

Extending CNDO/2 to Include Overlap Effects: Implementation and Comparative Analysis

Casey Tomlin and Jeremy Millford

Abstract

This study presents the implementation and analysis of an extended Complete Neglect of Differential Overlap (CNDO/2) method that explicitly accounts for overlap effects, comparable to the CNDO/S model. The standard CNDO/2, overlap-inclusive CNDO/S, and Modified Intermediate Neglect of Differential Overlap (MINDO/3) methods were implemented and tested on a series of hydrocarbon molecules. Bond displacement analysis was performed to assess how each method models potential energy surfaces. Results demonstrate that incorporating overlap significantly alters energy predictions. The overlap-inclusive method produces more negative total energies compared to standard CNDO/2. This work provides insight into the importance of overlap effects in semi-empirical quantum chemistry calculations.

1. Background

Semi-empirical quantum chemistry methods provide a balance between computational efficiency and accuracy by simplifying aspects of ab initio calculations and incorporating empirical parameters. The Complete Neglect of Differential Overlap (CNDO/2) method, developed by Pople and coworkers, is a foundational semi-empirical approach that simplifies calculations by neglecting differential overlap in electron repulsion integrals and using empirical parameters. Due to these simplifications, CNDO/2 tends to be more inaccurate than many of the methods that have spawned from it without offering many advantages in the way of speed.

A significant approximation in standard CNDO/2 is treating the overlap matrix (S) as an identity matrix, effectively assuming an orthonormal basis set. This approximation simplifies the computational procedure but may limit accuracy for systems where overlap effects are significant. CNDO/S, a spectroscopic variant of CNDO/2, explicitly includes overlap effects.

The Modified Intermediate Neglect of Differential Overlap (MINDO/3) represents a different semi-empirical approach with a different formulation. Contrary to Hartree-Fock theory, MINDO/3 takes care to calculate most parameters with distance in mind so those calculations tend to be more involved and theoretically more accurate than CNDO. MINDO neglects diatomic differential overlap but retains all one and two-center integrals, replacing them with simple analytic functions. Comparing these methods provides insight into the importance of overlap effects in semi-empirical calculations.

This study focuses on extending the CNDO/2 method to explicitly account for overlap effects, examining how this extension affects energy predictions for hydrocarbon molecules. In this paper we compare the results of CNDO/S and MINDO/3 with the standard ab initio Hartree-Fock theory and CNDO/2.

2. Objectives

The primary objectives of this study are to:

- Extend the CNDO/2 method to explicitly account for overlap effects, creating an implementation comparable to CNDO/S
- Compare the standard CNDO/2, overlap-inclusive CNDO/S, and MINDO/3 methods for calculating molecular energies of hydrocarbon systems

3. Methodology

3.1 CNDO/S Method: Theoretical Basis and Implementation

The CNDO/S method, developed by Del Bene and Jaffé (1968), is a spectroscopic variant of the original CNDO/2 approach. While CNDO/2 simplifies calculations by treating the overlap matrix S as an identity matrix, CNDO/S explicitly incorporates overlap effects. The key differences in our implementation include:

Explicit Overlap Matrix: In standard CNDO/2, the Hartree-Fock secular equation takes the form $FC = C\varepsilon$, where the overlap matrix S is assumed to be the identity matrix (I). This is a direct consequence of the Zero Differential Overlap (ZDO) approximation. In our CNDO/S implementation, we explicitly calculate and incorporate the overlap matrix, leading to the generalized eigenvalue problem: $FC = SC\varepsilon$. This requires solving a more complex eigenvalue problem but potentially provides a more realistic treatment of the molecular orbital coefficients.

Modified Resonance Integrals: Following Del Bene and Jaffé's approach, the resonance integrals are treated differently than in standard CNDO/2. For off-diagonal Fock matrix elements between orbitals on different centers A and B , we implemented: $F_{\mu\nu} = \beta \cdot S_{\mu\nu} - P_{\mu\nu} \cdot \gamma_{AB}$, where β is a parameter depending on the atoms involved, $S_{\mu\nu}$ is the overlap integral, $P_{\mu\nu}$ is the density matrix element, and γ_{AB} is the electron repulsion integral between centers A and B .

Orthogonalization Procedure: To solve the generalized eigenvalue problem $FC = SC\varepsilon$, our implementation followed this algorithm:

3.2 Modified Intermediate Neglect of Differential Overlap

The MINDO method advances beyond the CNDO approach by directly parameterizing integrals to fit experimental heats of atomization and molecular geometries, in contrast to conventional ab initio methods. After applying the neglect of diatomic differential overlap approximation, the remaining one- and two-center integrals are substituted with simplified parametric functions using parameters detailed in **Table 1**. Rather than employing fixed, predetermined values for atoms and orbitals, MINDO calculates these integrals to precisely align with the specific molecular context, explicitly incorporating quantifiable orbital overlap contributions.

