2D-block Geminals: Taming the combinatorics while releasing the strong orthogonality constraint

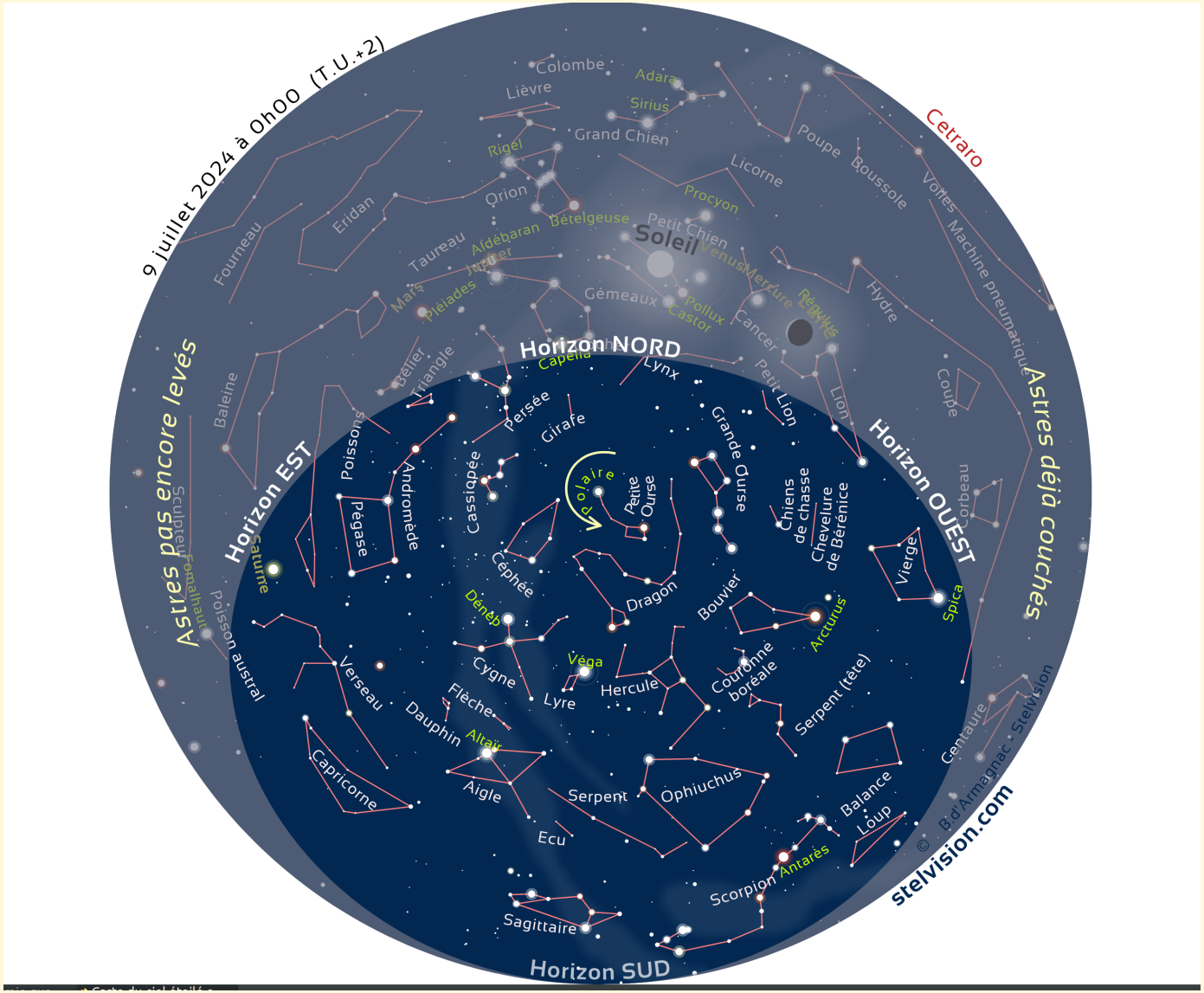
Patrick CASSAM-CHENAÏ

Laboratoire J.A. Dieudonné, Nice, France

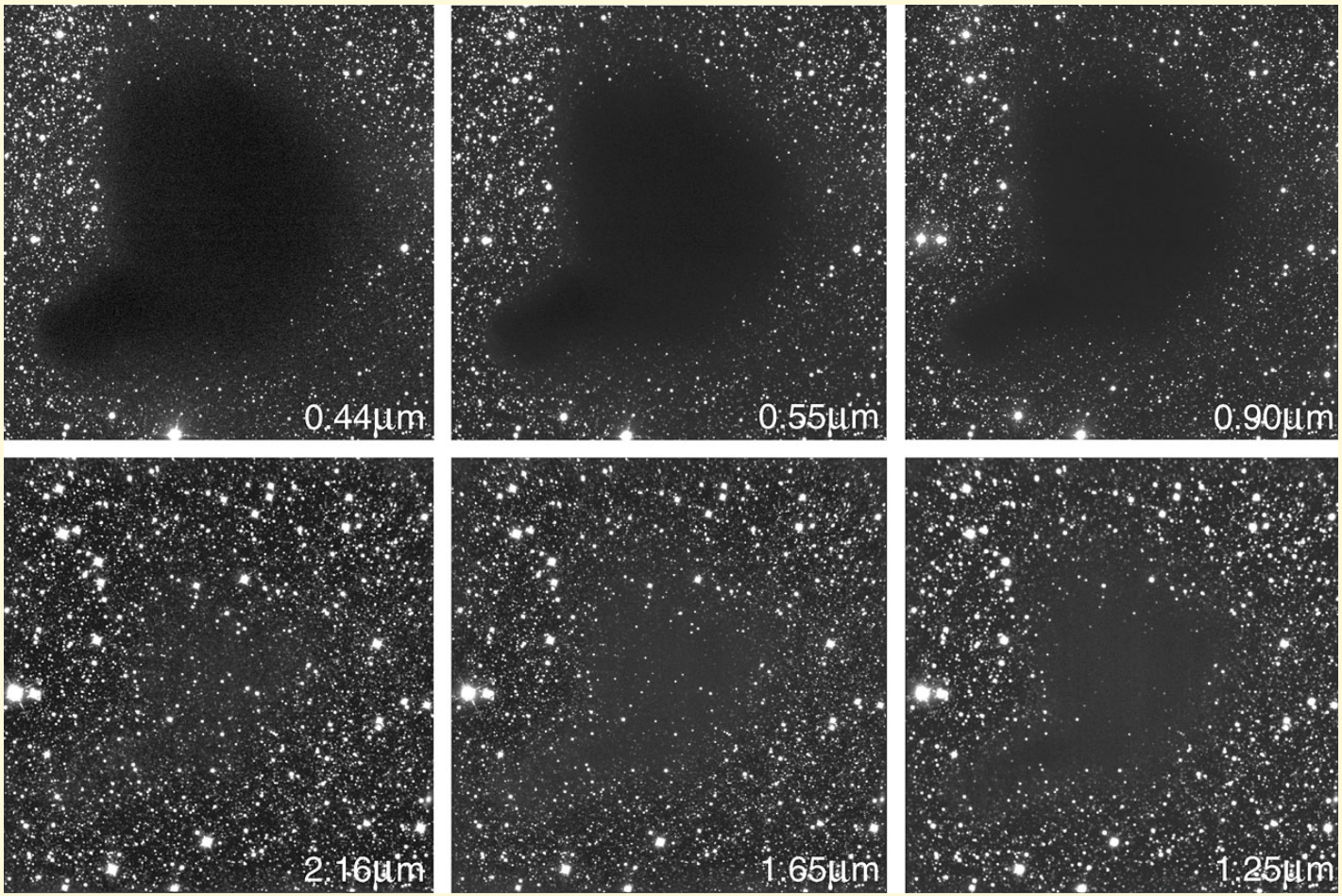
Cetraro 08/06/2024

Plan

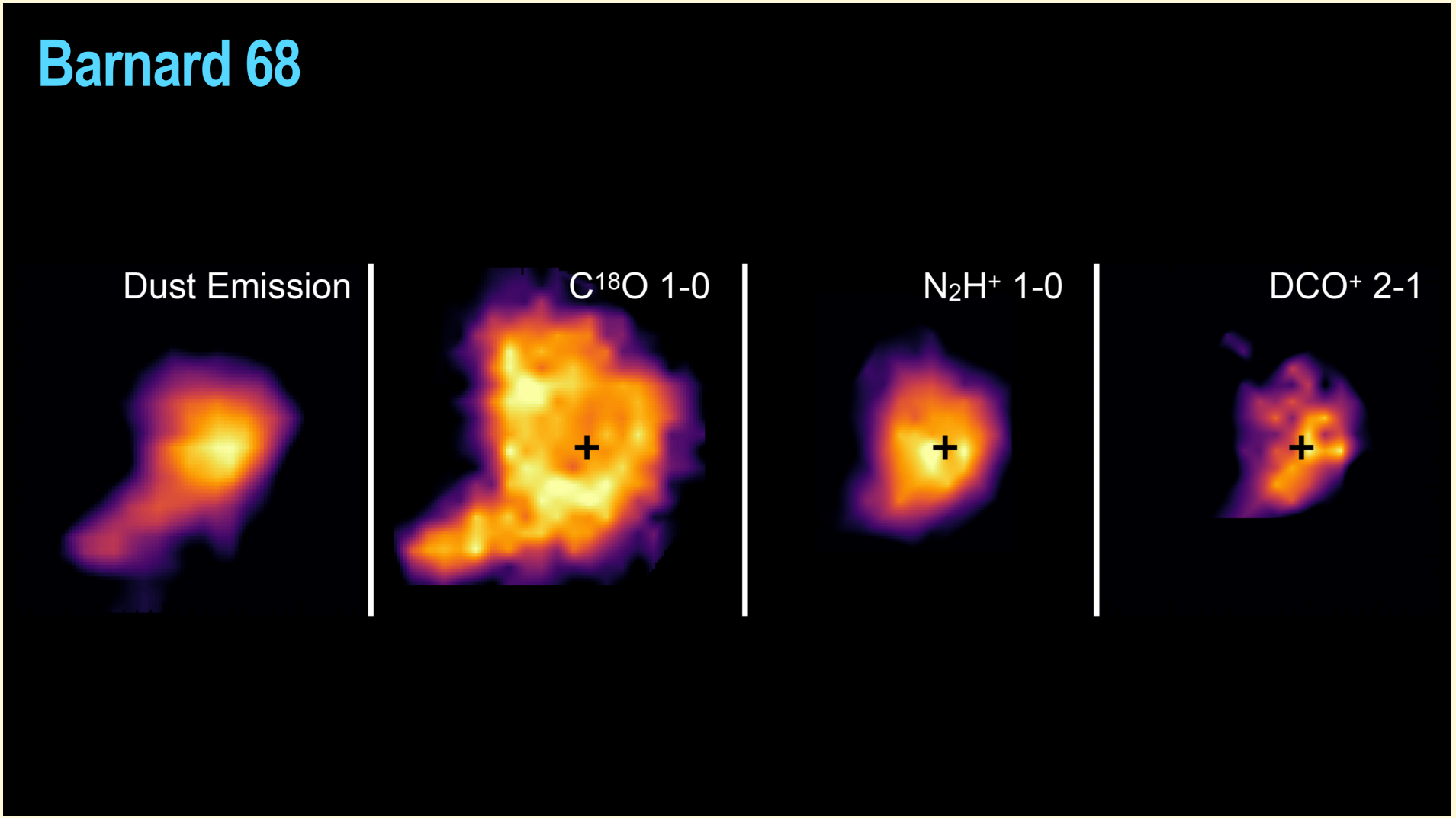
Sky map


9 juillet 2024 à 0h00 (T.U.+2)
Cetraro
Horizon NORD
Horizon EST
Horizon OUEST
Horizon SUD
Astres pas encore levés
Astres déjà couchés
Soleil
Colombe
Lièvre
Adara
Sirius
Grand Chien
Rigel
Orion
Bételgeuse
Procyon
Licorne
Poupe
Boussole
Voiles
Machine pneumatique
Eridan
Fourneau
Taureau
Aldébaran
Jupiter
Mars
Pléiades
Gémeaux
Pollux
Castor
Petit Chien
Vénus
Mercure
Régulus
Cancer
Hydre
Coupe
Corbeau
Lion
Petit Lion
Lynx
Capella
Persée
Girafe
Cassiopée
Polaire
Petite Ourse
Grande Ourse
Chiens de chasse
Chevelure de Bérénice
Vierge
Spica
Arcturus
Bouvier
Dragon
Céphée
Dénébola
Véga
Cygne
Lyre
Hercule
Couronne boréale
Serpent (tête)
Serpent
Ophiuchus
Balance
Loup
Centaure
Scorpion
Antarès
Sagittaire
Ecu
Aigle
Flèche
Dauphin
Altaïr
Capricorne
Verseau
Poisson austral
Fomalhaut
Sculpteur
Saturne
Baleine
Triangle
Bélier
Poissons
Andromède
Pégase
© B.d'Armagnac - Stelvision
stelvision.com

Barnard68



In emission

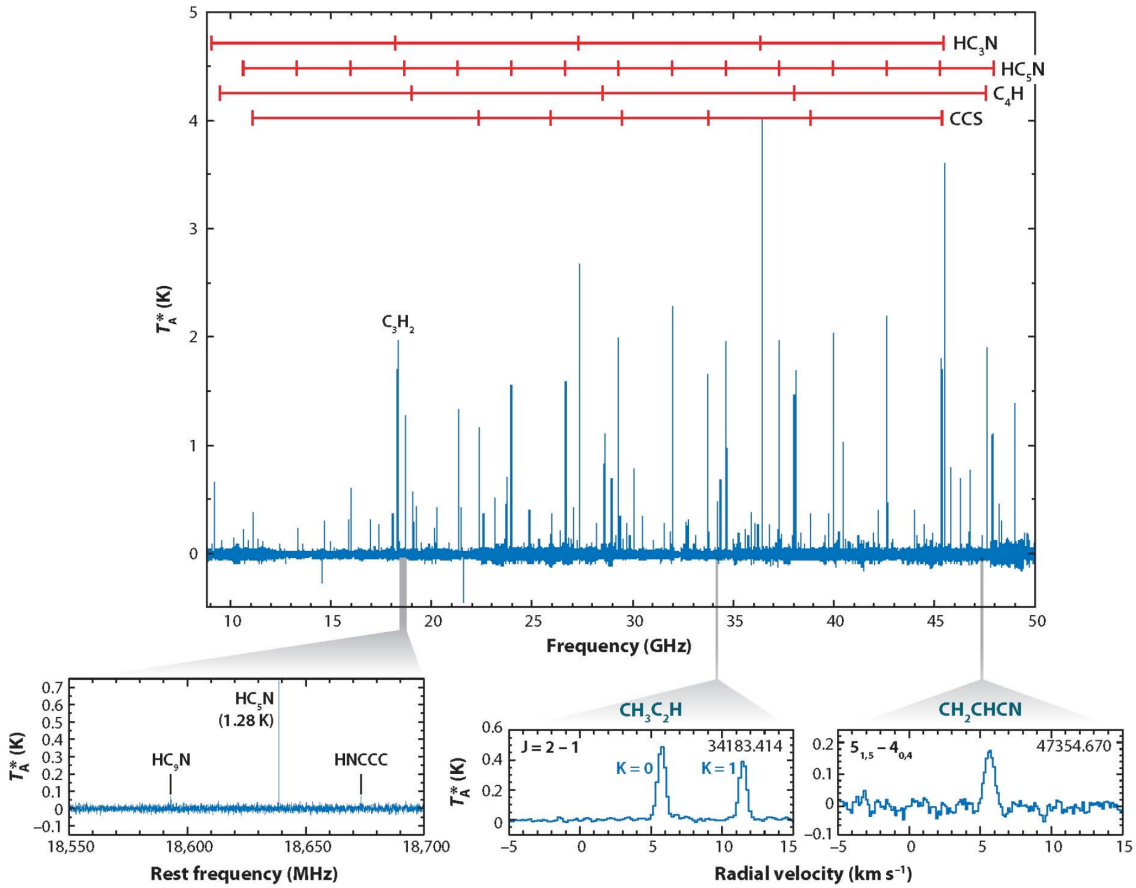


Identified inter/circumstellar molecules 06/2024

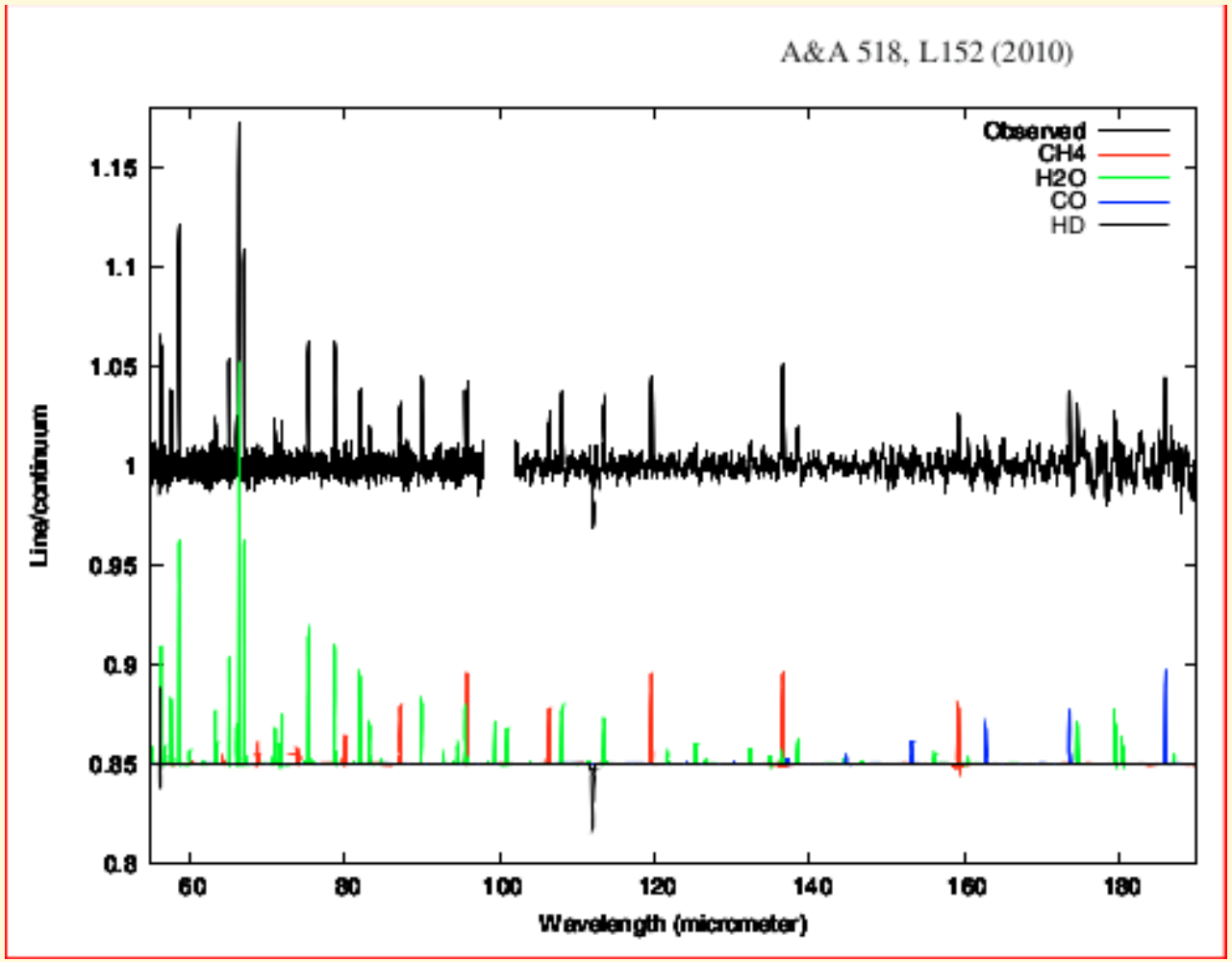
2 atoms	3 atoms	4 atoms	5 atoms	6 atoms	7 atoms	8 atoms	9 atoms	10 atoms	11 atoms	12 atoms	>12 atoms
H_2	C_3^*	c-C_3H	C_5^*	C_5H	C_6H	CH_3C_3N	CH_3C_4H	CH_3C_5N	HC_9N	c-$C_6H_6^*$	C_{60}^*
AlF	C_2H	l-C_3H	C_4H	l-H_2C_4	CH_2CHCN	$HC(O)OCH_3$	CH_3CH_2CN	$(CH_3)_2CO$	CH_3C_8H	n-C_3H_7CN	C_{70}^*
AlCl	C_2O	C_3N	C_4Si	$C_2H_4^*$	CH_3C_2H	CH_3COOH	$(CH_3)_2O$	$(CH_2OH)_2$	C_2H_5OCHO	i-C_3H_7CN	C_{60}^{+*}
C_2^{**}	C_2S	C_3O	l-C_3H_2	CH_3CN	HC_5N	C_7H	CH_3CH_2OH	CH_3CH_2CHO	$CH_3OC(O)CH_3$	$C_2H_5OCH_3$	c-C_6H_5CN
CH	CH_2	C_3S	c-C_3H_2	CH_3NC	CH_3CHO	C_8H_2	HC_7N	CH_3CHCH_2O	$CH_3C(O)CH_2OH$	1-c-C_5H_5CN	$HC_{11}N$
CH^+	HCN	$C_2H_2^*$	H_2CCN	CH_3OH	CH_3NH_2	CH_2OHCHO	C_8H	CH_3OCH_2OH	c-C_8H_6	2-c-C_5H_5CN	1-$C_{10}H_7CN$
CN	HCO	NH_3	CH_4^*	CH_3SH	c-C_2H_4O	l-HC_6H^+	$CH_3C(O)NH_2$	c-C_8H_4	$HOCH_2CH_2NH_2$	CH_3C_7N (?)	2-$C_{10}H_7CN$
