Multiconfigurational methods: past, present and the road ahead

Stefan Knecht

Content of the lectures

- Concepts

Past

- CASCI
- MCSCF/CASSCF
- MR Dynamical Correlation Approaches

Present: “Sample” CI coefficients using an advanced optimisation algorithm

- Density Matrix Renormalisation Group

Road ahead: Beyond “classical quantum chemistry”

- Quantum Chemistry on Quantum Computers

Content of the lectures

- Concepts

Past: kind of ...

- CASCI
- MCSCF/CASSCF
- MR Dynamical Correlation Approaches

Present and future?: “Sample” CI coefficients using an advanced optimisation algorithm

- Density Matrix Renormalisation Group

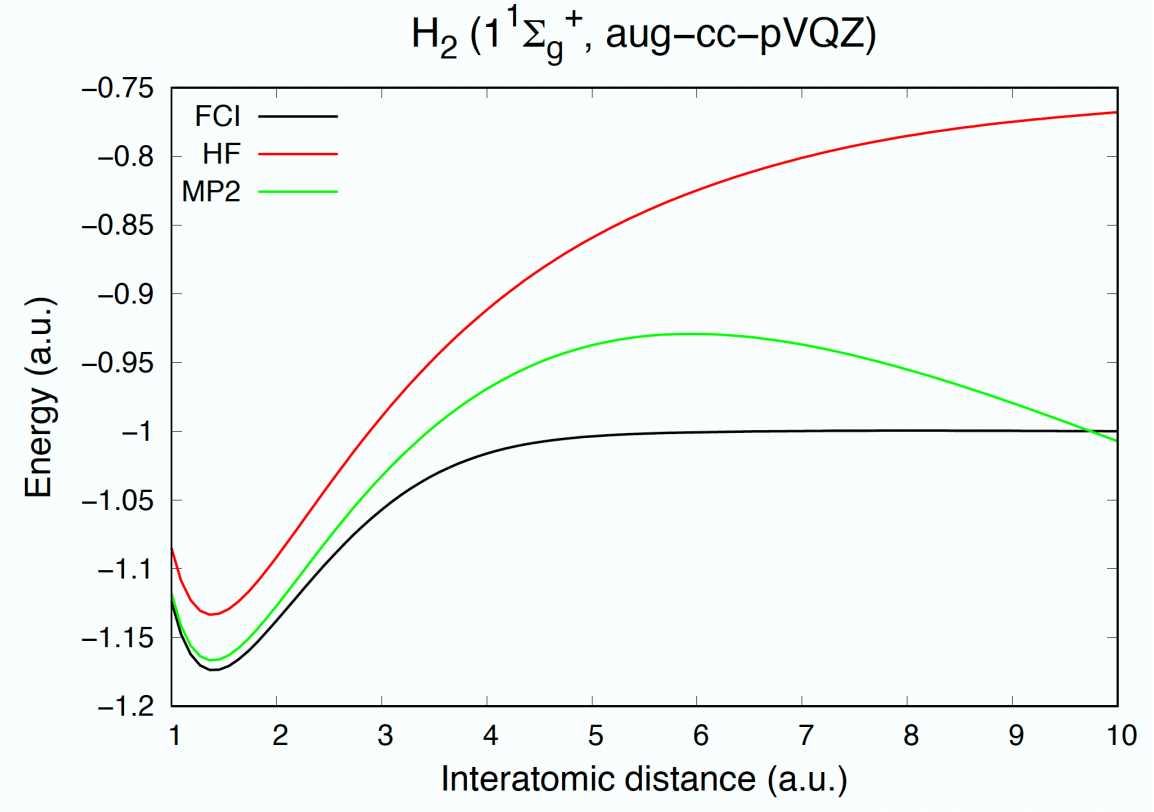
On the road to: Beyond “classical quantum chemistry”