The term U represents the one-center orbital energies for s and p orbitals, while the g terms

correspond to two-electron resonance integrals between orbitals localized on the same atom, effectively replacing the exact integrals typically precomputed in ab initio methods.

The Wolfsberg-Helmholtz factors integrate the effects of interatomic distances and intrinsic orbital energies into a single parameter, capturing the relative energetic variations associated with changes in bond lengths:

$$b_{AB} = \frac{H_{\mu\nu}}{S_{\mu\nu}}$$

Parameter	Hydrogen	Carbon	Nitrogen	Oxygen
U_s	-12.505	-51.79	-66.06	-91.73
U_p	-	-39.18	-56.40	-78.80
g_{ss}	12.848	12.23	13.59	15.42
g_{sp}	-	11.47	12.98	14.52
g_{pp}	-	11.08	12.66	14.48
$g_{pp'}$	-	9.84	11.59	12.98

Table 1: U and g parameter functions for commonly used atoms.

Mulliken orbital populations are computed by summing the distance-weighted electron densities, assigning electron density to each atomic orbital based on the density matrix and orbital overlap:

$$q_\mu = \sum_\nu (P_{\mu\nu} S_{\mu\nu})$$

The Fock matrix calculation integrates both resonance terms and electrostatic corrections, whereas the core Hamiltonian explicitly accounts for self-repulsion and interatomic charge interactions:

$$F_{\mu\nu} = S_{\mu\nu}(U_\mu + U_\nu) b_{AB} - p_{\mu\nu} \gamma_{AB}(R_{AB})$$

$$H_{\mu\mu} = U_\mu - (q_\mu - 1/2) \gamma_{AA} - \sum_{B \neq A} (q_B - Z_B) \gamma_{AB}$$

3.3 Implementation Framework

All three methods (CNDO/2, overlap-inclusive CNDO/S, and MINDO/3) were implemented in C++ with the Eigen linear algebra library. The core components included:

1. **Self-Consistent Field (SCF) procedure:** For all methods, the SCF procedure involved iteratively calculating the Fock matrix, solving the eigenvalue problem, and updating the

density matrix until convergence.

2. **Convergence techniques:** Due to the more challenging convergence behavior of the overlap-inclusive method, convergence acceleration techniques were implemented, including density matrix damping.
3. **Energy calculation:** Total energy was computed as the sum of electronic energy and nuclear repulsion energy, with method-specific formulations for the electronic component.
4. **Bond displacement analysis:** Bond distances were systematically varied from -0.2 to +0.2 Å in 0.05 Å increments, with energy calculations at each step.

3.4 Test Systems and Reference Data

The test set included the following hydrocarbon molecules:

- Hydrogen molecule (H_2)
- Methane (CH_4)
- Ethylene (C_2H_4)
- Ethane (C_2H_6)
- Benzene (C_6H_6)
- Butane (C_4H_{10})

For each molecule a total energy calculation was performed by adding the nuclear repulsion energy and electronic energy where the differences among all models was the electronic energy portion. To provide reference data for comparison, we also performed Hartree-Fock calculations using Psi4, a quantum chemistry software package, in Python. These calculations provided high-quality ab initio energy values against which the semi-empirical methods could be compared.

3.5 Implementations attempted

There are a number of implementations we spent significant time on but were ultimately unsuccessful at or either extracting results from.

For CNDO/S we attempted to tune parameters of beta H and C as well as scaling according to literature (Voityuk, 2005) but this method can easily result in overfitting and to implement such a method well many data points and re-evaluations are required. Due to complexity and SCF convergence issues we were unable to really delve into this method as much as we would have liked.

We Also attempted some methods to evaluate geometries using displacement and bond length evaluation but we were unfortunately unable to generate results we found meaningful, but not for a lack of trying.

4. Results and Discussion

4.1 Effect of Overlap Inclusion on Total Energies

Molecule	HF (Reference)	CNDO/2	CNDO/S	MINDO/3
H ₂	-30.4095	-40.5751	-35.1350	-18.6893
CH ₄	1093.7673	-181.2029	-905.0085	144.38
C ₂ H ₄	2097.2971	-247.2018	-1815.6449	599.002
C ₂ H ₆	2155.9255	-284.4417	-2374.8769	746.806
C ₄ H ₁₀	4280.3096	740.7236	-6822.2945	2356.14
C ₆ H ₆	6277.7466	-717.8334	-8943.6305	3267.23

Table 2: Total energies of hydrocarbons for Hartree-Fock, CNDO/2, CNDO/S and MINDO/3 methods.