CO	HCO^+	HCCN	HC_3N	HC_3NH^+	H_2CCHOH	CH_2CHCHO	C_8H^-	H_2CCCHC_3N	H_2CCCHC_4H	n-C_3H_7OH	c-C_9H_8
CO^+	HCS^+	$HCNH^+$	HCCNC	HCCCHO	C_6H^-	CH_2CCHCN	C_3H_6	C_2H_5NCO	$C_{10}H^-$ (2023)	i-C_3H_7OH	1-c-C_5H_5CCH
CP	HOC^+	HNCO	HCOOH	NH_2CHO	CH_3NCO	H_2NCH_2CN	CH_3CH_2SH	$C_2H_5NH_2$ (?)	$H_2C(CH_3)_3CN$ (2023)	$(CH_3)_2C{=}CH_2$ (2023)	2-c-C_5H_5CCH
SiC	H_2O	HNCS	H_2CNH	C_5N	HC_5O	CH_3CHNH	CH_3NHCHO	HC_7NH^+			c-$C_5H_4CCH_2$
HCl	H_2S	$HOCO^+$	H_2C_2O	l-HC_4H^+	$HOCH_2CN$	CH_3SiH_3	HC_7O	E-$CH_3CHCHCN$			2-C_9H_7CN
KCl	HNC	H_2CO	H_2NCN	l-HC_4N	HCCCHNH	$H_2NC(O)NH_2$	HCCCHCHCN	Z-$CH_3CHCHCN$			C_6CH_3CCH (2023)
NH	HNO	H_2CN	HNC_3	c-H_2C_3O	HC_4NC	HCCCH$_2$CN	H_2CCHC_3N	$CH_3C(CN)CH_2$			$CH_3OCH_2CH_2OH$ (2024)
NO	MgCN	H_2CS	SiH_4^*	H_2CCNH	c-C_3HCCH	HC_6NH^+	$H_2CCCHCCH$	CH_2CHCH_2CN			
NS	MgNC	H_3O^+	H_2COH^+	C_5N^-	l-H_2C_5	CH_2CHCCH	HOCHCHCHO (?)	$HOCH_2C(O)NH_2$ (2023)			
NaCl	N_2H^+	c-SiC_3	C_4H^-	HNCHCN	MgC_3N	MgC_6H		CH_3CH_2CCH (2024)			
OH	N_2O	CH_3^*	HC(O)CN	SiH_3CN	CH_2C_3N	$C_2H_5NH_2$					
PN	NaCN	C_3N^-	HNCNH	C_5S	NC_4NH^+ (2023)	$(CHOH)_2$					
SO	OCS	PH_3	CH_3O	MgC_4H	MgC_3N^+ (2023)	$HC_2(H)C_4$					
SO^+	SO_2	HCNO	NH_4^+	CH_3CO^+		C_7N^- (2023)					
SiN	c-SiC_2	HOCN	H_2NCO^+	C_3H_3		CH_3CHCO (2023)					
SiO	CO_2^*	HSCN	NCCNH$^+$	H_2C_3S		MgC_6H^+ (2023)					
SiS	NH_2	H_2O_2	CH_3Cl	HCCCHS							
CS	$H_3^{+(*)}$	C_3H^+	MgC_3N	C_5O							
HF	SiCN	HMgNC	NH_2OH	C_5H^+							
HD	AlNC	HCCO	HC_3O^+	HCCNCH$^+$							
FeO?	SiNC	CNCN	HC_3S^+	c-C_3C_2H							
O_2	HCP	HONO	H_2C_2S	HC_4S							
CF^+	CCP	MgC_2H	C_4S	$HMgC_3N$ (2023)							
SiH?	AlOH	HCCS	HC(O)SH	MgC_4H^+ (2023)							
PO	H_2O^+	HNCN	HC(S)CN	$H_2C_3H^+$ (2023)							
AlO	H_2Cl^+	H_2NC	HCCCO	H_2C_3N (2023)							
OH^+	KCN	$HCCS^+$	NaCCCN (2023)	$(HO)_2CO$ (2023)							
CN^-	FeCN	CH_3^+ (2023)	MgC_3N^+ (2023)	H_2CNCN (2024)							
SH^+	HO_2	HCNS (2024)		NCHCCS (2024)							
SH	TiO_2	$HOCS^+$ (2024)									
HCl^+	C_2N	HNSO (2024)									
TiO	Si_2C										
ArH^+	HS_2										
N_2	HCS										
NO^+?	HSC										
NS^+	NCO										
HeH^+	CaNC										
PO^+	NCS										
SiP ?	MgC_2										
FeC (2023)	HSO (2023)										
	CaC_2 (2024)										