- Quantum Chemistry on Quantum Computers

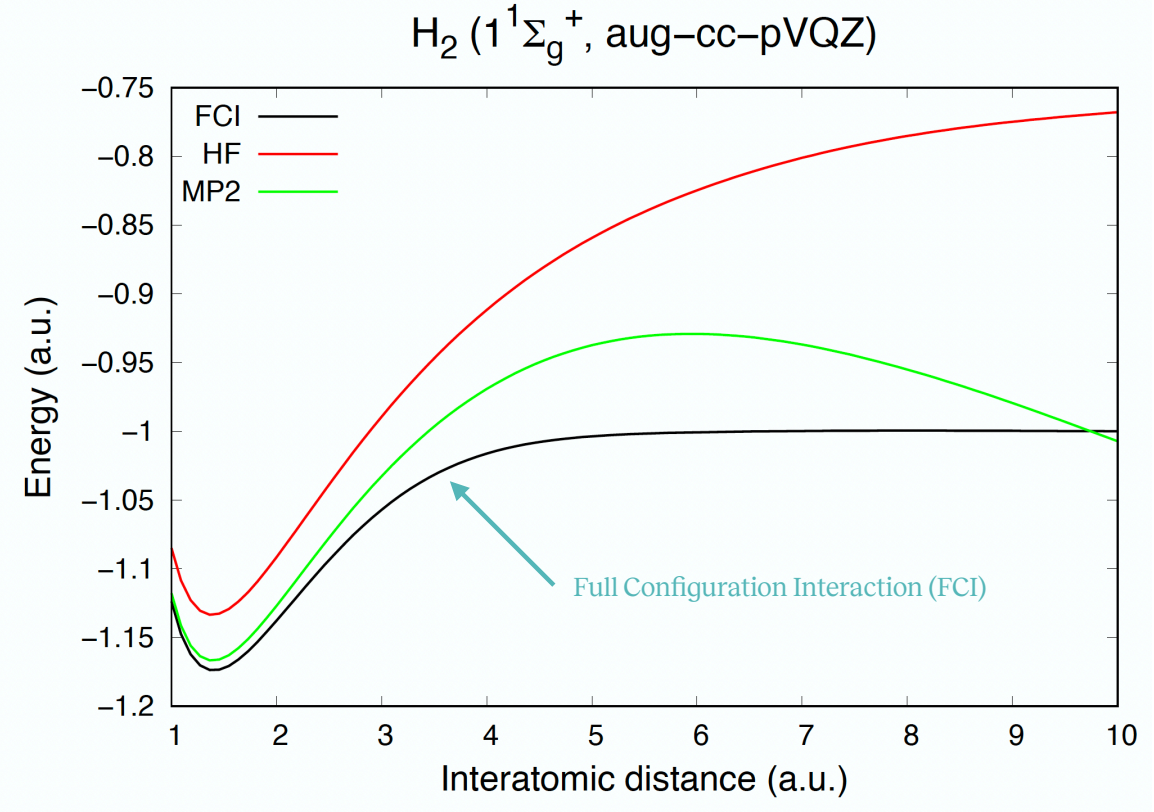
Concepts

First example: Things to learn from H_2...