The inclusion of overlap effects dramatically changes the energy predictions. For all molecules except H₂, the overlap-inclusive CNDO/S method produces much more negative (lower) total energies than both standard CNDO/2 and MINDO/3. This systematic effect becomes more pronounced with increasing molecular size.

For methane (CH₄), the overlap-inclusive calculation (-905.01 eV) yields an energy approximately 5 times more negative than the standard CNDO/2 result (-181.20 eV). This effect is even more dramatic for larger molecules like benzene, where the ratio increases to nearly 6.7. This may be due to a confounding error that is causing the calculation to increase a lot or to improper parameters set for the model. Seeing as CNDO/S is designed for unmatched orbital spins, this may be the cause of the model's drift from CNDO/2.

Interestingly the MINDO/3 values are all positive above H₂, as seen in **Table 2**, which reflect the tuning of the MINDO/3 parameters for predicting heats of atomization and molecular properties. This makes it less predictive for hydrocarbon electronic energies but may be able to be tuned for such a usage, but is outside the scope of this work.

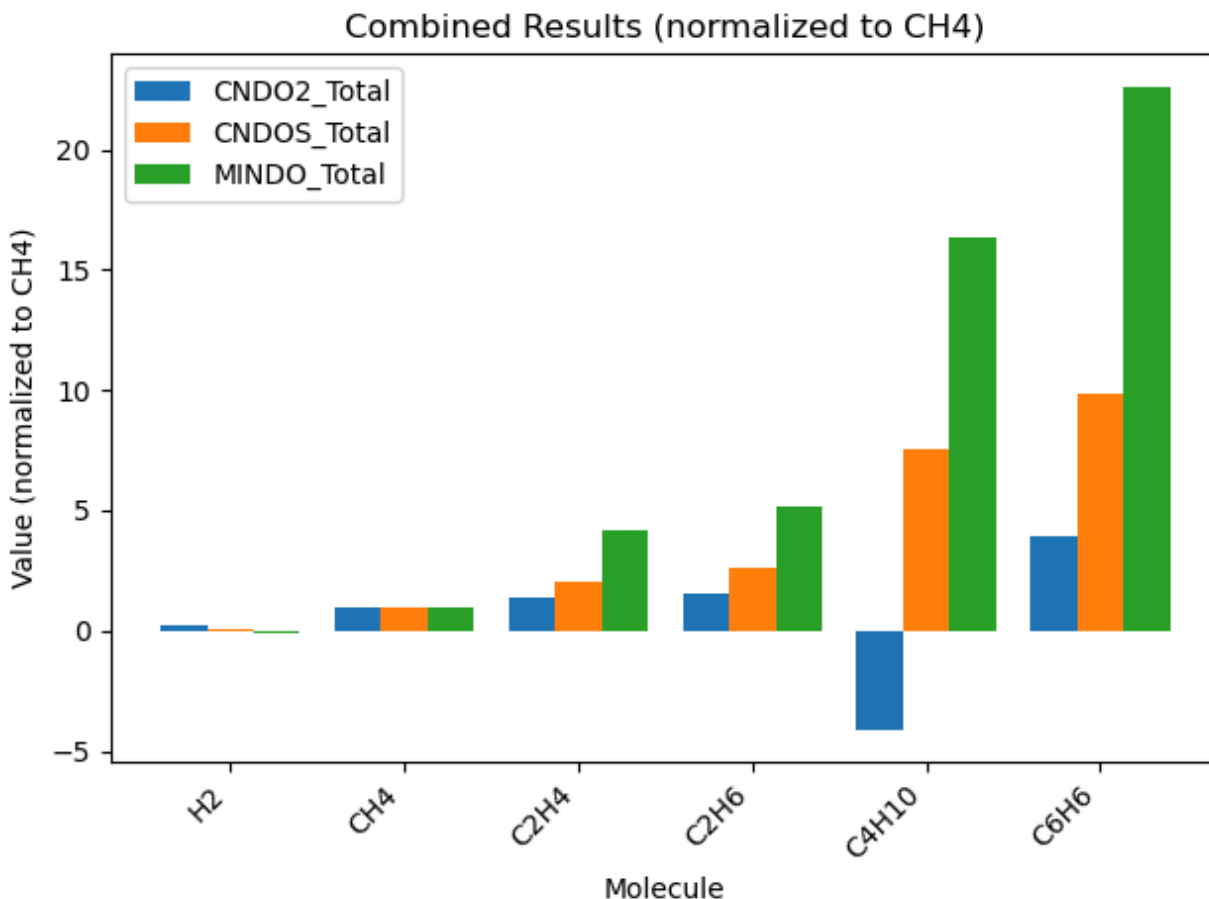


Figure 1: Combined results of total energies for various hydrocarbons from the CNDO/2, CNDOS and MINDO/3 models normalized to methane

