Artificial Neural Networks Applied as Molecular Wavefunction Solvers

Peng-Jian Yang,[*,†] Mahito Sugiyama,[*,‡,⊥] Koji Tsuda,[*,¶,§,#] and Takeshi Yanai[*,‖,†,⊥]

†*Department of Chemistry, Nagoya University, Furocho, Chikusa Ward, Nagoya, Aichi 464-8601, Japan*

‡*National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan*

¶*Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwa-no-ha, Kashiwa, Chiba 277-8561, Japan*

§*RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan*

‖*Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Furocho, Chikusa Ward, Nagoya, Aichi 464-8601, Japan*

⊥*JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*

#*Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, Ibaraki 305-0047, Japan*

E-mail: pj.yang_lester@berkeley.edu; mahito@nii.ac.jp; tsuda@k.u-tokyo.ac.jp; yanait@chem.nagoya-u.ac.jp

1 Indocyanine Green

Geometric information of indocyanine green (ICG).

The following cartesian coordinates are below in Å.

C	-4.92157956	0.00000000	-0.38353212
C	4.92157956	0.00000000	-0.38353212
C	-3.68982383	0.00000000	0.26035610
C	3.68982383	0.00000000	0.26035610
H	-3.69891585	0.00000000	1.35075428
H	3.69891585	0.00000000	1.35075428
C	-2.43392077	0.00000000	-0.35212041
C	2.43392077	0.00000000	-0.35212041
H	-2.36219457	0.00000000	-1.44140401
H	2.36219457	0.00000000	-1.44140401
C	-1.23445439	0.00000000	0.34845573
C	1.23445439	0.00000000	0.34845573
H	-1.25820279	0.00000000	1.44143873
H	1.25820279	0.00000000	1.44143873
C	0.00000000	0.00000000	-0.29583746
H	0.00000000	0.00000000	-1.39062126
N	6.07815258	0.00000000	0.30616006

N	-6.07815258	0.00000000	0.30616006
C	-5.23795288	0.00000000	-1.88074059
C	5.23795288	0.00000000	-1.88074059
C	-6.75866976	0.00000000	-1.86628572
C	6.75866976	0.00000000	-1.86628572
C	-7.18920017	0.00000000	-0.56005095
C	7.18920017	0.00000000	-0.56005095
C	-4.68582825	1.27180284	-2.54553645
C	4.68582825	1.27180284	-2.54553645
C	-4.68582825	-1.27180284	-2.54553645
C	4.68582825	-1.27180284	-2.54553645
H	-3.59206051	1.30800374	-2.49395633
H	3.59206051	1.30800374	-2.49395633
H	-3.59206051	-1.30800374	-2.49395633
H	3.59206051	-1.30800374	-2.49395633
H	-4.97642624	1.30680353	-3.60190181
H	4.97642624	1.30680353	-3.60190181
H	-4.97642624	-1.30680353	-3.60190181
H	4.97642624	-1.30680353	-3.60190181
H	-5.07555240	2.17255029	-2.05612490
H	5.07555240	2.17255029	-2.05612490
H	-5.07555240	-2.17255029	-2.05612490
H	5.07555240	-2.17255029	-2.05612490
C	-6.16433317	0.00000000	1.75433713
C	6.16433317	0.00000000	1.75433713
H	-5.67643344	-0.89249247	2.16544172
H	5.67643344	-0.89249247	2.16544172
H	-5.67643344	0.89249247	2.16544172
H	5.67643344	0.89249247	2.16544172
H	-7.21354079	0.00000000	2.05505060
H	7.21354079	0.00000000	2.05505060
C	-7.70415822	0.00000000	-2.92443976
C	7.70415822	0.00000000	-2.92443976
C	-8.54885597	0.00000000	-0.19358428
C	8.54885597	0.00000000	-0.19358428
C	-9.47478127	0.00000000	-1.20340429
C	9.47478127	0.00000000	-1.20340429
C	-9.08978513	0.00000000	-2.57009467
C	9.08978513	0.00000000	-2.57009467
C	-7.36692541	0.00000000	-4.30746846
C	7.36692541	0.00000000	-4.30746846
C	-10.07179899	0.00000000	-3.59440088
C	10.07179899	0.00000000	-3.59440088
H	-11.12407245	0.00000000	-3.30578536
H	11.12407245	0.00000000	-3.30578536
C	-8.34329908	0.00000000	-5.27097731
C	8.34329908	0.00000000	-5.27097731
H	-8.06094841	0.00000000	-6.32388059
H	8.06094841	0.00000000	-6.32388059
C	-9.71203297	0.00000000	-4.91642240
C	9.71203297	0.00000000	-4.91642240
H	-10.47414568	0.00000000	-5.69545938
H	10.47414568	0.00000000	-5.69545938
H	-6.32269021	0.00000000	-4.61207778
H	6.32269021	0.00000000	-4.61207778
H	-10.53880823	0.00000000	-0.96448775
H	10.53880823	0.00000000	-0.96448775
H	-8.86327452	0.00000000	0.84809704
H	8.86327452	0.00000000	0.84809704