Potential energy curve of H_2



Potential energy curve of H_2



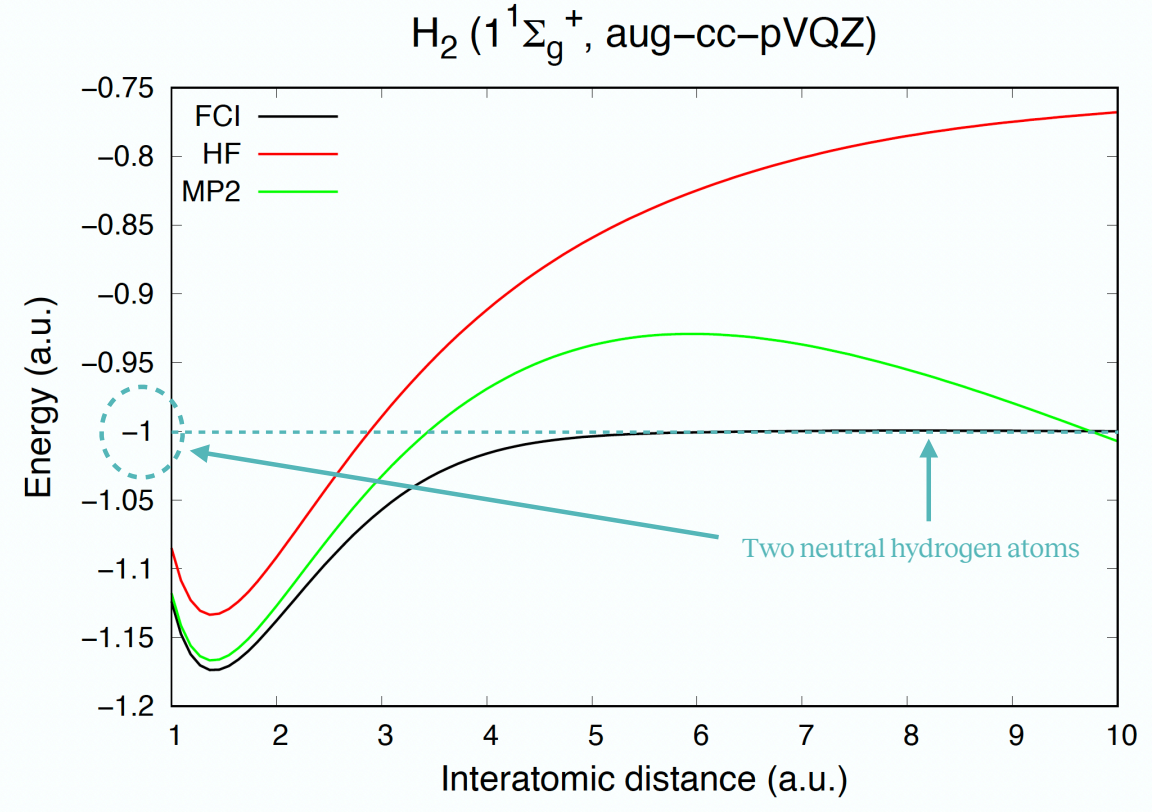
A quick reminder on CI

- Expand wave function as a linear combination of Slater determinants

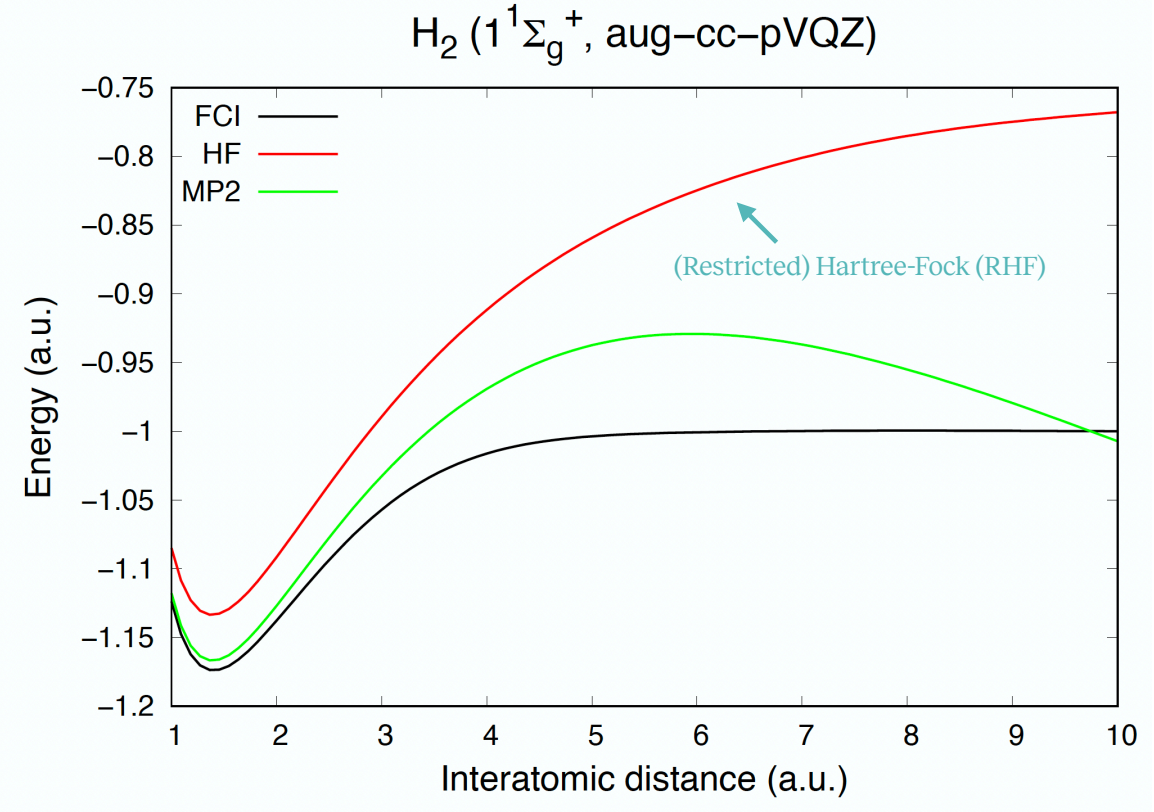
$$|\Psi\rangle = \sum_{\zeta} c_{\zeta} \left|\Phi_{\zeta}\right\rangle$$

- Orbital coefficients are fixed
- The parameters c_{ζ} are chosen to minimise the energy
- Leads to eigenvalue problem $Hc = Ec$ with $H_{\zeta\eta} = \left\langle \Phi_{\zeta} \middle| \hat{H} \middle| \Phi_{\eta} \right\rangle$
- If all determinants $\{\Phi\}$ of a given basis are included, we solve the Schrödinger equation projected onto the basis → full CI (FCI)

Potential energy curve of H_2



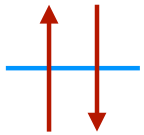
Potential energy curve of H_2



A quick reminder on RHF

- Minimize the energy of one Slater determinant or a symmetry/spin-defined combination of Slater determinants
- Individual unoccupied orbitals have a limited meaning (HOMO/LUMO → Koopman's theorem)

RHF wave function of H_2

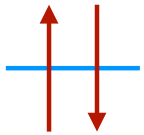


$\varphi_{1\sigma_g}$ Bonding orbital

$$\Phi^{RHF}(x_1, x_2) = \frac{1}{\sqrt{2}}\left(\varphi_{1\sigma_g.\alpha}(x_1)\varphi_{1\sigma_g.\beta}(x_2) - \varphi_{1\sigma_g.\alpha}(x_2)\varphi_{1\sigma_g.\beta}(x_1)\right)$$

$$= \varphi_{1\sigma_g}(\boldsymbol{r}_1)\varphi_{1\sigma_g}(\boldsymbol{r}_2) \times \frac{1}{\sqrt{2}}\left(\delta_{\sigma_1\alpha}\delta_{\sigma_2\beta} - \delta_{\sigma_2\alpha}\delta_{\sigma_1\beta}\right)$$