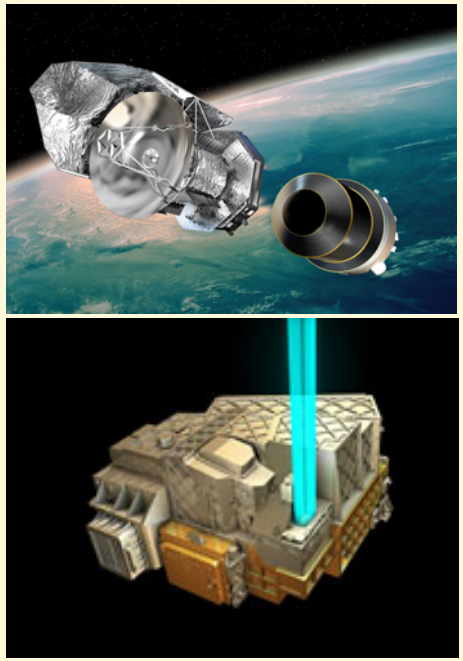
Unidentified molecules

TMC-1: scan from 10-50 GHz



Neptune's atmosphere abundancies





(*Herschel-PACS, 2009/10/30, Lellouch et al.*)

Schrödinger equation

Erwin Rudolf Josef Alexander Schrödinger
(1887 – 1961)

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \hat{H}\Psi(\mathbf{r},t)$$

Stationary Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \hat{H}\Psi(\mathbf{r},t) \xrightarrow{\Psi(\mathbf{r},t)=e^{-i\omega t}\psi(\mathbf{r})} \hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

où :

- $\hat{H}$: Hamiltonian: operator defined on a Hilbert space
- $\psi(\mathbf{r})$: wave function: vector of the Hilbert space $H = L^2(\Omega \subset \mathbb{R}^n)$
- E : Eigenvalue of the Hamiltonian

Solutions of the Stationary Schrödinger equation

- A given molecule (ex.: CO, CH_4) ⇝ a well-defined $\hat{H}$ (ex.: $\hat{H}_{CO}$, $\hat{H}_{CH_4}$)
- We solve the eigenvalue equation: $\hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r})$:

$$E_0 < E_1 < E_2 < \cdots < E_i < \cdots$$

$$\psi_0, \quad \psi_1, \quad \psi_2, \quad \cdots, \quad \psi_i, \quad \cdots$$

- Line positions ⇝ Possible wavelengths of emitted or absorbed photons

$$\lambda_{i\to j} = \frac{hc}{|E_i - E_j|}$$

- Transition intensity given by ψ_i, ψ_j.

- Electrons+nuclei eigenvalue problem $\rightsquigarrow$ Electrons only eigenvalue problem (+ nuclear motion only eigenvalue problem)

$$H_e[(Q_i)_i]\ \Psi = V_{BO}[(Q_i)_i]\ \Psi$$

- $H_e[(Q_i)_i]$ Hamiltonian for the electrons (fixed nuclei at positions $(Q_i)_i$)
- $\Psi \in \wedge^n \mathcal{H}$, n-electron wave function,
- $\mathcal{H}$ Hilbert space of the form: $\mathcal{H} := \mathcal{H}_o \otimes \mathcal{H}_s$,
$\mathcal{H}_s \equiv \mathbb{C}^2$ and $\mathcal{H}_o$ a finite dimensional vector subspace of $L^2(\mathbb{R}^3)$,
called the space of "orbitals", $dim\mathcal{H}_o = m$.

Mathematical references

* G. A. Hagedorn, A. Joye, *Comm. Math. Phys.* **223**, *583*, 2001.
"A Time-dependent Born-Oppenheimer Approximation with exponentially small error estimates"

* T. Jecko, *J. Math. Phys.* **55**, *053504*, 2014. *"On the mathematical treatment of the Born-Oppenheimer approximation"*

- $\mathcal{H}$: Hilbert space of one-electron functions , spanned by a set of "spin-orbitals" $(\psi_i)_i$, $dim\mathcal{H} = 2m$.
- $\wedge^n\mathcal{H}$: Hilbert space of n-electron functions , spanned by the "single configurations" $(\psi_{i_1} \wedge \psi_{i_2} \wedge \cdots \wedge \psi_{i_n})_{(i_1<i_2<\cdots<i_n)}$, $dim \wedge^n \mathcal{H} = \binom{2m}{n}$.

As required by Pauli's spin-statistic theorem:

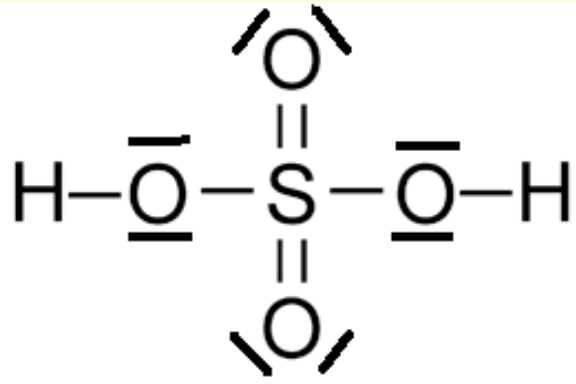
$$\forall \sigma \in \mathcal{S}_n \quad \psi_{i_1} \wedge \psi_{i_2} \wedge \cdots \wedge \psi_{i_n} = (-1)^{|\sigma|} \psi_{\sigma(i_1)} \wedge \cdots \wedge \psi_{\sigma(i_n)}$$

- $\mathcal{H} = \mathcal{H}_o \otimes \mathbb{C}^2$ is usually regarded as a sum of 2 orthogonal subspaces:

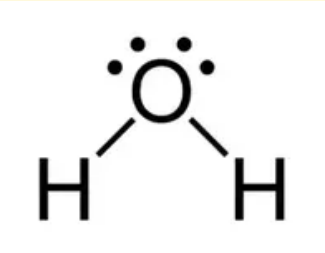
$$\mathcal{H} \equiv \mathcal{H}_\alpha \oplus \mathcal{H}_\beta, \qquad dim\mathcal{H}_\alpha = dim\mathcal{H}_\beta = m,$$

* $\mathcal{H}_\alpha$: Hilbert space of $\alpha \equiv (\frac{1}{2})$-spin-orbitals , spanned by $(\psi_i)_i$
* $\mathcal{H}_\beta$: Hilbert space of $\beta \equiv (-\frac{1}{2})$-spin-orbitals , spanned by $(\bar{\psi}_i)_i$

Lewis pairs of "classical chemistry"



