

# Transition probability approach for direct calculation of coefficients of Configuration Interaction wave function

Arijit Bag<sup>1\*</sup>

<sup>1</sup>Division of Chemical Sciences

Indian Institute of Science Education and Research Kolkata

Mohanpur, Nadia, WB, India 741246

To reduce the computation cost of Configuration Interaction (CI) method, a new technique is used to calculate the coefficients of doubly excited determinants directly from orbital energies, orbital overlap matrix and electron population obtained from Hartree Fock level run. This approach to approximate the coefficients of CI wave function is termed as transition probability approximated CI (TPA-CI). In principle, calculated dynamical electron correlation energy of TPA-CI and Full CI (FCI) are equivalent. It is observed that computed TPA-CI correlation energies of hydrogen, water, ammonia and ozone are very close to FCI values, within 5% error. The potential energy curve of the hydrogen molecule is also studied and it is found that the energy is minimum at its equilibrium bond length.

**Keywords:** Transition Probability Approximation, Configuration Interaction, Potential Energy Curve

## I. INTRODUCTION

Among the post Hartree-Fock (HF) [1, 2] methods, Configuration Interaction (CI) [3, 4] and Coupled Cluster (CC) [5–7] methods are most popular and treated as a benchmark for all computational calculations. Dynamical electron correlation has been treated properly in CI formulation. Yet we require good alternatives as CI is computationally too expensive to calculate Full CI energy even for a tri- or tetra- atomic molecule with a reasonable basis set. Coupled Cluster on the other hand, along with Møller-Plesset (MP) Perturbation theory [8, 9] are two good alternatives. Though full CI and full CC are equivalent, truncated CC has advantages over truncated CI as truncated CI is not size consistent, but truncated CC is size consistent. Not only that, due the presence of the exponential character of the CC wave function, truncated CC takes care of the contribution of the higher excited determinants. Thus, truncated CC is better than truncated CI at the same level of truncation. Presently, coupled cluster methods are frequently used for property calculation of small molecules [10–14]. But, due to high requirement of computing facilities neither CI nor CC is used for computer aided molecular modeling, drug designing or reaction mechanism studies where currently only density functional theory (DFT) based methods are used.

Energy calculation using CI only depends on the coefficients of doubly excited determinants. But, so far, there is no way to calculate coefficients of doubly excited determinants of CI wave function as equations we use to solve are coupled with other classes (singly excited, triply excited, etc.) of determinants. In general, we use iterative way to solve these equations which requires huge computational facilities. Any technique to obtain values of coefficients of doubly excited determinants of CI wave function would help to implement CI method for larger systems where currently only DFT based methods are applicable.

In the present approach, termed as Transition Probability Approximated CI (TPA-CI), the coefficients of doubly excited determinants are calculated using transition probability of electrons from occupied orbitals to the virtual orbitals. The transition probabilities are calculated from electron population, orbital energy and overlap integrals.

Theoretical details of TPA-CI are presented in the theory section. Computation methodology and computation cost analysis are presented in computation section. In results and discussion section TPA-CI correlation energy of the  $H_2$  molecule calculated in four different basis is presented. The potential energy curve of  $H_2$  molecule is also studied. TPA-CI correlation energies of three poly atomic molecules; water, ozone and ammonia, are also calculated and compared with respective FCI results.

## II. THEORY

CI wave function is as follows -

$$|\Psi\rangle = C_0 |\Psi_0\rangle + \sum_{a,r} C_a^r |\Psi_a^r\rangle + \sum_{a<b,r<s} C_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \dots \quad (1)$$

---

\*Electronic address: [bagarijit@gmail.com](mailto:bagarijit@gmail.com), [bagarijit@iiserkol.ac.in](mailto:bagarijit@iiserkol.ac.in)

