpclett.8b01343.

(54) Gieseking, R. L. M. Third-Order Nonlinear Optical Properties of Ag Nanoclusters: Connecting Molecule-like and Nanoparticle-like Behavior. *Chem. Mater.* 2019, 31, 6850–6859. https://doi.org/10.1021/acs.chemmater.9b01290.

(55) Caricato, M.; Mennucci, B.; Tomasi, J. Solvent Polarity Scales Revisited: A ZINDO-PCM Study of the Solvatochromism of Betaine-30. *Mol. Phys.* 2007, 104, 875–887. https://doi.org/10.1080/00268970500417994.
(56) Neese, F. Software Update: The ORCA Program System, Version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2018, 8 (1), e1327. https://doi.org/10.1002/wcms.1327.

(57) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16, Revision B.01. Gaussian, Inc.: Wallingford CT 2016.

(58) MOPAC2016; Stewart, J. J. P. MOPAC2016. Stewart Computational Chemistry: Colorado Springs, CO, USA.

(59) Stewart, J. J. P. MOPAC: A Semiempirical Molecular Orbital Program. *J. Comput. Aided Mol. Des.* 1990, 4, 1–105.

(60) Voityuk, A. A. Intermediate Neglect of Differential Overlap for Spectroscopy. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2013, 3 (5), 515–527. https://doi.org/10.1002/wcms.1141.

(61) Mataga, N.; Nishimoto, K. Electronic Structure and Spectra of Nitrogen Heterocycles. *Zeitschrift für Phys. Chemie* 1957, 13 (3–4), 140–157. https://doi.org/10.1524/zpch.1957.13.3_4.140.

(62) Klamt, A.; Schuurmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc. Perkin Trans.* 1993, 2, 799–805.

(63) Klamt, A. Calculation of UV/Vis Spectra in Solution. *J. Phys. Chem.* 1996, 100, 3349–3353.

(64) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* 1988, 38, 3098–3100. https://doi.org/10.1103/PhysRevA.38.3098.
(65) Dunning Jr., T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023. https://doi.org/10.1063/1.456153.

(66) Shu, Y.; Truhlar, D. G. Relationships between Orbital Energies, Optical and Fundamental Gaps, and Exciton Shifts in Approximate Density Functional Theory and Quasiparticle Theory. *J. Chem. Theory Comput.* **2020**. https://doi.org/10.1021/acs.jctc.0c00320.

(67) Aikens, C. M.; Li, S.; Schatz, G. C. From Discrete Electronic States to Plasmons: TDDFT Optical Absorption Properties of Agn (n = 10, 20, 35, 56, 84, 120) Tetrahedral Clusters. *J. Phys. Chem. C* **2008**, *112*, 11272–11279.

(68) Zhao, L.; Jensen, L.; Schatz, G. C. Pyridine-Ag20 Cluster: A Model System for Studying Surface-Enhanced Raman Scattering. *J. Am. Chem. Soc.* **2006**, *128*, 2911–2919. https://doi.org/10.1021/ja0556326.

(69) Chen, H.; Ratner, M. A.; Schatz, G. C. QM/MM Study of Photoinduced Reduction of a Tetrahedral Ag20+Cluster by a Ag Atom. *J. Phys. Chem. C* **2014**, *118*, 1755–1762. https://doi.org/10.1021/jp310149y.

(70) Walter, M.; Akola, J.; Lopez-Acevedo, O.; Jadzinsky, P. D.; Calero, G.; Ackerson, C. J.; Whetten, R. L.; Gronbeck, H.; Hakkinen, H. A Unified View of Ligand-Protected Gold Clusters as Superatom Complexes. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 9157–9162. https://doi.org/10.1073/pnas.0801001105.

(71) Tussupbayev, S.; Govind, N.; Lopata, K.; Cramer, C. J. Comparison of Real-Time and Linear-Response Time-Dependent Density Functional Theories for Molecular Chromophores Ranging from Sparse to High Densities of States. *J. Chem. Theory Comput.* **2015**, *11* (3), 1102–1109. https://doi.org/10.1021/ct500763y.
(72) Petek, H.; Bell, A. J.; Choi, Y. S.; Yoshihara, K.; Tounge, B. A.; Christensen, R. L. The 21Agstate of Trans,Trans-1,3,5,7-Octatetraene in Free Jet Expansions. *J. Chem. Phys.* 1993, 98, 3777–3794. https://doi.org/10.1063/1.464056.

(73) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. Benchmarks for Electronically Excited States: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* 2008, 128(13), 134110. https://doi.org/10.1063/1.2889385.

(74) Kulinich, A. V; Derevyanko, N. A.; Ishchenko, A. A. Synthesis, Structure, and Solvatochromism of Merocyanine Dyes Based on Barbituric Acid. *Russ. J. Gen. Chem.* 2006, 76, 1441–1457. https://doi.org/10.1134/s1070363206090167.

(75) Ishchenko, A. A.; Kulinich, A. V.; Bondarev, S. L.; Knyukshto, V. N. Electronic Structure and Fluorescent Properties of Malononitrile-Based Merocyanines with Positive and Negative Solvatochromism. *Opt. Spectrosc.* 2008, 104, 57–68. https://doi.org/10.1007/s11449-008-1008-y.

(76) Kulinich, A. V; Derevyanko, N. A.; Mikitenko, E. K.; Ishchenko, A. A. Merocyanines Based on 1,3-Indanedione: Electronic Structure and Solvatochromism. *J. Phys. Org. Chem.* 2011, 24, 732–742. https://doi.org/10.1002/poc.1821.

(77) Meyers, F.; Marder, S. R.; Pierce, B. M.; Bredas, J. L. Electric Field Modulated Nonlinear Optical Properties of Donor-Acceptor Polyenes: Sum-over-States Investigation of the Relationship between Molecular Polarizabilities (Alpha, Beta, and Gamma) and Bond Length Alternation. *J. Am. Chem. Soc.* 1994, 116, 10703–10714. https://doi.org/10.1021/ja00102a040.

(78) Gieseking, R. L.; Risko, C.; Marder, S. R.; Bredas, J.-L. Understanding the Relationships Among Molecular Structure, Excited-State Properties, and Polarizabilities of Pi-
Conjugated Chromophores. In *WSPC Reference on Organic Electronics*; 2016; Vol. 1, pp 393–419. https://doi.org/10.1017/CBO9781107415324.004.