Sulfuric acid



Water

APG wave function

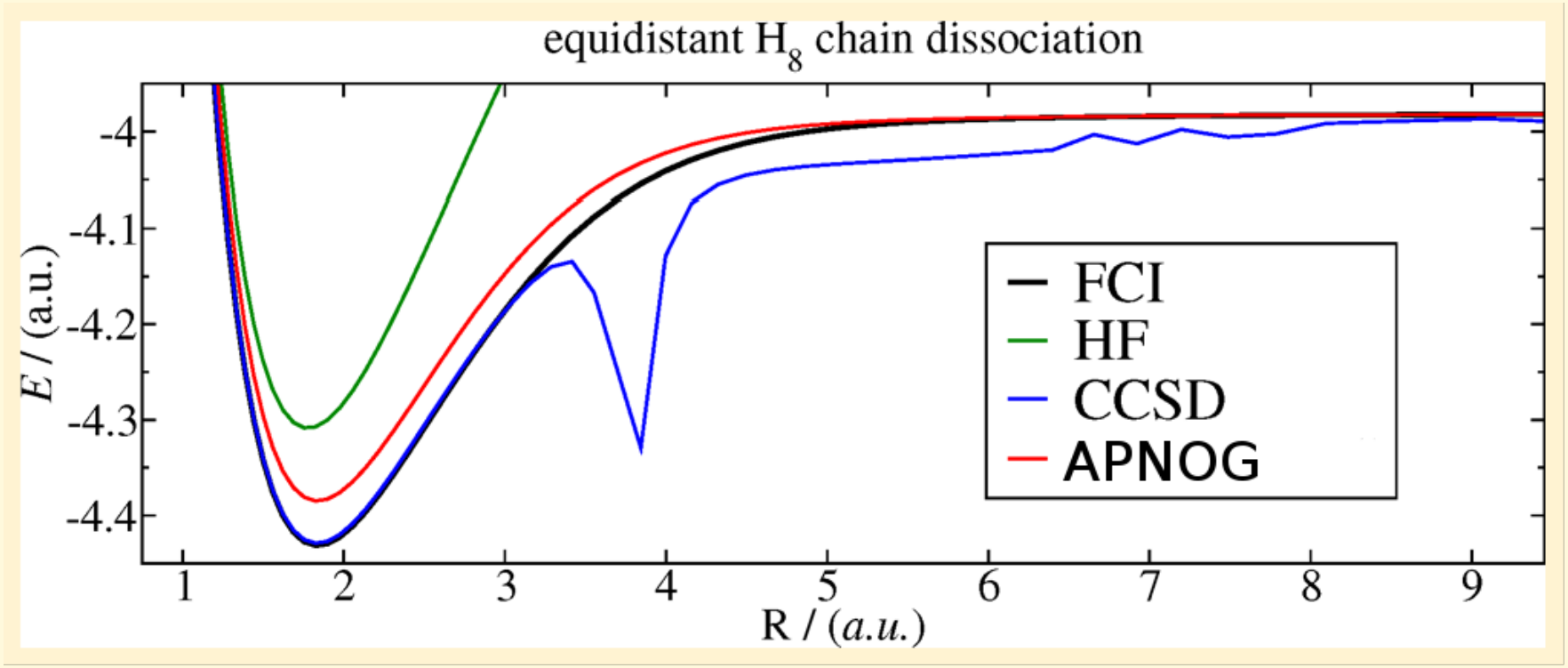
- geminal: $\Gamma \in \wedge^2 \mathcal{H}$

$$\Gamma = \sum_{i_1<i_2} c_{i_1 i_2}\ \psi_{i_1} \wedge \psi_{i_2}$$

- antisymmetrized product of geminals wave function: $\Phi_{APG} \in \wedge^{2k} \mathcal{H}$

$$\Phi_{APG} = \Gamma_1 \wedge \Gamma_2 \wedge \cdots \wedge \Gamma_k\ , \qquad \Gamma_1, \Gamma_2, \cdots, \Gamma_k \in \wedge^2 \mathcal{H}$$

Strongly correlated electronic systems



Limacher, Ayers, Johnson, De Baerdemacker, Van Neck, Bultinck, JCTC **9** *(2013) p.1394*

FCI: Full Configuration Interaction
HF: Hartree-Fock method
CCSD: Coupled-Cluster Single and Double
APNOG $\rightsquigarrow$ AP1roG : antisymmetrized Product of 1 reference orbital Geminals $\equiv$ pCCD

Antisymmetric Product of set-divided Geminals (APsetG)

$\rightsquigarrow$ General geminal: , $\psi_{i_1}, \psi_{i_2} \in \mathcal{H}$

$$\Gamma = \sum_{1 \le i_1 < i_2 \le 2m} c_{i_1 i_2} \ \psi_{i_1} \wedge \psi_{i_2}$$

Each geminal is parametrized by $m(2m-1)$ independent scalars c_{i_1,i_2}

$\rightsquigarrow$ APsetG geminal: , $\psi_{i_1} \in \mathcal{H}_\alpha, \bar{\psi}_{i_2} \in \mathcal{H}_\beta$ $\hat{S}_z$ eigenfunctions

$$\Gamma_{\text{set}} = \sum_{1 \le i_1, i_2 \le 2m} c_{i_1 i_2} \ \psi_{i_1} \wedge \bar{\psi}_{i_2} \qquad S_z = 0$$

Each APsetG geminal is parametrized by m^2 independent scalars c_{i_1,i_2}

$\rightsquigarrow$ Γ_{set} is parametrized by an $m \times m$ matrix: $C = (c_{i_1,i_2})$

Antisymmetric Product of set-divided Geminals (APsetG)

$\rightsquigarrow$ $\hat{S}^2$-Spin-adapted geminals

Singlet $\Rightarrow$ $c_{i_1 i_2} = c_{i_2 i_1}$ i.e. C symmetric

$$\Gamma_{\text{singlet}} = \sum_{1 \leq i_1 \leq m} c_{i_1 i_1} \; \psi_{i_1} \wedge \bar{\psi}_{i_1} + \sum_{1 \leq i_1 < i_2 \leq m} c_{i_1 i_2} \; \left(\psi_{i_1} \wedge \bar{\psi}_{i_2} + \psi_{i_2} \wedge \bar{\psi}_{i_1} \right)$$

Triplet $\Rightarrow$ $c_{i_1 i_2} = -c_{i_2 i_1}$ i.e. C antisymmetric

$$\Gamma_{\text{triplet}} = \sum_{1 \leq i_1 < i_2 \leq m} c_{i_1 i_2} \; \left(\psi_{i_1} \wedge \bar{\psi}_{i_2} - \psi_{i_2} \wedge \bar{\psi}_{i_1} \right)$$

$\rightsquigarrow$ Hermitian product = "overlap" formula: $\Gamma \leftrightarrow C$, $\Gamma' \leftrightarrow C'$

$$\langle \Gamma | \Gamma' \rangle = tr[C^\dagger C']$$

APsetG overlap formula

* General overlap closed formula: $\Gamma_i \leftrightarrow C_i$, $\Gamma'_j \leftrightarrow C'_j$

$$\langle \Gamma_1 \wedge \cdots \wedge \Gamma_k | \Gamma'_1 \wedge \cdots \wedge \Gamma'_k \rangle = \langle 0 | a(\Gamma_k) \cdots a(\Gamma_2) a(\Gamma_1) a^\dagger(\Gamma'_1) a^\dagger(\Gamma'_2) \cdots a^\dagger(\Gamma'_k) | 0 \rangle =$$

$$\sum_{\substack{0 \le N_{k,0}, \ldots, N_{k,k} \le k \\ \sum_{i=0}^{k} N_{k,i} = \sum_{i=0}^{k} i N_{k,i} = k}} (-1)^{N_{k,0}} \sum_{\sigma, \sigma' \in \mathfrak{S}_k} \prod_{i=1}^{k} \frac{\prod_{j=1}^{N_{k,i}} tr\left[C^\dagger_{\sigma\left(\sum_{p=0}^{i-1} p N_{k,p} + (j-1)i+1\right)} C'_{\sigma'\left(\sum_{p=0}^{i-1} p N_{k,p} + (j-1)i+1\right)} \cdots C^\dagger_{\sigma\left(\sum_{p=0}^{i-1} p N_{k,p} + ji\right)} C'_{\sigma'\left(\sum_{p=0}^{i-1} p N_{k,p} + ji\right)} \right]}{i^{N_{k,i}} \, N_{k,i}!},$$

* Ex.: $k = 3 \rightsquigarrow$ 3 partitions of k

- $N_{3,1} = 3,\ N_{3,2} = 0,\ N_{3,3} = 0 \quad \rightsquigarrow \quad 3 * 1 + 0 * 2 + 0 * 3 = 3 = 1 + 1 + 1, \quad N_{3,0} = 0$

$$\sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{tr\left[C^\dagger_{\sigma(1)} C'_{\sigma'(1)}\right] \, tr\left[C^\dagger_{\sigma(2)} C'_{\sigma'(2)}\right] \, tr\left[C^\dagger_{\sigma(3)} C'_{\sigma'(3)}\right]}{3!}$$

- $N_{3,1} = 1,\ N_{3,2} = 1,\ N_{3,3} = 0 \quad \rightsquigarrow \quad 1 * 1 + 1 * 2 + 0 * 3 = 3 = 0 + 1 + 2, \quad N_{3,0} = 1$

$$- \sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{tr\left[C^\dagger_{\sigma(1)} C'_{\sigma'(1)}\right] \, tr\left[C^\dagger_{\sigma(2)} C'_{\sigma'(2)} C^\dagger_{\sigma(3)} C'_{\sigma'(3)}\right]}{2}$$

- $N_{3,1} = 0,\ N_{3,2} = 0,\ N_{3,3} = 1 \quad \rightsquigarrow \quad 0 * 1 + 0 * 2 + 1 * 3 = 3 = 0 + 0 + 3, \quad N_{3,0} = 2$

$$\sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{tr\left[C^\dagger_{\sigma(1)} C'_{\sigma'(1)} C^\dagger_{\sigma(2)} C'_{\sigma'(2)} C^\dagger_{\sigma(3)} C'_{\sigma'(3)}\right]}{3}$$

APsetG overlap formula

$$\langle \Phi_{APG} | \Phi'_{APG} \rangle = \langle \Gamma_1 \wedge \cdots \wedge \Gamma_k | \Gamma'_1 \wedge \cdots \wedge \Gamma'_k \rangle$$

Scaling of the general overlap formula $\rightsquigarrow$ Exponential computational cost

$$\sum_{\substack{0 \leq N_{k,0},\ldots,N_{k,k} \leq k \\ \sum_{i=0}^{k} N_{k,i} = \sum_{i=0}^{k} i N_{k,i} = k}} \quad \underset{k \to +\infty}{nb.\ of\ terms \longrightarrow} \quad \frac{1}{4k\sqrt{3}} \exp\left(\pi \sqrt{\frac{2k}{3}}\right)$$

Antisymmetrized Product of Interacting Geminals

APIG All C_i, C'_i matrices are diagonal

$$\forall i,\ C_i = \begin{pmatrix} \lambda_1^i & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_m^i \end{pmatrix},\ C'_i = \begin{pmatrix} \lambda'^i_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda'^i_m \end{pmatrix}$$