RHF wave function of H_2



$\varphi_{1\sigma_g}$ Bonding orbital

$$\Phi^{RHF}(x_1, x_2) = \frac{1}{\sqrt{2}}\left(\varphi_{1\sigma_g.\alpha}(x_1)\varphi_{1\sigma_g.\beta}(x_2) - \varphi_{1\sigma_g.\alpha}(x_2)\varphi_{1\sigma_g.\beta}(x_1)\right)$$

$$= \varphi_{1\sigma_g}(\boldsymbol{r}_1)\varphi_{1\sigma_g}(\boldsymbol{r}_2) \times \frac{1}{\sqrt{2}}\left(\delta_{\sigma_1\alpha}\delta_{\sigma_2\beta} - \delta_{\sigma_2\alpha}\delta_{\sigma_1\beta}\right)$$

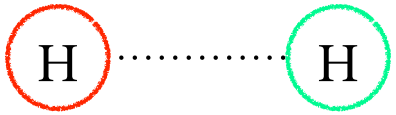
RHF wave function of stretched H_2 in a minimal basis

Bonding orbital

$$\varphi_{1\sigma_g}(\boldsymbol{r}) = \frac{1}{\sqrt{2}}\left(s_A(\boldsymbol{r}) + s_B(\boldsymbol{r})\right)$$

$$= \varphi_{1\sigma_g}(\boldsymbol{r}_1)\varphi_{1\sigma_g}(\boldsymbol{r}_2) \times \frac{1}{\sqrt{2}}\left(\delta_{\sigma_1\alpha}\delta_{\sigma_2\beta} - \delta_{\sigma_2\alpha}\delta_{\sigma_1\beta}\right)$$

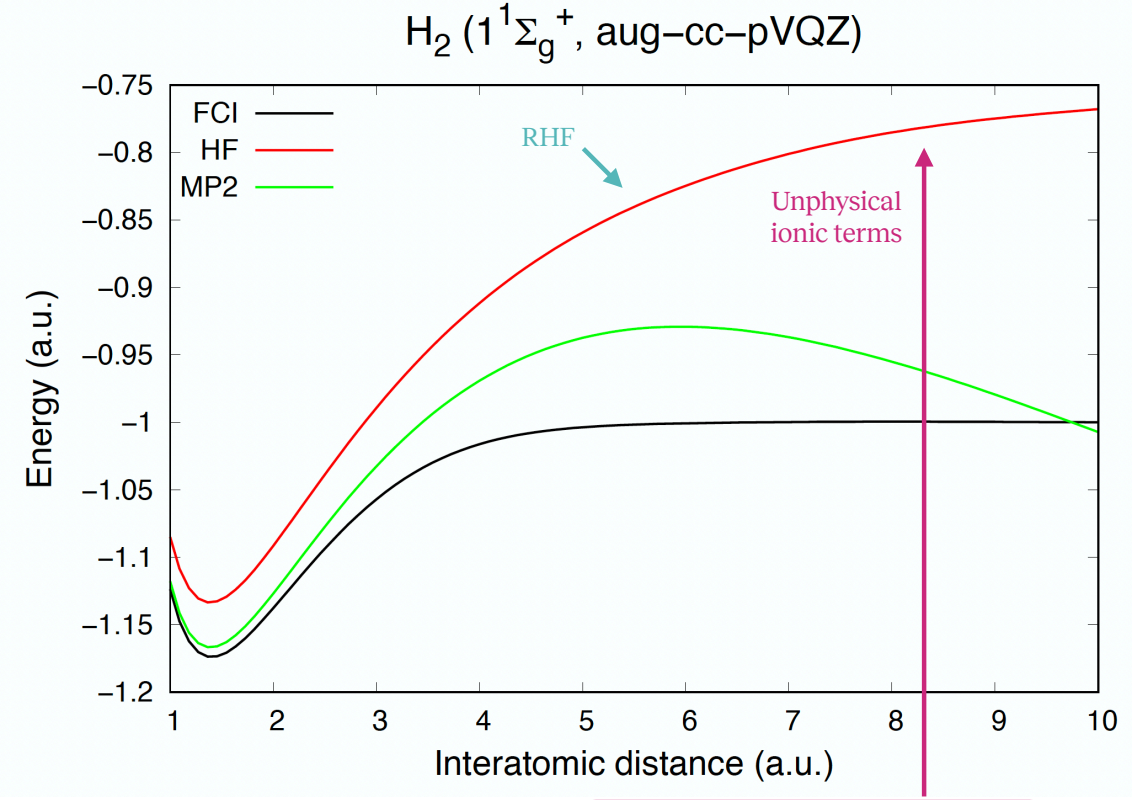
RHF wave function of stretched H_2 in a minimal basis



$$\frac{1}{2}\left(s_A(\boldsymbol{r_1})s_B(\boldsymbol{r_2}) + s_A(\boldsymbol{r_2})s_B(\boldsymbol{r_1}) + s_A(\boldsymbol{r_1})s_A(\boldsymbol{r_2}) + s_B(\boldsymbol{r_1})s_B(\boldsymbol{r_2})\right)$$

$$= \varphi_{1\sigma_g}(\boldsymbol{r}_1)\varphi_{1\sigma_g}(\boldsymbol{r}_2) \times \frac{1}{\sqrt{2}}\left(\delta_{\sigma_1\alpha}\delta_{\sigma_2\beta} - \delta_{\sigma_2\alpha}\delta_{\sigma_1\beta}\right)$$