Total energies of ICG for CAS(4,4) calculated with the deterministic approach.

Model	Hidden nodes	Total energies (in E_h) Seed=1000	Seed=2000	Speed=3000
CMO				
RBM	1	-1512.87947	-1512.87937	-1512.87947
RBM	2	-1512.87977	-1512.87962	-1512.87977
RBM	3	-1512.88011	-1512.88001	-1512.88040
RBM	4	-1512.88041	-1512.88042	-1512.88036
RBM	5	-1512.88042	-1512.88042	-1512.88042
RBM	6	-1512.88042	-1512.88042	-1512.88042
RBM	7	-1512.88042	-1512.88042	-1512.88042
RBM	8	-1512.88042	-1512.88042	-1512.88042
RBM	9	-1512.88042	-1512.88042	-1512.88042
RBM	10	-1512.88042	-1512.88042	-1512.88042
RBM	20	-1512.88042	-1512.88042	-1512.88042
RBM	40	-1512.88042	-1512.88042	-1512.88042
BM2	0	-1512.87983	-1512.87983	-1512.87983
BM3	0	-1512.88041	-1512.88041	-1512.88041
LMO				
RBM	1	-1512.87351	-1512.87360	-1512.87518
RBM	2	-1512.87831	-1512.87755	-1512.87845
RBM	3	-1512.87993	-1512.88023	-1512.88023
RBM	4	-1512.88038	-1512.88016	-1512.88039
RBM	5	-1512.88042	-1512.88036	-1512.88041
RBM	6	-1512.88042	-1512.88042	-1512.88041
RBM	7	-1512.88042	-1512.88042	-1512.88042
RBM	8	-1512.88042	-1512.88042	-1512.88042
RBM	9	-1512.88042	-1512.88042	-1512.88042
RBM	10	-1512.88042	-1512.88042	-1512.88042
RBM	20	-1512.88042	-1512.88042	-1512.88042
RBM	40	-1512.88042	-1512.88042	-1512.88042
BM2	0	-1512.87937	-1512.87937	-1512.87937
BM3	0	-1512.88010	-1512.88010	-1512.88010
CASCI		-1512.88042		
HF		-1512.87513		

Total energies of ICG for CAS(4,4) calculated with the stochastic approach.

Model	Hidden nodes	MO type	Seed	E (in E_h)	Standard deviation
RBM	5	CMO	1000	-1512.88041	4.6E-07
			2000	-1512.88037	1.3E-06
			3000	-1512.88025	9.8E-09
	5	LMO	1000	-1512.88040	4.7E-07
			2000	-1512.88040	4.8E-07
			3000	-1512.88004	2.5E-06
	10	CMO	1000	-1512.88042	3.9E-08
			2000	-1512.87794	5.2E-09
			3000	-1512.88033	5.4E-09
	10	LMO	1000	-1512.88042	1.3E-07
			2000	-1512.88042	8.0E-08
			3000	-1512.88042	2.3E-09

Canonical and localized forms of active orbitals of ICG used for CAS(4e,4o)