- Seniority-0 geminals $\Gamma_i = \sum_{1 \le j \le m} \lambda_j^i\ \phi_j \wedge \bar{\phi}_j$
- $tr\left[C_{a_1}^\dagger C'_{b_1} \cdots\ C_{a_i}^\dagger C'_{b_i}\right] \to \sum_{j=1}^{m} \lambda_j^{a_1} \cdots \lambda_j^{a_i} \lambda'^{b_1}_j \cdots \lambda'^{b_i}_j$
- Scaling still $\frac{1}{4k\sqrt{3}} \exp\left(\pi\sqrt{\frac{2k}{3}}\right) \rightsquigarrow$ AP1roG, Richardson-Gaudin states

Antisymmetrized Geminal Power (AGP)

aka Projected Bardeen-Cooper-Schrieffer wave function $\forall i, C_i = C_0$

- $\sum_{\sigma,\sigma' \in \mathfrak{S}_k} \to (k!)^2$
- $tr\Big[C^\dagger_{\sigma(\sum_{p=0}^{i-1} pN_{k,p}+(j-1)i+1)} C_{\sigma'(\sum_{p=0}^{i-1} pN_{k,p}+(j-1)i+1)} \cdots C^\dagger_{\sigma(\sum_{p=0}^{i-1} pN_{k,p}+ji)} C_{\sigma'(\sum_{p=0}^{i-1} pN_{k,p}+ji)}\Big]$

$$\to tr[(C_0^\dagger C_0)^i]$$

$\rightsquigarrow$ Without loss of generality: $\forall i,j, \quad (C_0)_{i,j} = \delta_{i,j}\eta_i$

- $$tr[(C_0^\dagger C_0)^i] = \sum_{j=1}^{m} (\eta_j^2)^i$$
- $$\langle \underbrace{\Gamma_0 \wedge \cdots \wedge \Gamma_0}_{k \text{ times}} | \underbrace{\Gamma_0 \wedge \cdots \wedge \Gamma_0}_{k \text{ times}} \rangle = (k!)^2 \times \Big(\sum_{1 \le i_1 < \ldots < i_k \le m} \eta_{i_1}^2 \cdots \eta_{i_k}^2 \Big)$$
- Ex. $k = 3$, $$\Big(\frac{\sum_{i=1}^{m} \eta_i^6}{3} - \frac{\sum_{i=1}^{m} \eta_i^2 \times \sum_{j=1}^{m} \eta_j^4}{2} + \frac{\sum_{i=1}^{m} \eta_i^2 \times \sum_{j=1}^{m} \eta_j^2 \times \sum_{k=1}^{m} \eta_k^2}{6} \Big)$$

Antisymmetrized Product of Strongly-orthogonal Geminals

APSG aka 1-orthogonal geminals wave function

$$\forall i,j \; i \neq j, \; C_i^\dagger C_j = C_i^\dagger C'_j = C_i'^\dagger C'_j = 0$$

- Only the partition $k = 1 + 1 + \cdots + 1$ contributes
- $\sum_{\sigma,\sigma' \in \mathfrak{S}_k} \frac{tr\left[C^\dagger_{\sigma(1)} C'_{\sigma'(1)}\right] \; \cdots \; tr\left[C^\dagger_{\sigma(k)} C'_{\sigma'(k)}\right]}{k!} \rightarrow \sum_{\sigma \in \mathfrak{S}_k} \rightarrow k!$
- $\langle \Gamma_1 \wedge \cdots \wedge \Gamma_k | \Gamma'_1 \wedge \cdots \wedge \Gamma'_k \rangle = tr\left[C_1^\dagger C'_1\right] \; \cdots \; tr\left[C_k^\dagger C'_k\right]$

Antisymmetrized Product of Strongly-orthogonal Geminals

More explicitly

$$C_i = \begin{pmatrix} 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \lambda_1^i & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \lambda_{d_i}^i & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

- Restricted Hartree-Fock: $\forall i, \quad d_i = 1$

Strong orthogonality shortcomings

Ground state energies in Hartree at "experimental" geometry (STO-3G calculations)

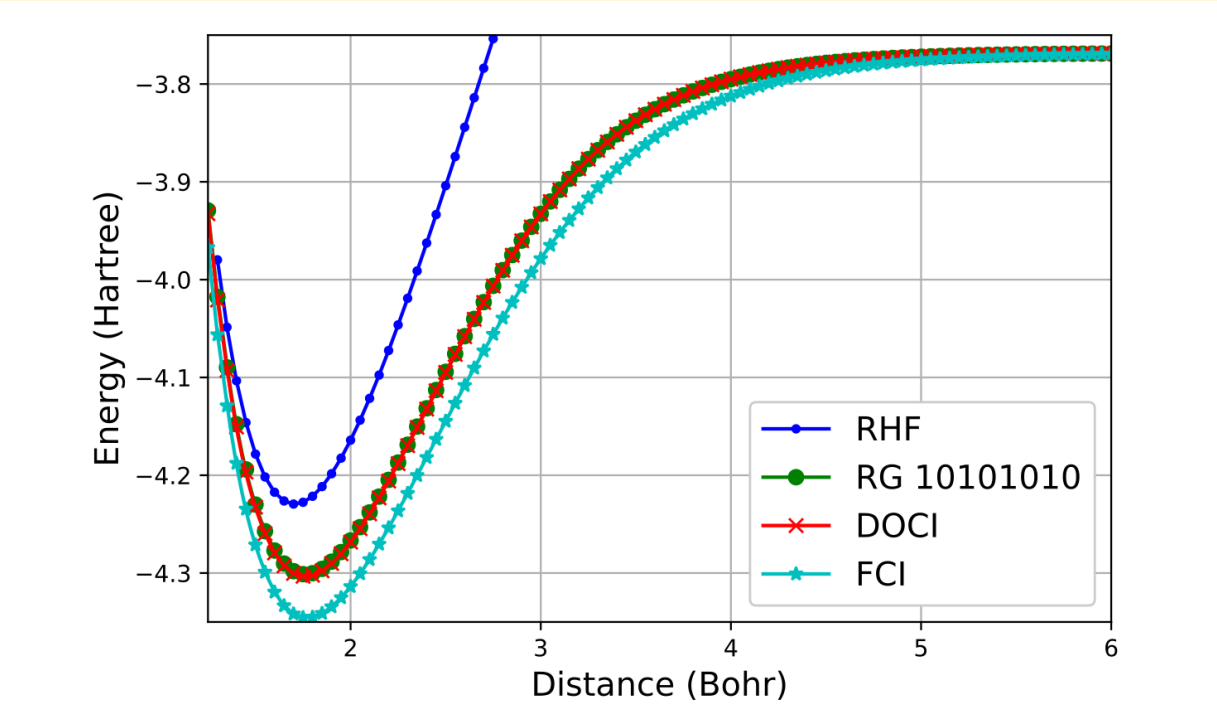
System	LiH	Be	Li_2	BeH_2	BH	Be_2
E^0_{FCI}	-7.882392	-14.403655	-14.667340	-15.594861	-24.809945	-28.804345
E^0_{GSCF}	-7.882372	-14.403655	-14.667114	-15.594715	-24.809938	-28.803212
with 2-orthogonality	-7.882368	-14.403654	-14.667090	-15.594703	-24.809920	-28.803080
APSG	-7.882203	-14.403630	-14.666584	-15.588630	-24.807908	-28.781789
E^0_{RHF}	-7.862002	-14.351880	-14.638725	-15.559405	-24.752780	-28.698990

Electric dipole moments

System	LiH	BH
D^z_{FCI}	-4.6201	0.6138
D^z_{GSCF}	-4.6197	0.6138
with 2-orthogonality	-4.6189	0.6142
D^z_{APSG}	-4.6269	0.6861
D^z_{RHF}	-4.8578	0.9569

P. Cassam-Chenaï, V. Rassolov, *Chemical Physics Letters* **487**, *147–152*, 2010.

Seniority zero shortcomings



Fecteau, et al., *The Journal of Chemical Physics* **156**, *194103*, 2022.

RHF: Restricted Hartree-Fock method

Ridchardson-Gaudin

DOCI: doubly occupied configuration interaction (with optimized orbitals)

FCI: Full Configuration Interaction

Extended Permutation Invariant 2-Orthogonal 2D-blocks

- $1D$ or $2D$-block diagonal matrices:

$$C_i = \begin{pmatrix} \lambda_1^i \tau_1^i & 0 & 0 & 0 \\ 0 & \lambda_2^i \tau_2^i & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_m^i \tau_m^i \end{pmatrix}$$

$$\lambda_j^i \ scalars$$

$$and \ \tau_k^i \in \{Id_1\}$$

$$or \ \tau_k^i \in \{Id_2, \sigma_x, i\sigma_y, \sigma_z\}$$

$$or \ \tau_k^i \in \{\sigma(\alpha), \sigma_x, i\sigma_y\}$$

$$\text{with } \sigma(\alpha) = \sqrt{2} \begin{pmatrix} Sin(\alpha) & 0 \\ 0 & Cos(\alpha) \end{pmatrix}$$

- if $\tau_k^i = \tau_k^j$ and $i \neq j$ then $\lambda_i \lambda_j = 0$

Singlet geminals with two 2D-block types

$$\forall k, \tau_k^i \in \{\sigma(\alpha), \sigma_x\}$$

$\rightsquigarrow$ Minimalistic way to go beyond strong orthogonality and seniority 0 approximations: $\sigma_x \rightsquigarrow \left(\phi_k \wedge \bar{\phi}_{k+1} + \phi_{k+1} \wedge \bar{\phi}_k\right)$ i.e. singlet "mono-excitations" to the next orbital.

$\rightsquigarrow$ computational cost is under control (easy traces, only $\frac{3\times m}{2}$ parameters to optimize)

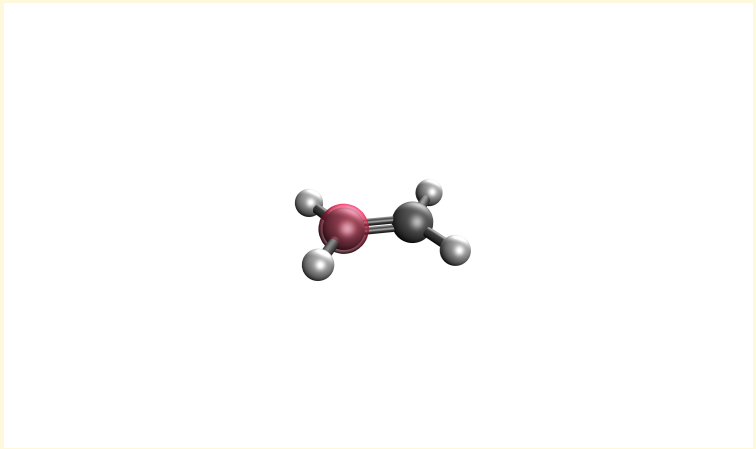
$\rightsquigarrow$ flexibility: includes APSG, hence HF, GVB, ...