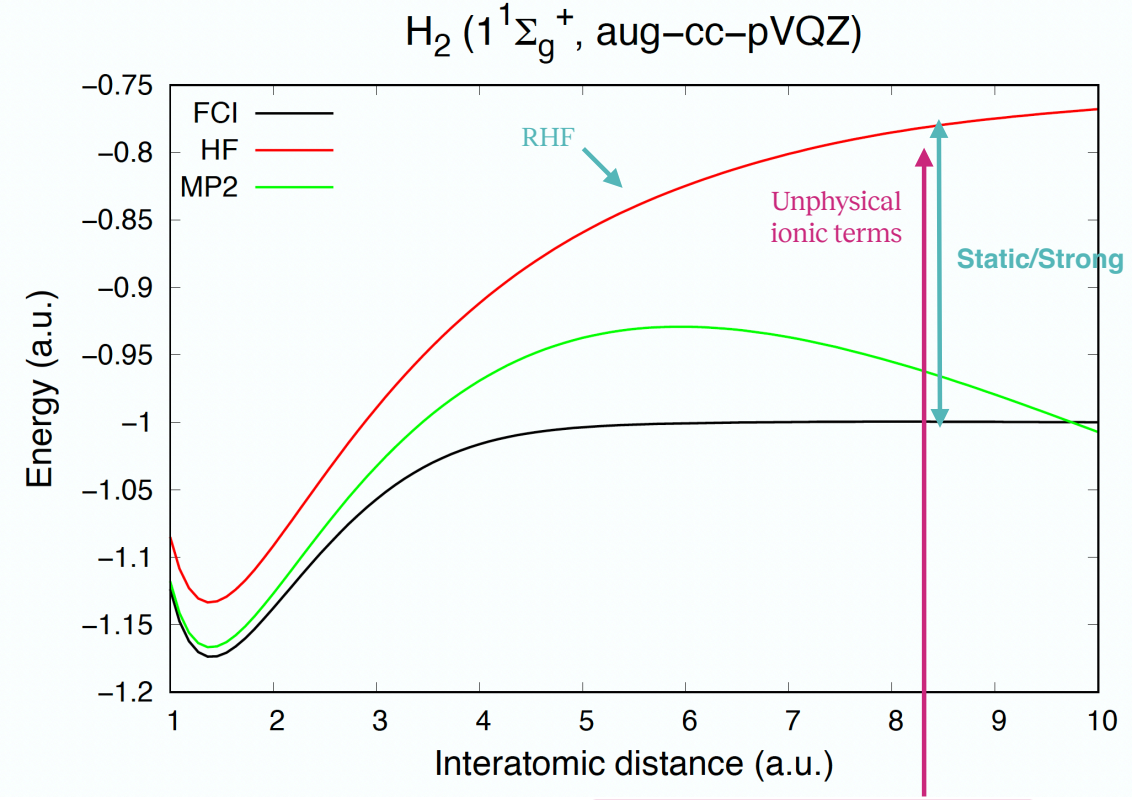
Potential energy curve of H_2



$$\frac{1}{2}\left(s_A(\boldsymbol{r_1})s_B(\boldsymbol{r_2}) + s_A(\boldsymbol{r_2})s_B(\boldsymbol{r_1}) + s_A(\boldsymbol{r_1})s_A(\boldsymbol{r_2}) + s_B(\boldsymbol{r_1})s_B(\boldsymbol{r_2})\right)$$