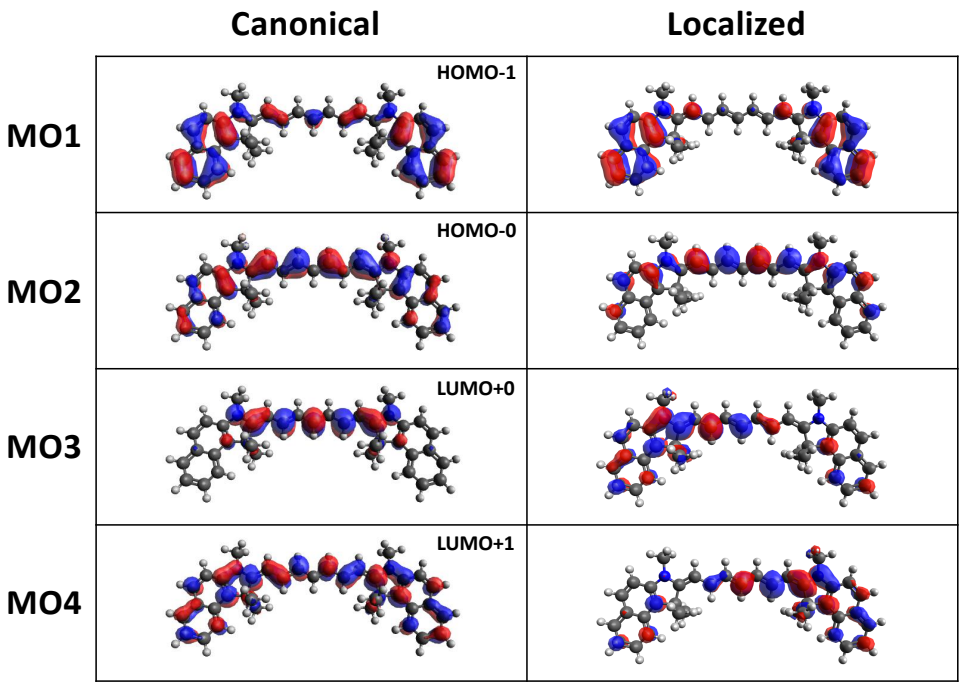

Figure 1: Plots of canonical and localized forms of active orbitals of ICG used for CAS(4e,4o). Pipek-Mezey method were used for orbital localization.

Total energies of ICG for CAS(6,6) calculated with the deterministic approach.

Model	Hidden nodes	Total energies (in E_h) Seed=1000	Seed=2000	Speed=3000
CMO				
RBM	5	-1512.88500	-1512.88583	-1512.88568
RBM	10	-1512.88664	-1512.88675	-1512.88673
RBM	20	-1512.88681	-1512.88681	-1512.88681
RBM	40	-1512.88681	-1512.88681	-1512.88681
BM2	0	-1512.88551	-1512.88567	-1512.88560
BM3	0	-1512.88665	-1512.88663	-1512.88666
LMO				
RBM	5	-1512.87414	-1512.86314	-1512.87199
RBM	10	-1512.88473	-1512.88521	-1512.88478
RBM	20	-1512.88384	-1512.88306	-1512.88645
RBM	40	-1512.88681	-1512.88681	-1512.88681
BM2	0	-1512.87325	-1512.87325	-1512.87325
BM3	0	-1512.88491[a]	N/A[b]	-1512.88492[a]
CASCI		-1512.88682		

[a] Unconverged with F_θ being approximately 8×10^{-4}. [b] Unconverged with F_θ above 8×10^{-3}.

Total energies of ICG for CAS(8,8) calculated with the deterministic approach.

Model	Hidden nodes	Total energies (in E_h) Seed=1000	Seed=2000	Speed=3000
CMO				
RBM	10	-1512.89858	-1512.89807	-1512.89854
RBM	20	-1512.89986	-1512.89974	-1512.89981
RBM	30	-1512.90018	-1512.90023	-1512.90008
RBM	40	-1512.90026	-1512.90028	-1512.90028
RBM	50	-1512.90034	-1512.90031	-1512.90035
RBM	60	-1512.90035	-1512.90035	-1512.90030
RBM	70	-1512.90031	-1512.90031	-1512.90034
RBM	80	-1512.90032	-1512.90030	-1512.90031
RBM	90	-1512.90036	-1512.90031	-1512.90033
RBM	100	-1512.90026	-1512.90030	-1512.90038
CASCI		-1512.90055		