In **Figure 1** the trends indicate that the predicted energies for the methods produced for this paper had higher energies relative to CH₄. CNDO/2 shows an anomaly when calculating C₄H₁₀ which may be explained by the limited basis set of the method showing a larger relative difference with respect to CH₄. MINDO/3 on the other hand shows larger relative differences which point to parameters that have been tuned for calculations other than total energy, potentially heats of atomization. CNDOS shows the lowest, most normal differences indicating it may be the best predictor for the given task.

Comparing with the Hartree-Fock reference values, we observe that all semi-empirical methods significantly underestimate the total energies, with MINDO/3 producing results closest to the HF values in terms of general trend but still with substantial differences. This highlights the inherent limitations of semi-empirical approaches compared to ab initio methods. The fine-tuned adjustments of the parameters due to more specialized calculations for orbital dependencies means that MINDO/3 can better accommodate various molecules and provide more accurate energies to experiment.

This substantial difference in electronic energy can be attributed to:

1. Explicit treatment of non-orthogonality between basis functions

2. Different treatment of electron-electron interactions when overlap is included

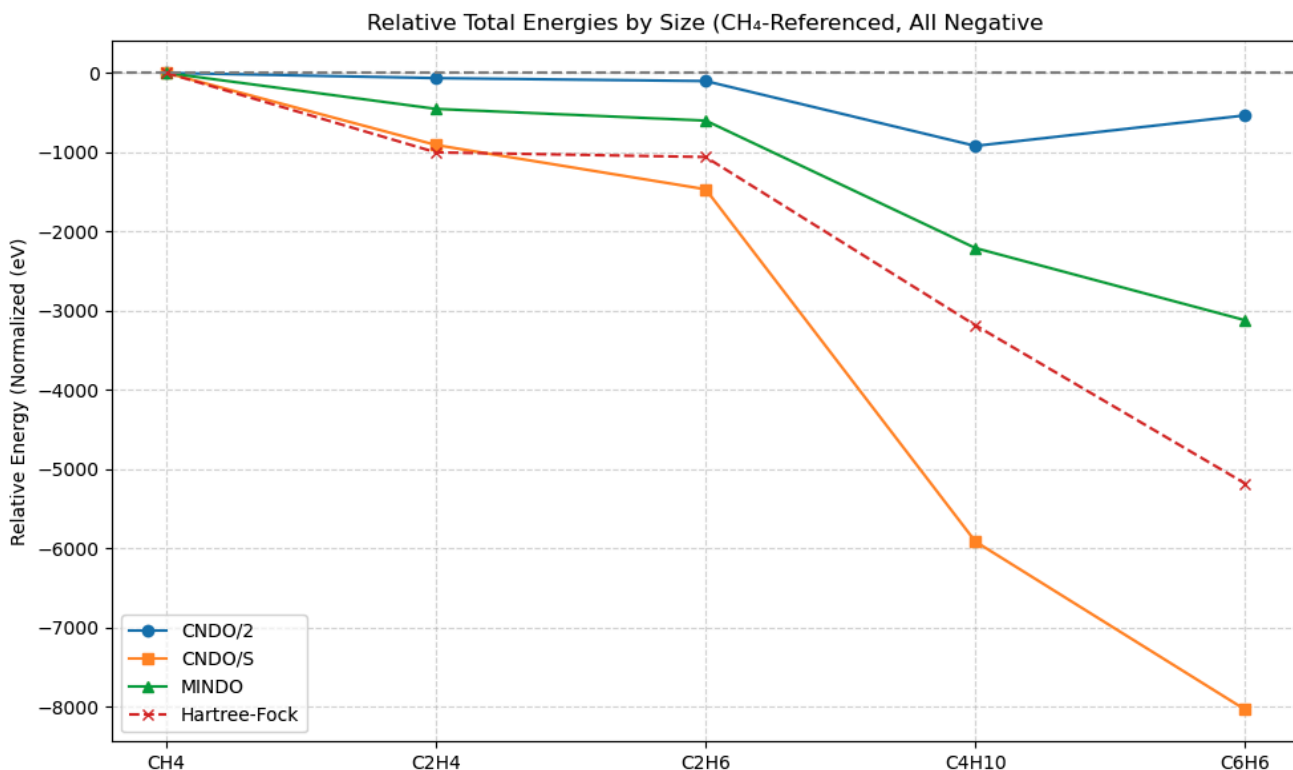


Figure 2: Relative total energies compared across models, normalized to CH₄

We can see from **Figure 2**, that both CNDO/S and MINDO follow the trend of our benchmark Hartree Fock values leading to more negative values as structure becomes more stable. For visualization we normalized the sign of the energy.