Now, from Brillouin's theorem we get  $\langle \Psi_0 | H | \Psi_a^r \rangle = 0$ . The triple and higher excitations do not mix with  $|\Psi_0\rangle$  as they differ from  $|\Psi_0\rangle$  by more than two orbitals. Thus, the expression for CI correlation energy ( $E_{corr}$ ) is

$$E_{corr} = \sum_{a,b,r,s} C_{ab}^{rs} \langle \Psi_0 | H | \Psi_{ab}^{rs} \rangle \quad (2)$$

where,  $a$  and  $b$  are occupied orbitals and  $r$  and  $s$  are virtual orbitals. In Equation 2, different co-efficients of doubly excited determinants like  $C_{ab}^{rs}$  are unknown variables. In CI, these unknown variables are evaluated by an iterative method which requires diagonalization and renormalization. Not only that, doubly excited determinants are coupled with other excited determinants. Thus, computation time and storage requirement increases exponentially with increase of number of basis. In present research work an alternative method is used to find the values of these coefficients. In this approach values of the coefficients of doubly excited determinants like  $C_{ab}^{rs}$  are calculated from the transition probability of electrons from  $a, b$  (occupied) orbitals to  $r, s$  (virtual) orbitals. Reason is that the excited determinants in CI wave function are constructed by replacing one or more occupied atomic orbitals (A.O.) depending upon the order of excitations, with the same number of virtual orbitals. Thus, coefficient of any specific order excited determinant would be the transition probability of that excitation. In the present approach (TPA-CI), it is considered that the transition probability of electrons from orbitals  $a$  and  $b$  which belong to occupied orbitals, to orbitals  $r$  and  $s$  which belong to unoccupied orbitals, is proportional to the electron density at  $a$  and  $b$ , orbital overlap between  $a$  and  $r$  ( $S_a^r$ ) and  $b$  and  $s$  ( $S_b^s$ ) and partition of electron distribution probability between  $a$  and  $r$ , and  $b$  and  $s$ . Partition of electron distribution probability is taken as an exponential function of the orbital energy difference *i. e.*  $e^{(E_a-E_r)}$  and  $e^{(E_b-E_s)}$ . For non-degenerate states, the transition probability is inversely proportional to their orbital energy difference. If we consider proportionality constant as unity we get an expression for the coefficients of doubly excited determinants as follows -

$$C_{ab}^{rs} = \frac{\rho_a \rho_b S_a^r S_b^s e^{-\Delta E_{ar} \Delta E_{bs}}}{\Delta E_{ar} \Delta E_{bs}} \quad (3)$$

where  $\rho_a$  and  $\rho_b$  are the electron densities of orbitals  $a$  and  $b$ ,  $\Delta E_{ar}$  and  $\Delta E_{bs}$  are orbital energy difference between  $a$  and  $r$  and  $b$  and  $s$  respectively. Thus, the correlation energy expression of TPA-CI is as follows -

$$E_{corr} = \frac{1}{8} \sum_{a,b,r,s} C_{ab}^{rs} \langle rs | ab \rangle \quad (4)$$

As values of co-efficients of doubly excited determinants are evaluated using equation 3 we can say, in principle, calculated correlation energy using equation 4 is an approximation to FCI correlation energy.

### III. COMPUTATION

#### A. Computational details

In this work GAMESS [15] (Version = 6 SEP 2001 (R5) from Iowa State University) is used for HF calculation. Mulliken charge population is used as electron densities of respective atomic orbitals. Total energy is the sum of HF energy and TPA-CI correlation energy. Overlap matrix is taken from the GAMESS out put file. Two electron integrals in molecular orbital basis are used in equation 4. Calculation of co-efficients of doubly excited determinants and correlation energy are done using in home code.

#### B. Computation cost analysis

In TPA-CI computation, first, we have to run HF; similar to other post HF methods. Thus, computation time and requirement up to this point is same as CI, CC and Perturbation methods. After HF run, we have to compute two electron integrals in molecular orbital basis from two electron integrals in atomic basis and MO coefficients. This part is the most time taking part in TPA-CI calculation. But, time and computation cost is same as other methods. Calculation of coefficients of doubly excited determinants and correlation energy is done in four nested 'do' loops among which two outer loops run from 1 to  $m$  where  $m$  is the number of occupied orbitals and two inner loops run from  $m+1$  to  $N$  where  $N$  is the total number of basis. Thus, at this stage we don't need any storage for correlation energy calculation. Though, we can store the coefficients of doubly excited determinants for further calculation. But, in TPA-CI, we get these coefficients directly. There is no need for diagonalization or reconstruction as it requires in CI calculation. Time requires for calculation of correlation energy in TPA-CI is proportional to  $m^2(N-m)^2$  which is much less than  $N^4$  for large value of  $N$ . On the other hand, in general, time required for energy calculation in CI and CC are proportional to  $N^5$ .