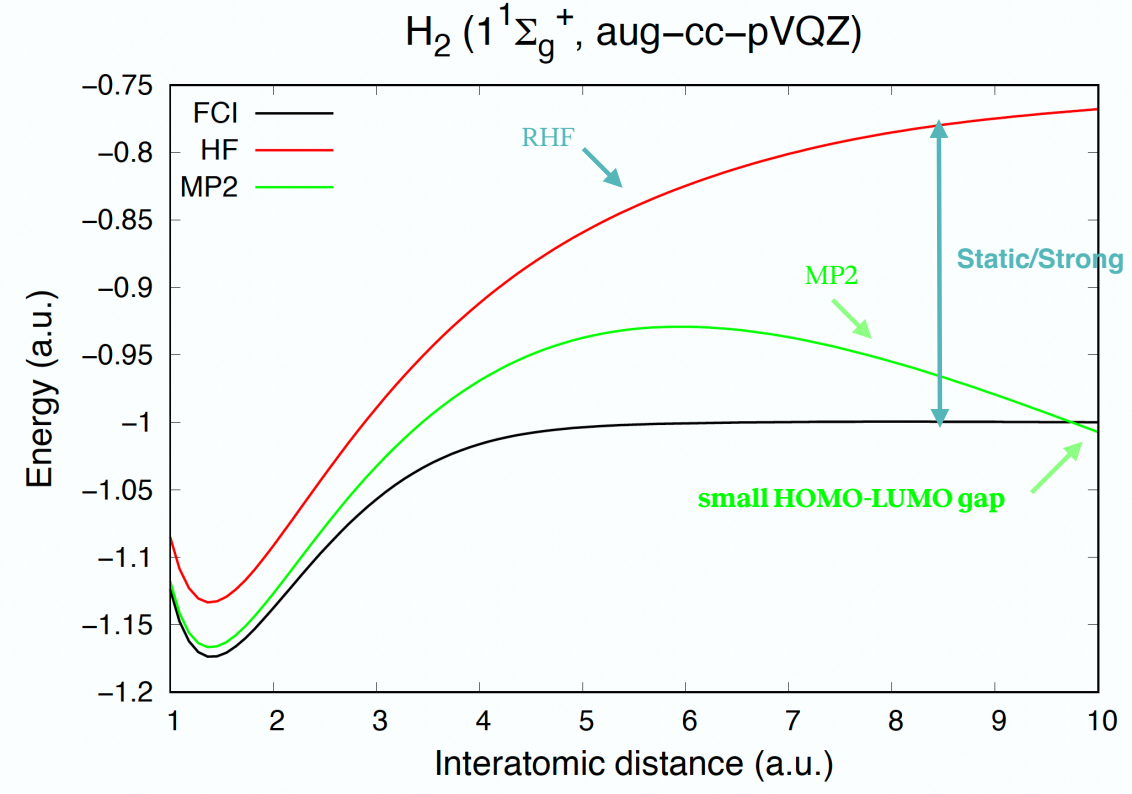
$H^- \cdots H^+ \qquad H^+ \cdots H^-$

Potential energy curve of H_2

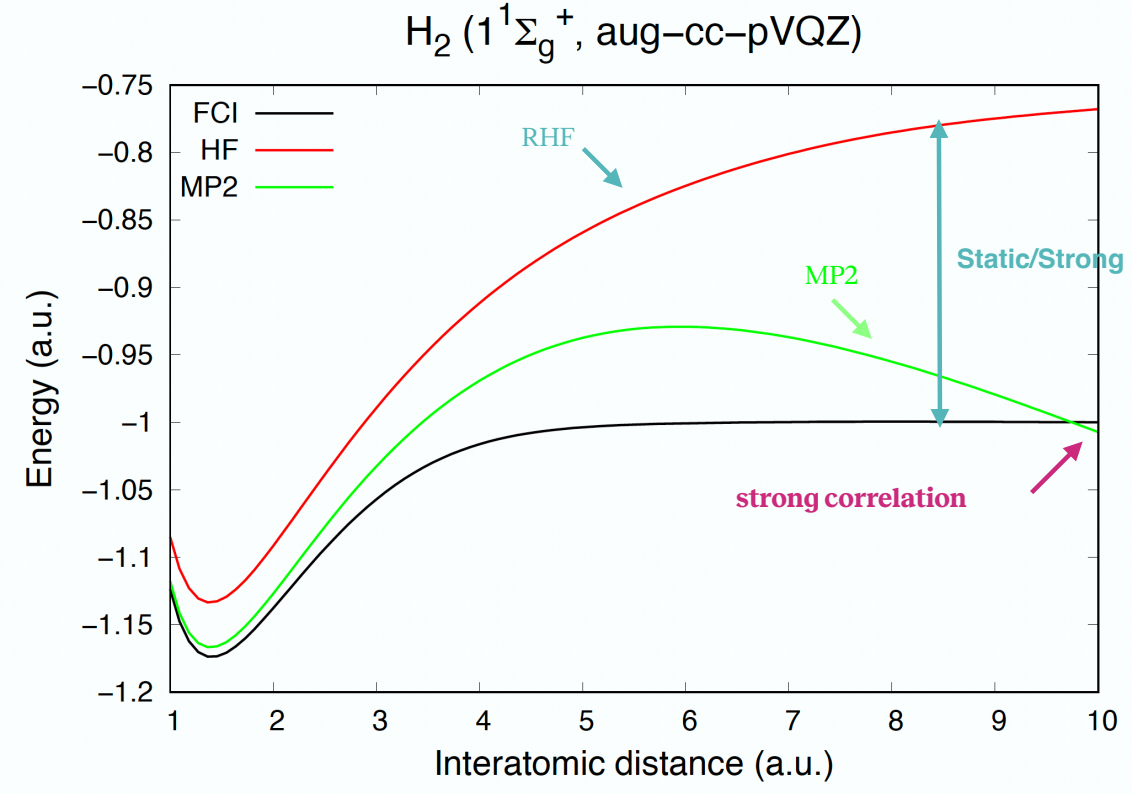


$$\frac{1}{2}\left(s_A(\boldsymbol{r_1})s_B(\boldsymbol{r_2}) + s_A(\boldsymbol{r_2})s_B(\boldsymbol{r_1}) + s_A(\boldsymbol{r_1})s_A(\boldsymbol{r_2}) + s_B(\boldsymbol{r_1})s_B(\boldsymbol{r_2})\right)$$