4.2 Computational Implications of Overlap Inclusion

Including overlap effects significantly impacts the computational behavior of the methods:

SCF Convergence:

- Standard CNDO/2: Typically converged within 25-40 iterations
- Overlap-inclusive CNDO/S: Often required 100+ iterations and sometimes did not fully converge within the standard 150-iteration limit. We attempted implementing level scaling, DIIS, and perturbation in delta to achieve better convergence but didn't have success. CNDO/S is known to have difficulties converging. We did plot delta values to check that values were at least stabilizing.
- MINDO/3: Required 40-70 iterations typically

Convergence Stability: The overlap-inclusive method showed oscillatory convergence behavior that required stronger damping. As MINDO/3 takes into account more molecule

information it tends to converge faster and with greater accuracy than the CNDO methods.

Computational Cost: For methane, CNDO/S took over twice as long to run compared to standard CNDO/2, while MINDO/3 required moderately more time than CNDO/2. This highlights the additional computational expense introduced by including overlap in CNDO/S. These observations highlight a fundamental trade-off: including overlap effects potentially improves the physical basis of the model but introduces significant computational challenges. Despite the requirement of MINDO/3 to perform more computations, it tends to run in shorter time than CNDO/S and with greater precision over the SCF loop since the additional parameters tend to add few complexities.

5. Conclusions

This study on extending CNDO/2 to include overlap effects leads to several important observations:

1. **Energy Calculations:** Explicit inclusion of overlap in the CNDO/2 framework (creating the CNDO/S method) dramatically changes energy predictions, resulting in much more negative total energies. This effect becomes more pronounced with increasing molecular size. We expect the values to become more negative as stability increases.
2. **Computational Impact:** Including overlap significantly increases computational cost and introduces convergence challenges. The overlap-inclusive calculations often require more iterations and show oscillatory convergence behavior.
3. **Method Comparison:** The three methods (CNDO/2, CNDO/S, and MINDO/3) show substantial differences in their energy predictions, with all semi-empirical methods producing relative energies that are proportional to the Hartree-Fock reference values from Psi4.

These findings illustrate how the explicit treatment of overlap effects significantly influences semi-empirical calculations, affecting energetic predictions and computational behavior. While no semi-empirical method approaches the accuracy of ab initio calculations, the comparison between methods highlights the trade-offs involved in different approximations and their effects on computational results.

6. Future Work

Several promising directions for further investigation include:

- Investigating alternative approaches to handling overlap that might maintain accuracy while reducing computational cost.
- Exploring the effects of overlap inclusion on other molecular properties beyond energies
- Developing more robust convergence techniques specifically designed for overlap-inclusive methods

7. References

1. Pople, J.A.; Beveridge, D.L. Approximate Molecular Orbital Theory; McGraw-Hill: New York, 1970.
2. Del Bene, J.; Jaffe, H.H. Use of the CNDO Method in Spectroscopy. I. Benzene, Pyridine, and the Diazines. *J. Chem. Phys.* 1968, 48, 1807-1813.
3. 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.
4. Szabo, A.; Ostlund, N.S. Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory; Dover Publications: Mineola, NY, 1996.
5. Jensen, F. Introduction to Computational Chemistry, 3rd ed.; Wiley: Hoboken, NJ, 2017.
6. Parrish, R.M.; Burns, L.A.; Smith, D.G.A.; Simmonett, A.C.; DePrince, A.E.; Hohenstein, E.G.; Bozkaya, U.; Sokolov, A.Y.; Di Remigio, R.; Richard, R.M.; et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* 2017, 13, 3185-3197.
7. Voityuk, A. A. *Thermochemistry of Hydrocarbons. Back to Extended Hückel Theory*. *J. Phys. Chem. A* 2005, 109 (20), 4429–4433
8. Bingham R, Dewar M and Lo D 1975 Additions and Corrections - Ground States of Molecules. XXV. MINDO/3. An Improved Version of the MINDO Semiempirical SCF-MO Method. *J. Am. Chem. Soc.* **97** 4787–4787