TABLE I: TPA-CI energy of Hydrogen molecule in different basis sets

Basis	RHF Energy (a.u.)	TPA-CI correlation energy (a.u.)	FCI correlation energy (a.u.)	Error in correlation energy (%)
STO-3g	-1.1167	0.02396	0.02056	16.57
6-31G	-1.1267	0.02611	0.02494	4.72
cc-pVDZ	-1.1287	0.03593	0.03469	3.55
cc-pVTZ	-1.1330	0.03839	0.03947	2.74

#### IV. RESULTS AND DISCUSSION

The TPA-CI correlation energy of hydrogen molecule in the four different basis is calculated and compared with FCI results to test the efficacy of the present method. The potential energy curve of hydrogen molecules is studied to check whether the present theory is consistent with distorted geometry or not. The TPA-CI correlation energy of ozone, water and ammonia are also computed and compared with FCI correlation energy to test the applicability of the present method for polyatomic molecules.

##### A. Comparison of TPA-CI and FCI correlation energy

Hydrogen molecule is very small and hence FCI calculation in different basis is possible. Thus, this molecule is used for comparison of TPA-CI and FCI correlation energy. STO-3G, 631-G, cc-pVDZ and cc-pVTZ basis are used for this purpose. An equilibrium bond length, which is 1.401 a.u.[16] is used for all calculations. Calculated results are presented in Table I. TPA-CI correlation energies agree very well with FCI correlation energies except a STO-3G basis, which is 16.57% higher than corresponding FCI correlation energy. It is observed that TPA-CI correlation energies of STO-3G, 631-G and cc-pVDZ basis are higher than corresponding FCI correlation energies though the percentage of error decreases with increase of number of basis functions. This may be due to high electron density in the occupied orbitals when the number of basis is less. But the TPA-CI correlation energy of hydrogen molecules in cc-pVTZ basis is less than the FCI correlation energy and also error is only 2.74%. Thus, we may conclude that for larger basis TPA-CI and FCI yields equivalent results.

##### B. Study of potential energy curve

To test the accuracy of the present approach at different distorted geometry, potential energy curve of hydrogen molecules is studied taking 13 different  $H-H$  bond length, both sides of its equilibrium bond length. cc-pVDZ basis is used for this purpose. It is observed that potential energy curve of HF is above TPA-CI curve and FCI curve which is expected. But TPA-CI curve is above FCI curve for distorted bond length and below FCI curve around the equilibrium bond length which is very interesting. Nature of TPA-CI curve is also different from HF and FCI curve. HF and FCI curve is very smooth all along the plot while, TPA-CI curve has a pocket like structure around equilibrium bond length. The reason is still unknown. It may be due to the contribution of singly, triply or higher excited determinants to the ground state energy at any distorted geometry. In the present work, TPA-CI ground state energy is calculated taking only the coefficient of doubly excited determinant. Thus, at equilibrium geometry both FCI and TPA-CI yields nearly same energy. But, for distorted geometries while FCI calculation includes contributions from other excited determinants TPA-CI calculation takes only the contribution of doubly excited determinants only. Thus, we observe discrepancies between FCI and TPA-CI energies for distorted geometries. Equilibrium bond length calculated from potential energy curve of TPA-CI (1.4a.u.) matches very well with experimental value (1.401a.u. [16]) compare to HF (1.425a.u.) and FCI (1.45a.u.).

##### C. TPA-CI correlation energy for polyatomic molecules

TPA-CI is also implemented for polyatomic molecules. Ozone, water and ammonia are used for this purpose. For all three molecules, first geometry optimization is done at the HF level with same basis and then TPA-CI and FCI calculation is performed