2 Dinitrogen

CAS(6e,6o) MOs of the dinitrogen ground state $X^1\sum_g^+$

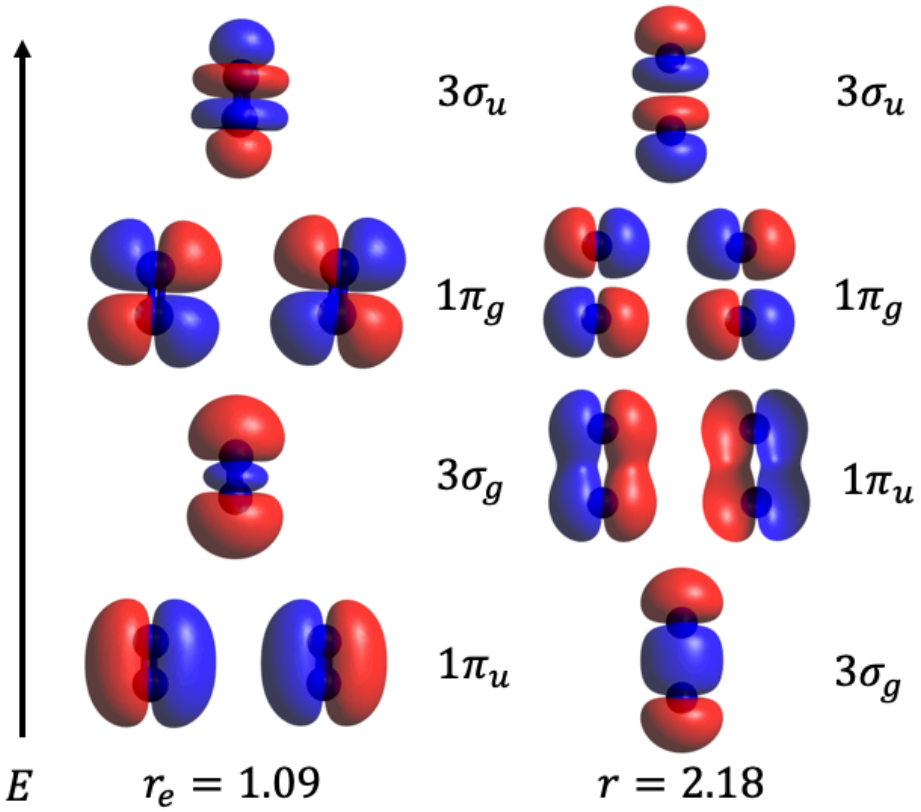

Figure 2: The active orbitals used for CAS-CI calculations of the N_2 molecule of the $X^1\sum_g^+$ state; $(1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 1\pi_u^4 2\sigma_g^2)$; change in orbital energy and shape is shown for increased bond length. Units given in Å.

Total energies of dinitrogen for CAS(6,6) in CMO basis.

	Energy/ E_h			ΔE / E_h		
r / Å	CASCI	RHF	RBM (MCMC)	RBM (det)	BM2 (det)	BM3 (det)
.545	-102.33974	0.03173	< 1e-6	< 1e-6	4e-6	< 1e-6
1.09	-107.61734	0.12381	1.1e-4	< 1e-6	1.3e-4	1e-6
1.64	-107.50162	0.34717	2.0e-4	< 1e-6	1.9e-4	< 1e-6
2.18	-107.43267	0.67003	1.9e-4	3.9e-4	2.7e-3	5e-6
2.73	-107.43579	0.89174	3.8e-4	1.9e-4	3.0e-4	4.6e-5
3.27	-107.43720	1.00077	8.8e-5	5.7e-5	8.8e-5	1.6e-5

Total energies of dinitrogen for CAS(6,6) in CMO basis with RBM (det).

bond length / Å	E / E_h	ΔE from CASCI
.545	-102.33974	3.24e-07
.763	-106.44497	1.34e-06
.872	-107.15464	2.02e-08
.981	-107.48673	1.47e-08
1.09	-107.61734	3.47e-11
1.20	-107.64505	6.32e-06
1.31	-107.62432	2.71e-07
1.42	-107.58443	3.64e-08
1.53	-107.54080	3.52e-08
1.64	-107.50162	1.01e-08
1.96	-107.43935	4.63e-05
2.07	-107.43392	2.98e-04
2.18	-107.43228	3.87e-04
2.29	-107.43171	1.11e-03
2.40	-107.43288	7.21e-04
2.73	-107.43560	1.88e-04
3.27	-107.43726	5.41e-05