$\rightsquigarrow$ implementation in TONTO (actually all block types are implemented)

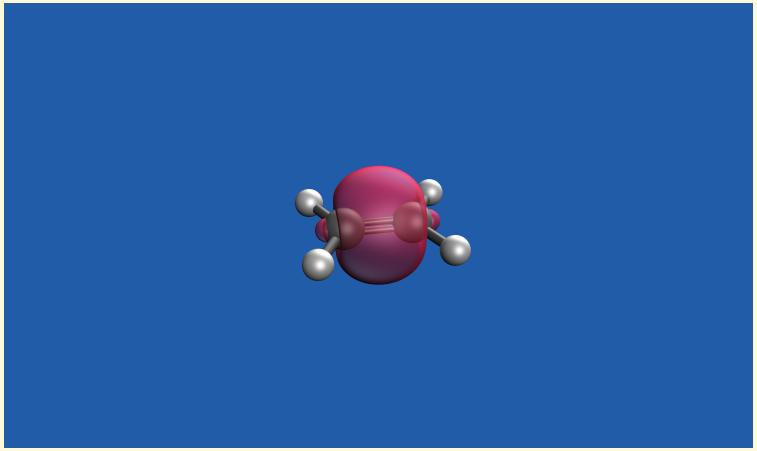
N2 example input file

```
{
name=     N2
basis_name= 6-31G(d)
basis_directory= /home/cassam/QCP/Perso-ware/tonto/basis_sets
use_spherical_basis= TRUE
charge=   0
multiplicity= 1
atoms= {
keys= { label=  { units= angstrom }  pos=  }
data= {
   N  0.0d0          0.0d0         -0.55025d0
   N  0.0d0          0.0d0         +0.55025d0
}
}
put
scfdata= {
   kind=            rhf
}
pauli_block_geminals= {
guess_type= readin
block_partition= 0 14 0
readin_mos= yes  !to read in molecular orbitals instead of using canonical Hartree-Fock orbitals
make_integrals= TRUE !an <ij|kl> archive file will be used
opt_gd2_angles= TRUE
grad_tol= 0.00000001
coef_tol= 0.00000001
bra_is_ket= TRUE
bra_set=
   GD2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2
 +1.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0 -0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0
   ID2    GD2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2
 +0.0d0 +1.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0 -0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0
   ID2    ID2    GD2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2
 +0.0d0 +0.0d0 +1.0d0 +0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0 -0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0
   ID2    ID2    ID2    GD2    GD2    GD2    GD2    ID2    ID2    ID2    ID2    ID2    ID2    ID2
 +0.0d0 +0.0d0 +0.0d0 +1.0d0 +1.0d0 -1.0d0 +1.0d0 -0.0d0 -0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0
   ID2    ID2    ID2    ID2    ID2    ID2    ID2    GD2    GD2    GD2    ID2    ID2    ID2    ID2
 +0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 +0.0d0 +0.0d0 -1.0d0 -1.0d0 -1.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0
   ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    ID2    GD2    GD2    SX     SX
 +0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 -0.0d0 +0.0d0 +0.0d0 +0.0d0 +0.0d0 +1.0d0 -1.0d0 -0.1d0 +0.1d0
   SX     SX     SX     SX     SX     SX     SX     SX     SX     SX     SX     SX     GD2    GD2
 +0.1d0 -0.1d0 +0.1d0 -0.1d0 +0.1d0 -0.1d0 +0.1d0 -0.1d0 -0.1d0 -0.1d0 +0.1d0 +0.1d0 +1.0d0 -1.0d0
 gd2_info= 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0 1.5d0
}
pbg_optimization
put_pauli_block_geminals
}
```

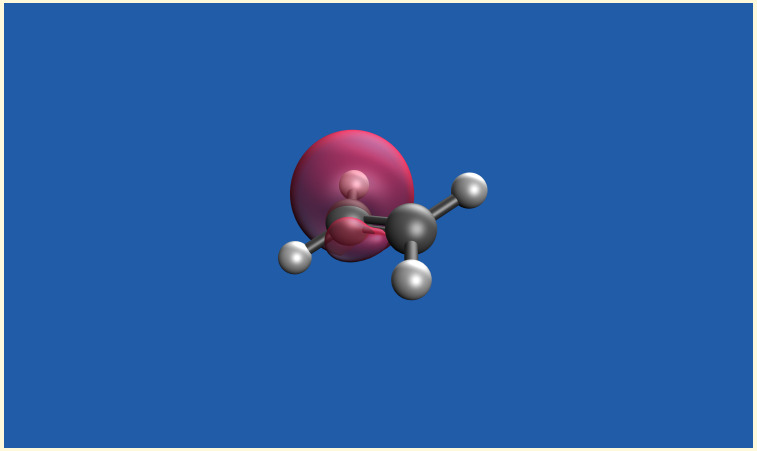
Ethylene geminal isodensity surfaces



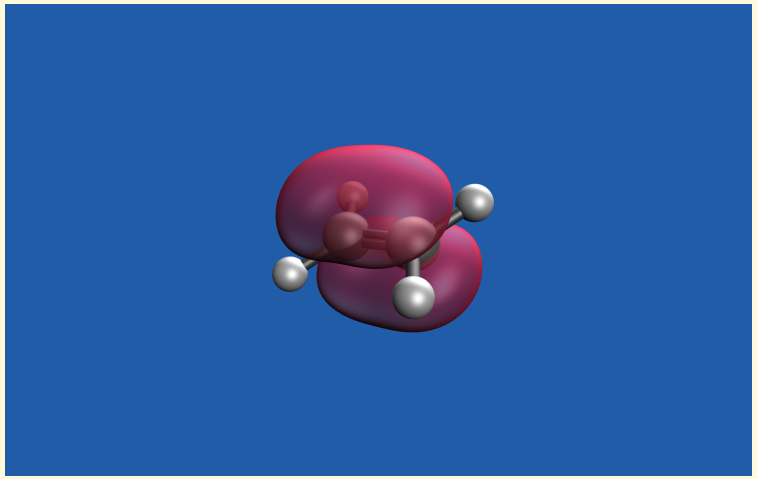
core geminal



C-C σ geminal



C-H σ geminal



C-C π geminal

Choosing $\sigma(\alpha)$ locations

- Location fixed to be able to reproduce at least the APSG ansatz

$$g_i^{APSG} = \underbrace{D_1^i\ \phi_{i_1} \wedge \bar{\phi}_{i_1} + D_2^i\ \phi_{i_2} \wedge \bar{\phi}_{i_2}}_{\lambda_1^i \sigma(\alpha_1)} + \cdots + \underbrace{D_{2k-1}^i\ \phi_{i_{2k-1}} \wedge \bar{\phi}_{i_{2k-1}} + D_{2k}^i\ \phi_{i_{2k}} \wedge \bar{\phi}_{i_{2k}}}_{\lambda_k^i \sigma(\alpha_k)}$$

$$|D_1^i| > |D_2^i| > \cdots > |D_{2k-1}^i| > |D_{2k}^i|$$

$$D_1^i = \sqrt{2}\lambda_1^i\ Sin(\alpha_1), \quad D_2^i = \sqrt{2}\lambda_1^i\ Cos(\alpha_1), \quad \cdots$$

- Blocks are arranged in decreasing order of geminal coefs, D_j^i while geminal are ordered in increasing energies.

- If "Arai" sub-Hilbert space $\mathcal{H}_o^i := span(\phi_{i_1}, \cdots, \phi_{i_p})$, is of odd dimension: $p = 2k + 1$ (odd number of pairs in the expansion)
$\rightsquigarrow$ add a 1D-block for the less populated pair with coef. D_{2k+1}^i

Choosing σ_x locations

- Still $(N_{geminals}-1)^{N_{blocks}}$ possibilities to choose the σ_x's positions!

$$g_i^{APSG} = \underbrace{D_1^i\ \phi_{i_1} \wedge \bar{\phi}_{i_1} + D_2^i\ \phi_{i_2} \wedge \bar{\phi}_{i_2}}_{\text{Block 1 of geminal i}} + \cdots + \underbrace{D_{2k-1}^i\ \phi_{i_{2k-1}} \wedge \bar{\phi}_{i_{2k-1}} + D_{2k}^i\ \phi_{i_{2k}} \wedge \bar{\phi}_{i_{2k}}}_{\text{Block k of geminal i}}$$

- However, in general $|D_1^i| \gg |D_2^i| > \cdots > |D_{2k-1}^i| > |D_{2k}^i|$

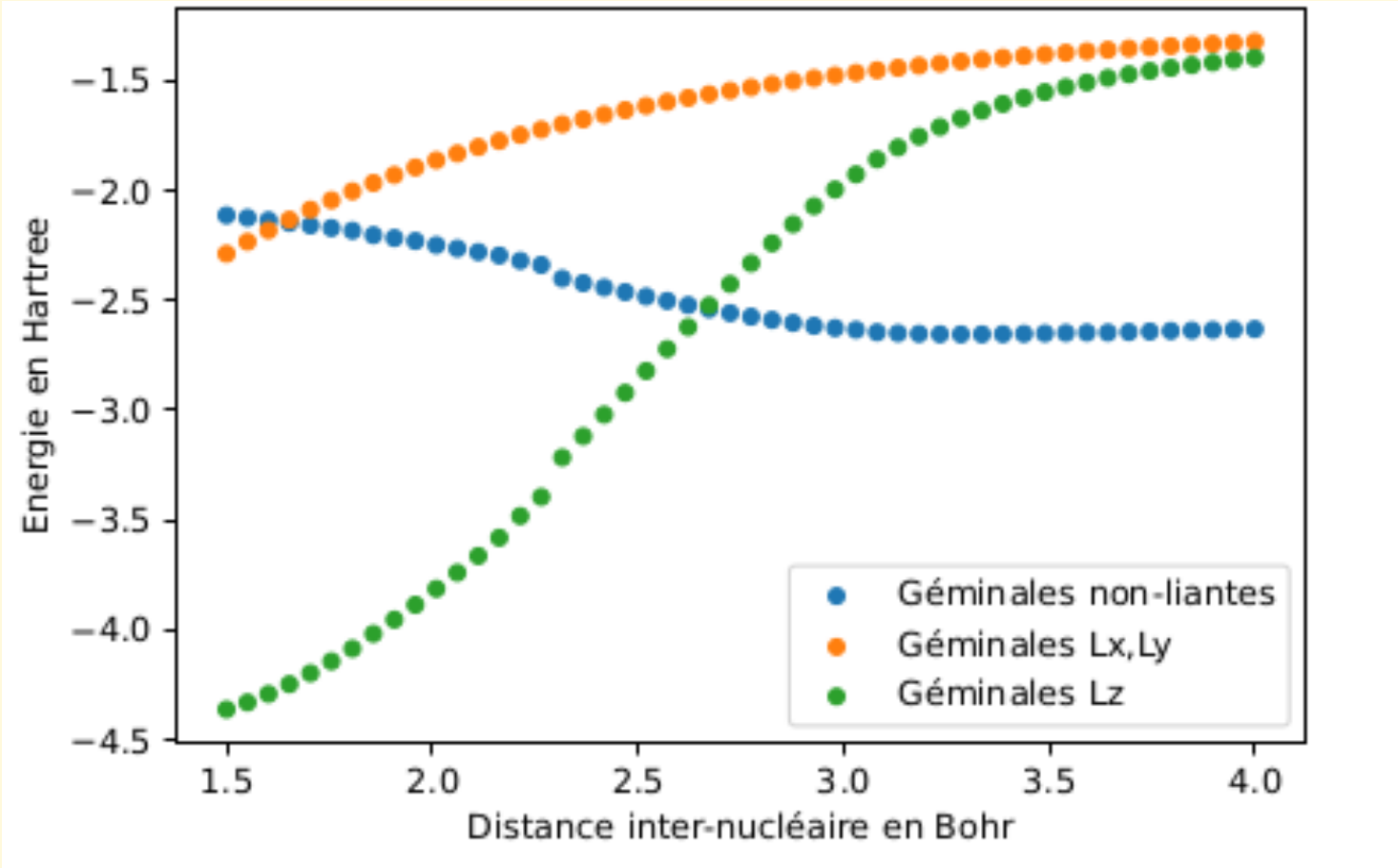
- It is useless to excite an orbital which is weakly populated