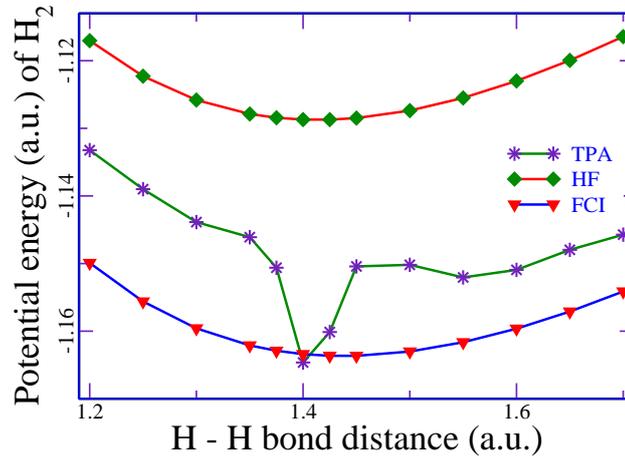


FIG. 1: Potential energy curve of hydrogen molecule

TABLE II: TPA-CI correlation energy of ozone, water and ammonia

Molecule	Basis	TPA-CI correlation energy (a.u.)	FCI correlation energy (a.u.)	Error in correlation energy (%)
$O_3$	DZ(dunning)	0.22416	0.23807	5.84
$NH_3$	631-G	0.12353	0.12846	3.84
$H_2O$	cc-pVDZ	0.06955	0.07168	2.97

separately. Calculated results are presented in Table II. For all three molecules error in TPA-CI correlation energy with respect to FCI is very small. We can get 94 – 97% of correlation energy using TPA-CI technique. Here we observe that TPA-CI correlation energies are less than the FCI correlation energy though small basis are used. This may be due to the use of optimized geometry which may be different from exact geometry and we have observed that (described in the previous subsection) at non-equilibrium geometry TPA-CI correlation energy is less than the FCI correlation energy irrespective of basis size.

## V. CONCLUSIONS

From present study, we can conclude that this new approach which is termed as TPA-CI, is applicable for electronic structure calculation of molecules. Its accuracy in dynamical electron correlation energy calculation is equivalent to FCI at moderate sizes of basis. As a requirement of computation for TPA-CI technique is much less than FCI, we can use it for bigger molecules. Its application limit is same as DFT. Thus, this approach would promote CI method to use for real problems. We need further development in computation (coding) for property calculation using this approach.

## VI. ACKNOWLEDGMENTS

Author thanks IISER Kolkata for financial support and Dr. Pradip Ghorai, author's post doctorate research guide to give an opportunity to publish this work independently.

- 
- [1] Fock V., *Z. Physik* 1930, **61**, 126.
  - [2] Pauli W., *Z. Physik* 1925, **31**, 756.
  - [3] Boys S. F., *Proc. Roy. Soc. (London)* 1950, **A201**, 125.
  - [4] Pople J. A. and Binkley J. S., *Int. J. Quant. Chem. Symp.* 1976, **10**, 1.
  - [5] Cizek J., *J. Chem. Phys.* 1966, **45**, 4256.

- [6] Cizek J., *Adv. Chem. Phys.* 1969, **14**, 35.
- [7] Purvis D. and Bartlett R. J., *J. Chem. Phys.* 1982, **76**, 1910.
- [8] Møller C. and Plasset M. S., *Phys.Rev.* 1934, **46**, 618.
- [9] Leininger M. L., Allen W. D., Schaefer H. F. and C.D.Sherrill C. D., *J.Chem.Phys.* 2000, **112**, 9213.
- [10] Bag A., Manohar P. U., Vaval, N. and Pal S., *J. Chem. Phys.* 2009, **131** 024102.
- [11] Bag A., Manohar P. U. and Pal S., *Computing Letters* 2007, **3**, 351.
- [12] Manohar P. U., Shamasundar K. R., Bag A., Vaval N., Pal S., *Challenges and advances in computational chemistry and physics* 2010, **11**, 375.
- [13] Bag A., Bhattacharya S. and Pal S. , *Recent Advances in Spectroscopy :theoretical, astrophysical and experimental perspectives* edited by R. K. Chaudhury, M. V. Mekkaden, A. V. Ravveendran, A. Satyanarayanan, 2010, 99.
- [14] Bag A., *Linear response approach in FSMRCC for molecular property* LAMBERT Academic Publishing, Deutschland, Germany, 2016.
- [15] GAMESS, General Atomic and Molecular Electronic Structure System; *J. Comput. Chem.* 1993, **14**, 1347.
- [16] Kolos W. and Wolniewicz L., *J. Chem. Phys.* 1968, **49**, 404.