Potential energy curve of H_2



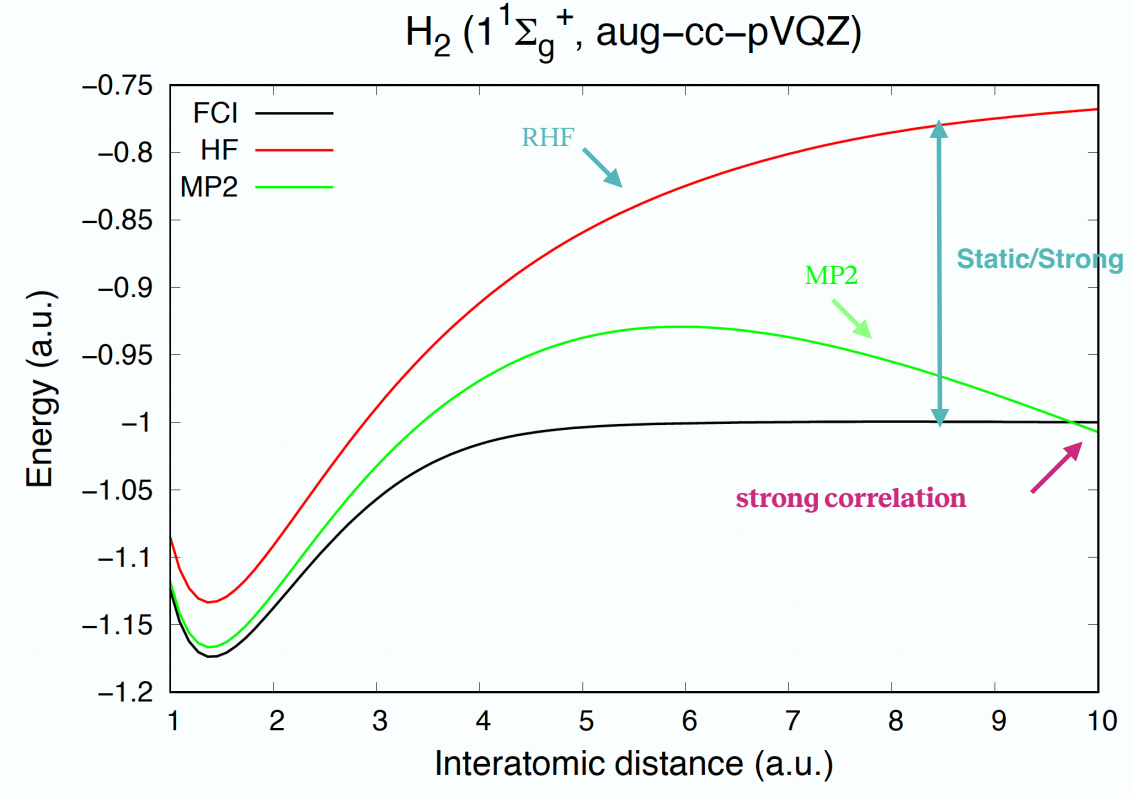
Potential energy curve of H_2



NB: "Static" usually describes why a single-reference picture fails (see following slides) while "strong" describes how badly it fails. Often used synonymously.

Near-degeneracies give rise to static correlation!

Potential energy curve of H_2

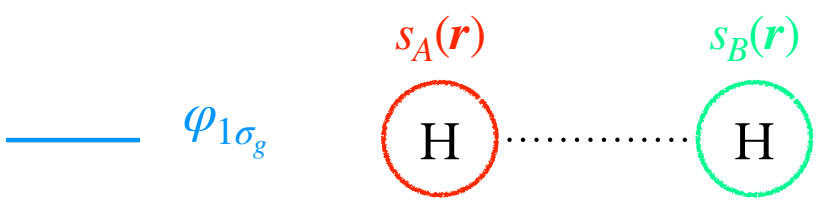


NB: "Static" usually describes why a single-reference picture fails (see following slides) while "strong" describes how badly it fails. Often used synonymously.

Near-degeneracies give rise to static correlation!

Better ansatz?

Multi-configurational wave function



Bonding orbital

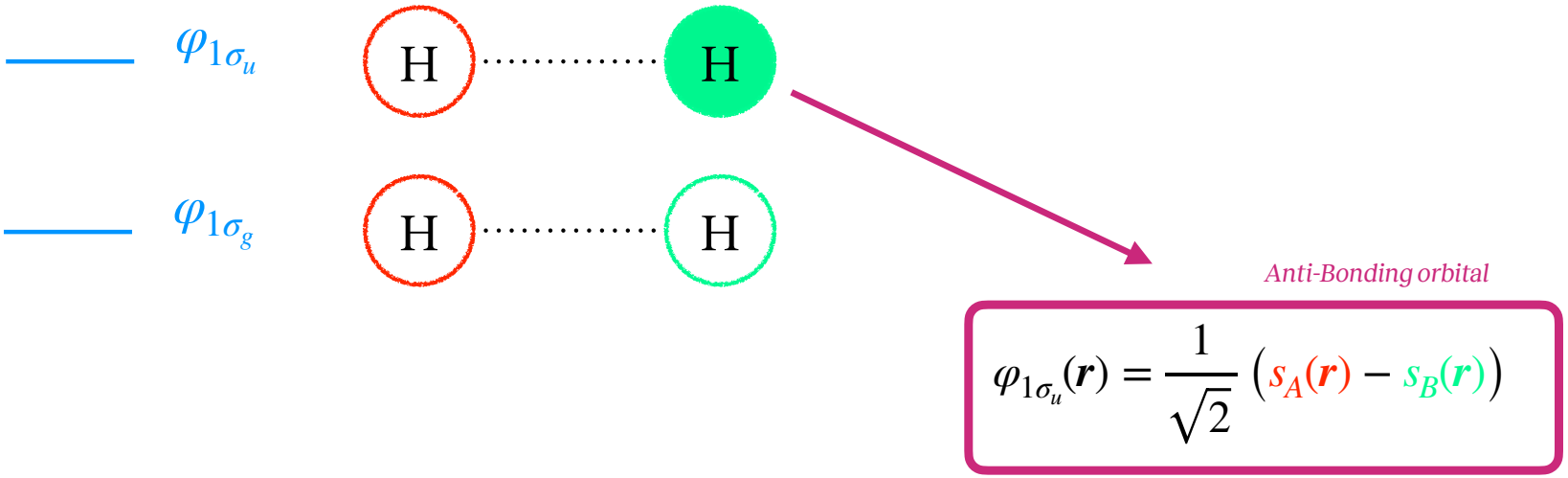
$$\varphi_{1\sigma_g}(\boldsymbol{r}) = \frac{1}{\sqrt{2}}\left(s_A(\boldsymbol{r}) + s_B(\boldsymbol{r})\right)$$