$\rightsquigarrow$ σ_x restricted to blocks that are the first block of one of the geminals

- Huge reduction down to $(N_{geminals}-1)^{N_{geminals}}$ possibilities to choose the positions of the σ_x's

N2 APSG Guess

color code

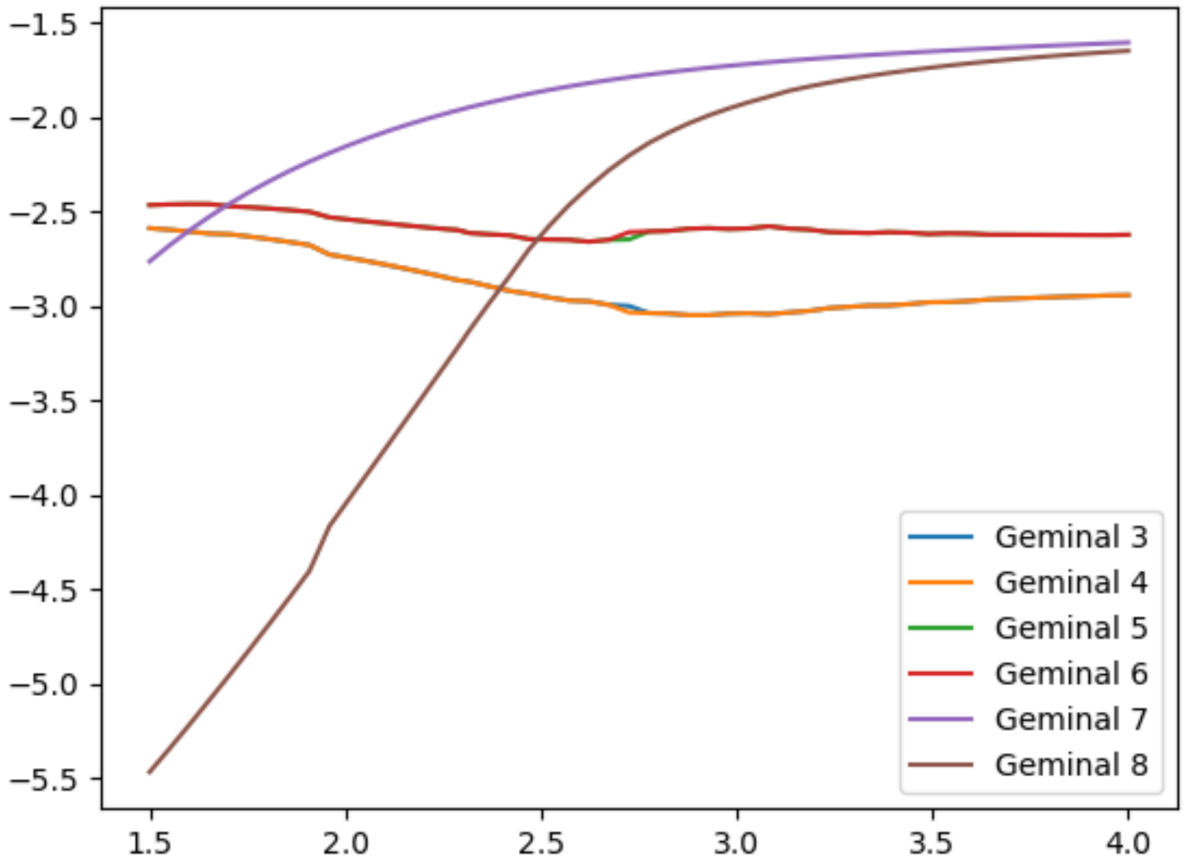


Valence geminal energies

O2 APSG Guess

color code

Valence geminal energies



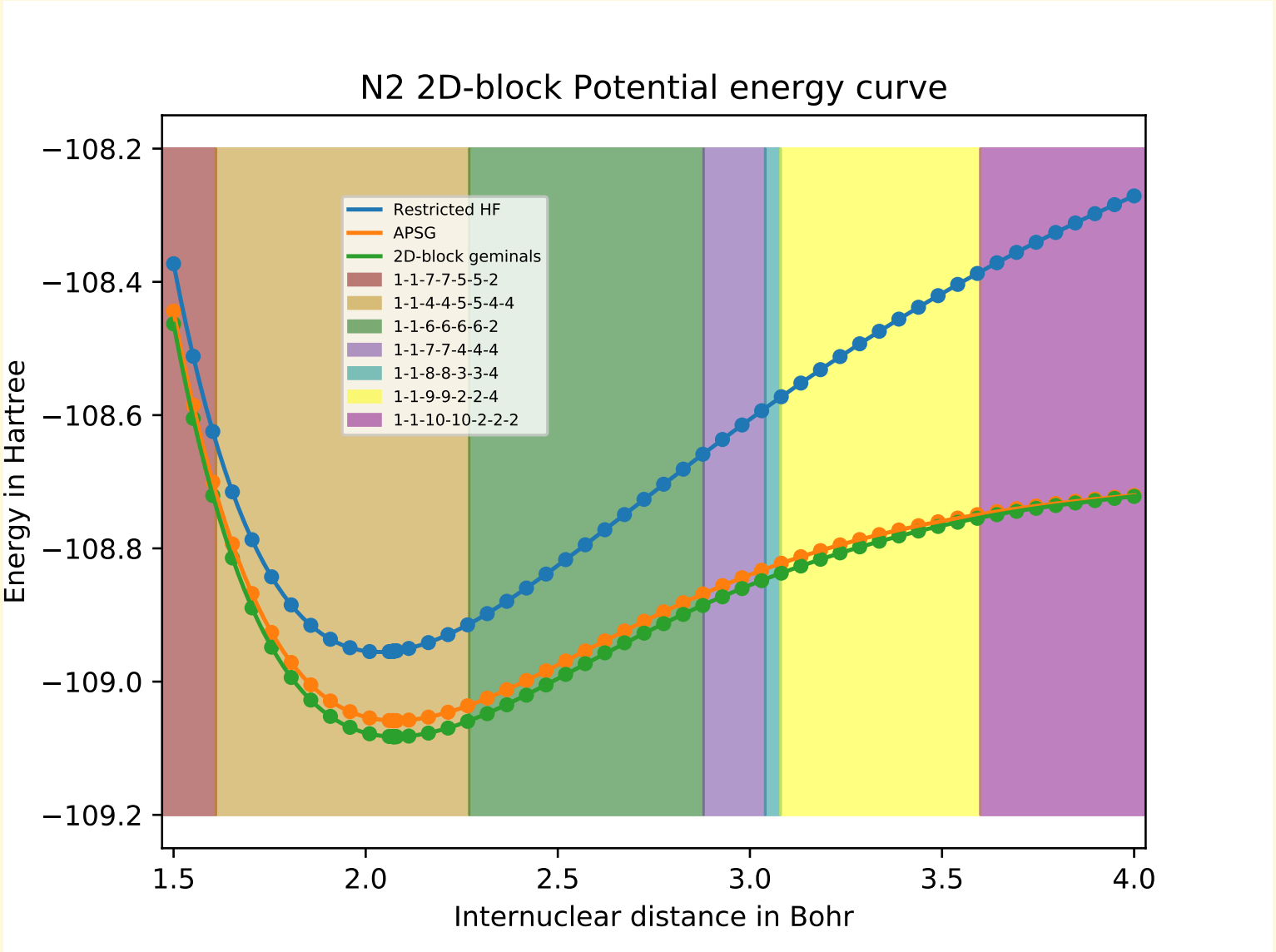
2D, 2 types, Block APG calculations

Lowest singlet energies from $2D$-block, $\{\sigma(\alpha), \sigma_x\}$-type calculations at geometries derived from experiment (in Hartree)

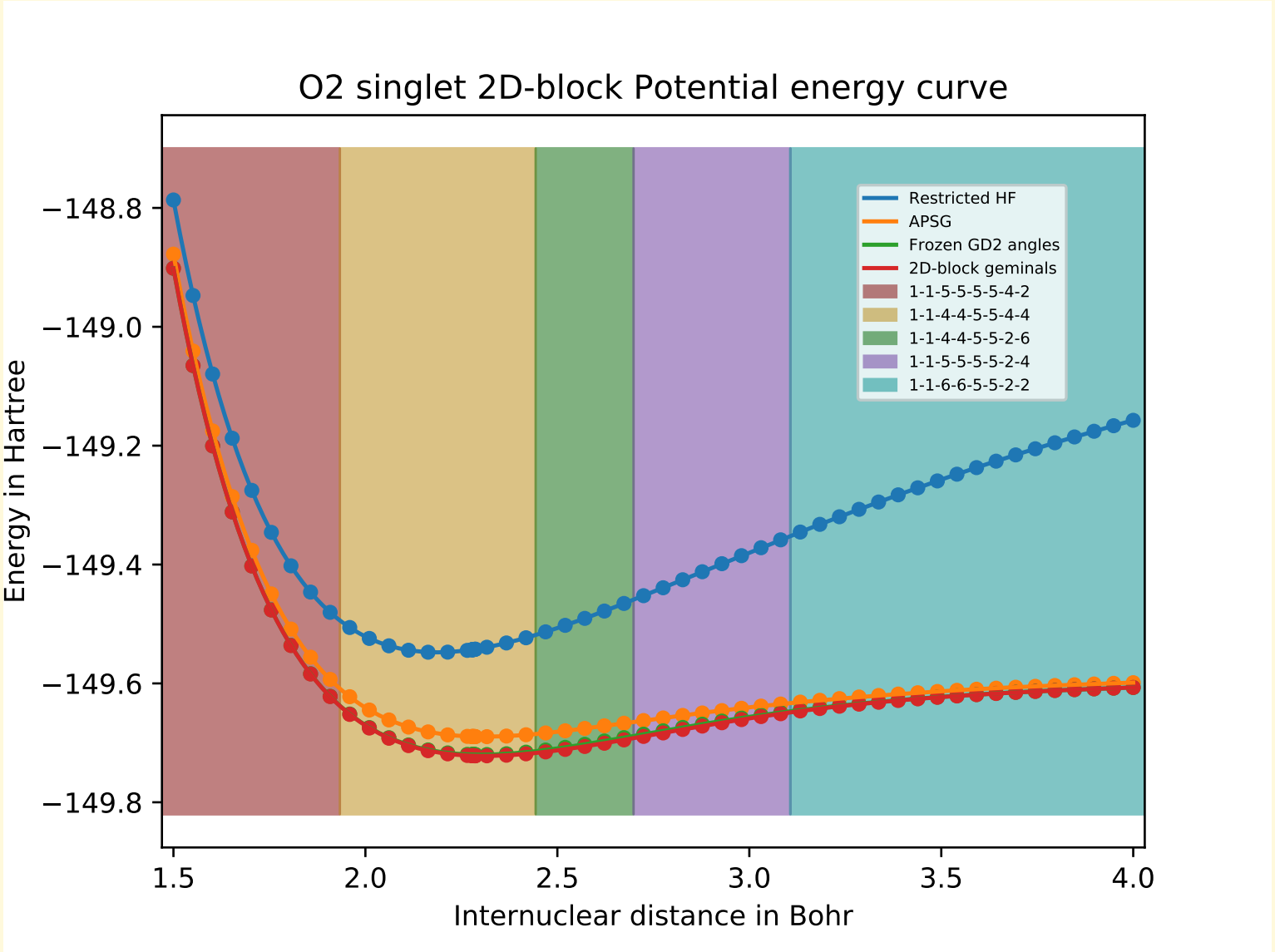
Homonuclear diatomics	N_2	N_2	O_2	F_2
r_e^{exp} $(\AA)$	1.098	1.098	1.2075	1.412
Basis set	cc-pVTZ	cc-pVDZ	cc-pVDZ	cc-pVDZ
$E^0_{\{\sigma(\alpha),\sigma_x\}}$	-109.134390	-109.082362	-149.718718	-198.853436
E^0_{APSG}	-109.110630	-109.058601	-149.689436	-198.840387
E^0_{RHF}	-108.983412	-108.954087	-149.542930	-198.685664
nb. of geminals	7	7	8	9
nb. of blocks	30	14	14	14
nb. of parameters	70/90	35/42	36/42	42/42
nb. terms in overlap	64974	360	343	1116
cpu time (s)	0.4	0.01	0.01	0.02

