

Fragmentation of NOON-based diagnostic for assessing the accuracy of SD FNDMC in complexes with hydrogen bonds.

Martin Šulka¹ and Matúš Dubecký^{1,2}

¹ Slovak University of Technology in Bratislava, Faculty of Materials Science and Technology in Trnava, Advanced Technologies Research Institute, Jána Bottu 25, 917 24 Trnava, Slovakia
² Department of Physics, Faculty of Science, University of Ostrava, 30. dubna 22, 70103 Ostrava, Czech Republic

(3)

(5)

(6)



Introduction

Fixed-node diffusion Monte Carlo (FNDMC) calculations of large noncovalent systems rely on compact yet effective trial wave function Ψ_T , very often composed of single Slater determinant (SD). The nonequivalence in Slater determinants of interacting complex and its non-interacting subsystems may lead to innacurate FNDMC energy differences. Recently, Dubecký[1] introduced a measure of SD nonequivalence, based on the Euclidean distance between the natural orbital occupation number (NOON) vector of the SD and the exact solution in the NOON vector space

$$d(\overrightarrow{\lambda}, \overrightarrow{\lambda}^{SD}) = \frac{1}{\sqrt{N}} \|\overrightarrow{\lambda} - \overrightarrow{\lambda}^{SD}\| = \sqrt{\frac{\sum_{i} (\lambda_{i} - \lambda_{i}^{SD})^{2}}{N}}.$$
 (1)

obtained from the qualitatively correct correlated size-consistent post-HF computa-

Results

Figure 1: Fragmentation scheme of (a) formic acid dimer "FAD", (b) rotated formic acid dimer "rot-FAD", (c) methane-diol "MD" that represents a hydrogenated FAD, (d) adenine thymine model complex "ATM", (e) succinamide dimer complex "SAD", and, (f) butadiene-argon complex "But-Ar".



tion, where N is the number of explicitly correlated electrons in the post-HF treatment. Thus, d measures the nondynamic correlation content of a given state. Measure of SD non-equivalence, δ , is defined as

$$\delta = d^C - d^S,\tag{2}$$

obtained from scaled distances d^S and d^C for supersystem S and its non-interacting counterpart C, evaluated at CCSD/aug-VDZ level. For truly large systems, its estimation, however, becomes intractable. The goal of this work is to see if one can approximate δ (Eq. 2) by $\tilde{\delta}$ from fragment calculations. This would help in significant reduction of CPU cost and would make it possible to predict FNDMC inaccuracy in large complexes.

Methods

For a complex *AB*, consisting of fragments *A* and *B*, let us assume that *A* consists of subsystems *A'* and *A''* and there is a single covalent bond between them, A = A'A'' and similarly, B = B'B'', so that complexes may well be fragmented without the conjugation breaking. We need to settle how to compute $\delta(A'A'', B'B'', A'A''B'B'')$. Number of noncovalent complexes were chosen as test cases in development of fragmentation technique. These were fragmented into fragments as shown in Figure 1. The resulting fragments were H-saturated and partly re-optimized (unsaturated part of fragment remained rigid) in subsequent calculations. δ values were evaluated at CCSD/cc-pVDZ level. Both δ and Interaction Energy (IE) were decomposed to many-body terms



Table 1: Many-body decomposition of δ and Interaction Energy (IE) in kJ/mol evaluated at CCSD/cc-pVDZ level.