P. Cassam-Chenaï, T. Perez, D. Accomaso, *The Journal of Chemical Physics* **158**, *074106*, 2023.

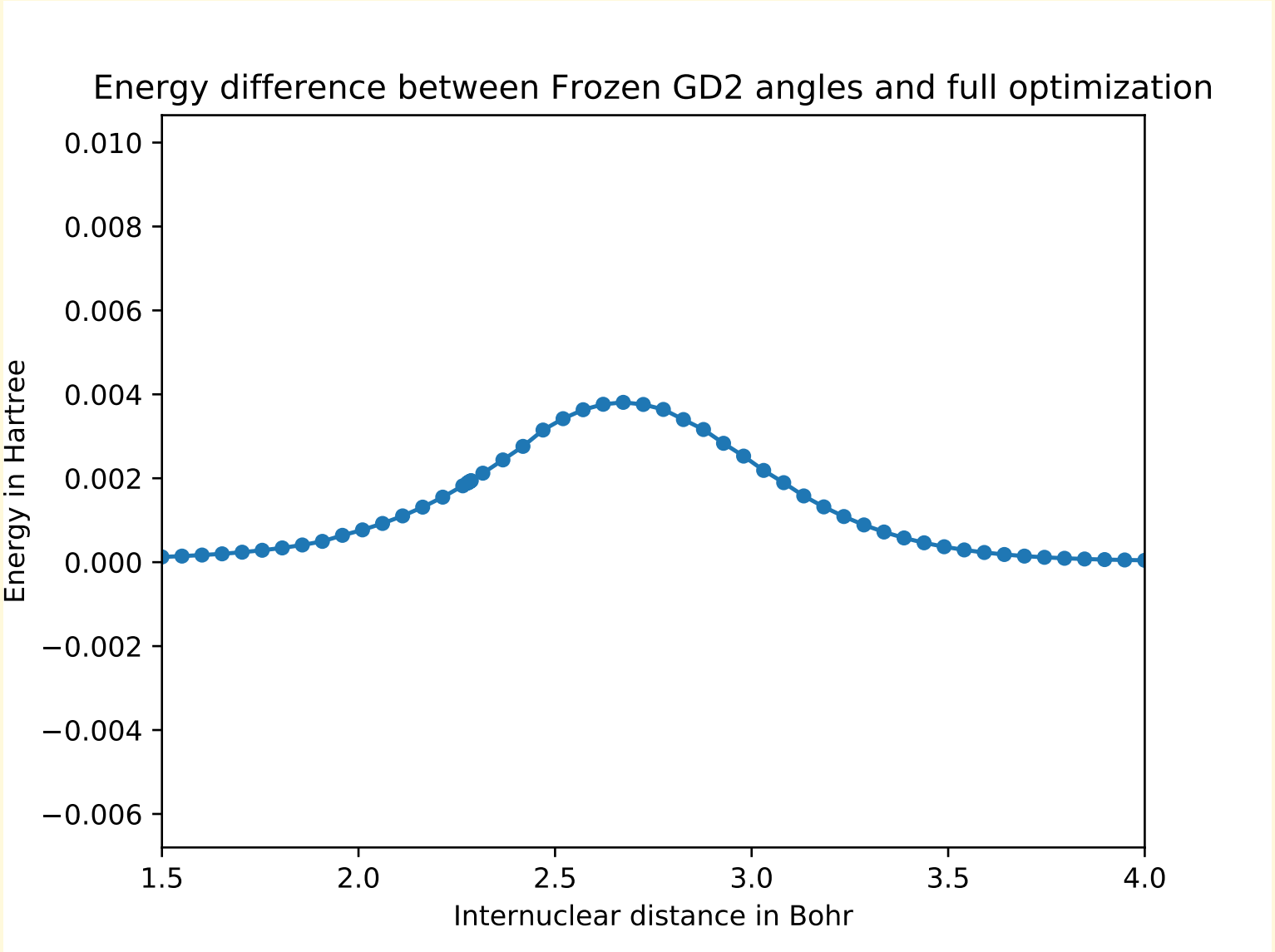
N2 PECs (cc-pVDZ basis)



O2 singlet PECs (cc-pVDZ basis)



O2 singlet frozen angles (cc-pVDZ basis)



O2 singlet calculations

	RHF	APSG	FA-2D-EPI2O	2D-EPI2O	exp.(NIST)
r_e ($\AA$)	1.1572	1.2220	1.2175	1.2224	1.2269
E_{min} (Hartree)	-149.548	-149.690	-149.720	-149.722	
$E_0 - E_{min}$ (cm^{-1})	992.51	704.37	756.18	740.86	738.53
ω_e (cm^{-1})	1989.07	1420.23	1520.99	1488.97	1483.50
$E_1 - E_0$ (cm^{-1})	1969.42	1376.83	1487.50	1459.13	1457.7

r_e equilibrium distance
E_{min} minimum energy
$E_0 - E_{min}$ Zero point energy
ω_e harmonic frequency
$E_1 - E_0$ fundamental frequency.

Lessons to take home

1. Proof of concept that 2D-block geminals is more accurate than APSG
2. Partitionning of MOs into blocks dictated by the APSG Arai space partitioning and decreasing orbital occupation ordering seems optimal
3. σ_α-matrix type and 1D-block to reproduce APSG
4. Simple rules to choose the best geminal for the mono-excitation of each block i.e the σ_x-matrix type $\rightsquigarrow$ tamed combinatorics

Perspectives

1. TO DO: Release the constraint of only one σ_x-matrix type per 2D-block to avoid some symmetry breaking $\rightsquigarrow$ Implementation on quantum computers
2. TO DO: Investigate the hierarchy 3D-block, 4D-block $\cdots$
3. TO DO: Investigate properties such as dipole moments (to compute line intensities)
4. Other ideas to simplify the general overlap formula?

APsetG overlap formula

* General overlap closed formula: $\Gamma_i \leftrightarrow C_i$, $\Gamma'_j \leftrightarrow C'_j$

$$\langle \Gamma_1 \wedge \cdots \wedge \Gamma_k | \Gamma'_1 \wedge \cdots \wedge \Gamma'_k \rangle = \langle 0 | a(\Gamma_k) \cdots a(\Gamma_2) a(\Gamma_1) a^\dagger(\Gamma'_1) a^\dagger(\Gamma'_2) \cdots a^\dagger(\Gamma'_k) | 0 \rangle =$$

$$\sum_{\substack{0 \le N_{k,0}, \ldots, N_{k,k} \le k \\ \sum_{i=0}^{k} N_{k,i} = \sum_{i=0}^{k} i N_{k,i} = k}} (-1)^{N_{k,0}} \sum_{\sigma, \sigma' \in \mathfrak{S}_k} \prod_{i=1}^{k} \frac{\prod_{j=1}^{N_{k,i}} tr\left[C^\dagger_{\sigma\left(\sum_{p=0}^{i-1} p N_{k,p} + (j-1)i+1\right)} C'_{\sigma'\left(\sum_{p=0}^{i-1} p N_{k,p} + (j-1)i+1\right)} \cdots C^\dagger_{\sigma\left(\sum_{p=0}^{i-1} p N_{k,p} + ji\right)} C'_{\sigma'\left(\sum_{p=0}^{i-1} p N_{k,p} + ji\right)} \right]}{i^{N_{k,i}} \, N_{k,i}!},$$

* Ex.: $k = 3 \rightsquigarrow$ 3 partitions of k

– $N_{3,1} = 3,\ N_{3,2} = 0,\ N_{3,3} = 0 \quad \rightsquigarrow \quad 3 * 1 + 0 * 2 + 0 * 3 = 3 = 1 + 1 + 1, \quad N_{3,0} = 0$

$$\sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{tr\left[C^\dagger_{\sigma(1)} C'_{\sigma'(1)}\right] \, tr\left[C^\dagger_{\sigma(2)} C'_{\sigma'(2)}\right] \, tr\left[C^\dagger_{\sigma(3)} C'_{\sigma'(3)}\right]}{3!}$$

– $N_{3,1} = 1,\ N_{3,2} = 1,\ N_{3,3} = 0 \quad \rightsquigarrow \quad 1 * 1 + 1 * 2 + 0 * 3 = 3 = 0 + 1 + 2, \quad N_{3,0} = 1$

$$- \sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{tr\left[C^\dagger_{\sigma(1)} C'_{\sigma'(1)}\right] \, tr\left[C^\dagger_{\sigma(2)} C'_{\sigma'(2)} C^\dagger_{\sigma(3)} C'_{\sigma'(3)}\right]}{2}$$

– $N_{3,1} = 0,\ N_{3,2} = 0,\ N_{3,3} = 1 \quad \rightsquigarrow \quad 0 * 1 + 0 * 2 + 1 * 3 = 3 = 0 + 0 + 3, \quad N_{3,0} = 2$

$$\sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{tr\left[C^\dagger_{\sigma(1)} C'_{\sigma'(1)} C^\dagger_{\sigma(2)} C'_{\sigma'(2)} C^\dagger_{\sigma(3)} C'_{\sigma'(3)}\right]}{3}$$

We acknowledge the SIGAMM Mesocentre for hosting the CONVIV code project and the JAD laboratory computer facilities (SMP computer with ScaleMP aggregation system administrated by J.-M. Lacroix and R. Ruelle).
This work was supported by the French government through the France 2030 investment plan managed by the National Research Agency (ANR), as part of the Initiative of Excellence Université Côte d'Azur under reference number ANR-15-IDEX-01. The authors are grateful to the Université Côte d'Azur's Center for High-Performance Computing (OPAL infrastructure) for providing resources and support.