	FAD		rot-FAD		ATM		SAD	
fragment	IE	$ ilde{\delta}$	IE	$ ilde{\delta}$	IE	$ ilde{\delta}$	IE	$ ilde{\delta}$
A'B'	-19.64	0.000331	-4.25	-0.000059	-8.41	0.000263	-33.12	0.000144
A'B''	-9.25	0.000542	-18.01	0.000602	-14.92	0.000174	-1.02	-0.000071
A″B′	-9.25	0.000542	-0.02	-0.000064	-3.43	-0.000067	-4.09	0.000019
A''B''	2.04	-0.000227	-2.13	-0.000099	-3.94	0.000214	-33.12	0.000144
$ ilde{\delta}^2$	-36.10	0.001188	-24.41	0.000380	-30.71	0.000585	-71.35	0.000236
$\sum ilde{\delta}^i$	-36.10	0.001188	-24.41	0.000380	-30.71	0.000585	-71.35	0.000236
A'A''B'	-17.42	-0.000088	2.30	-0.000014	-5.73	-0.000090	0.52	-0.000050
A'A''B''	-1.00	-0.000061	-4.26	-0.000135	-14.87	0.000002	0.80	-0.000023
A'B'B''	-17.42	-0.000088	-6.70	-0.000123	-10.23	0.000090	0.80	-0.000023
A''B'B''	-1.00	-0.000061	3.00	0.000055	-6.93	0.000106	0.52	-0.000050
$ ilde{\delta}^3$	-36.85	-0.000297	-5.66	-0.000218	-37.76	0.000108	2.63	-0.000145
$\sum ilde{\delta}^i$	-72.95	0.000891	-30.07	0.000163	-68.47	0.000693	-68.72	0.000091
A'A"B'B"	-11.74	-0.000112	0.10	0.000051	-3.81	-0.000118	-1.46	0.000033
$ ilde{\delta}^4$	-11.74	-0.000112	0.10	0.000051	-3.81	-0.000118	-1.46	0.000033
$\sum ilde{\delta}^i$ / Total	-84.69	0.000780	-29.97	0.000214	-72.28	0.000575	-70.18	0.000124

 $\delta = \tilde{\delta}^2 + \tilde{\delta}^3 + \tilde{\delta}^4$

It has to be noted that intra-molecular fragment interactions, i.e. A'A" and B'B" interactions do not contribute to the overall δ . Each m-body term is $\tilde{\delta}$ of a given m-body fragment, from which the sum of all containing lower than m-body $\tilde{\delta}$ are subtracted. In case of ATM, two body terms are

$$\tilde{\delta}_{A'B'} = d^{C}_{A'B'} - d^{S}_{A'B'} \qquad \tilde{\delta}_{A'B''} = d^{C}_{A'B''} - d^{S}_{A'B''}$$

$$\tilde{\delta}_{A''B''} = d^{C}_{A''B'} - d^{S}_{A''B'} \qquad \tilde{\delta}_{A''B''} = d^{C}_{A''B''} - d^{S}_{A''B''}$$
(4)

an example of a three-body term, from which the contained two-body terms must be subtracted

 $\tilde{\delta}_{(A'A'')B'} = d^{C}_{(A'A'')B'} - d^{S}_{(A'A'')B'} - (\tilde{\delta}_{A'B'} + \tilde{\delta}_{A''B'})$

and four-body term, from which all possible lower-body terms must be subtracted



Figure 2: Basis set convergence of $\tilde{\delta}^2$ in comparison with the overall δ evaluated in cc-pVDZ "DZ" and cc-pVTZ "TZ" basis, extrapolated to CBS cc-pVXZ(D,T) "XZ".



 $\tilde{\delta}_{(A'A'')(B'B'')} = d^C_{(A'A'')(B'B'')} - d^S_{(A'A'')(B'B'')} - (\tilde{\delta}^2 + \tilde{\delta}^3)$

References

[1] M. Dubecký: Bias cancellation in one-determinant fixed-node diffusion monte carlo: Insights from fermionic occupation numbers, *Phys. Rev. E*, 95, 033308 (2017).
[2] M. Šulka, M. Dubecký: Fragmentation of natural orbital occupation numbers-based diag-

nostic of differential multireference character in complexes with hydrogen bonds. *J Comput Chem.*, **42**, 475–483 (2021).

Acknowledgement

We acknowledge financial support from Czech Science Foundation (No. 18-24321Y), University of Ostrava (IRP201826), Slovak grants APVV-15-0105 and VEGA 1/0223/19, the computational resources (projects ITMS 26230120002 and 26210120002) and funding by the ERDF -"University Scientific Park Campus MTF STU - CAMBO" ITMS: 26220220179.

Conclusions

In this study, by performing the many-body decomposition of δ in a number of non-covalent complexes, we show that

- δ can be decomposed similarly as the interaction energy, when total δ can be reconstructed by summing up the individual m-body terms
- largest contribution comes from the two-body terms; at the same time there is a fair correlation between the individual δ and IE of a given fragment
- δ^2 , that includes all two-body fragment interactions, may thus serve as the upper/lower limit estimate of the overall δ of HB/dispersion interacting complexes, including its sign. This will help in significant reduction of CPU cost and makes it possible to predict FNDMC inaccuracy in large complexes.