For simplicity, we assume an orthonormal AO basis

$$S = \left\langle s_A(\boldsymbol{r}) \middle| s_B(\boldsymbol{r}) \right\rangle = 0$$

and consider the spatial part only

Multi-configurational wave function

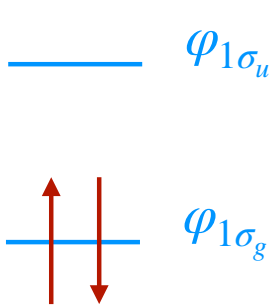


For simplicity, we assume an orthonormal AO basis

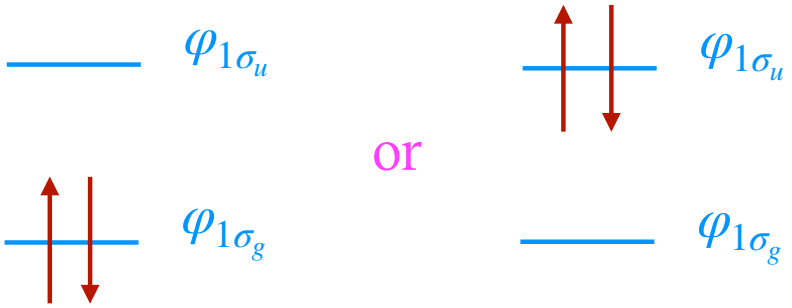
$$S = \left\langle s_A(\boldsymbol{r}) \middle| s_B(\boldsymbol{r}) \right\rangle = 0$$

and consider the spatial part only

Multi-configurational wave function



$$|\Psi\rangle = \left|\left(1\sigma_g\right)^2\right\rangle$$



$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left[\left|\left(1\sigma_g\right)^2\right\rangle - \left|\left(1\sigma_u\right)^2\right\rangle\right]$$

$$\equiv \frac{1}{\sqrt{2}}\left(\varphi_{1\sigma_g}(\boldsymbol{r}_1)\varphi_{1\sigma_g}(\boldsymbol{r}_2) - \varphi_{1\sigma_u}(\boldsymbol{r}_1)\varphi_{1\sigma_u}(\boldsymbol{r}_2)\right)$$

Multi-configurational wave function

$$\varphi_{1\sigma_u}(\boldsymbol{r}) = \frac{1}{\sqrt{2}}\left(s_A(\boldsymbol{r}) - s_B(\boldsymbol{r})\right)$$

$$|\Psi\rangle \equiv \frac{1}{\sqrt{2}}\left(\varphi_{1\sigma_g}(\boldsymbol{r}_1)\varphi_{1\sigma_g}(\boldsymbol{r}_2) - \varphi_{1\sigma_u}(\boldsymbol{r}_1)\varphi_{1\sigma_u}(\boldsymbol{r}_2)\right)$$

$$\varphi_{1\sigma_g}(\boldsymbol{r}) = \frac{1}{\sqrt{2}}\left(s_A(\boldsymbol{r}) + s_B(\boldsymbol{r})\right)$$

Multi-configurational wave function

$$|\Psi\rangle \equiv \frac{1}{\sqrt{2}}\left(\varphi_{1\sigma_g}(\boldsymbol{r}_1)\varphi_{1\sigma_g}(\boldsymbol{r}_2) - \varphi_{1\sigma_u}(\boldsymbol{r}_1)\varphi_{1\sigma_u}(\boldsymbol{r}_2)\right)$$

$$= \frac{1}{\sqrt{2}}\left(s_A(\boldsymbol{r}_1)s_B(\boldsymbol{r}_2) + s_A(\boldsymbol{r}_2)s_B(\boldsymbol{r}_1)\right)$$

$$\mathrm{H}\cdots\mathrm{H} \qquad \mathrm{H}\cdots\mathrm{H}$$

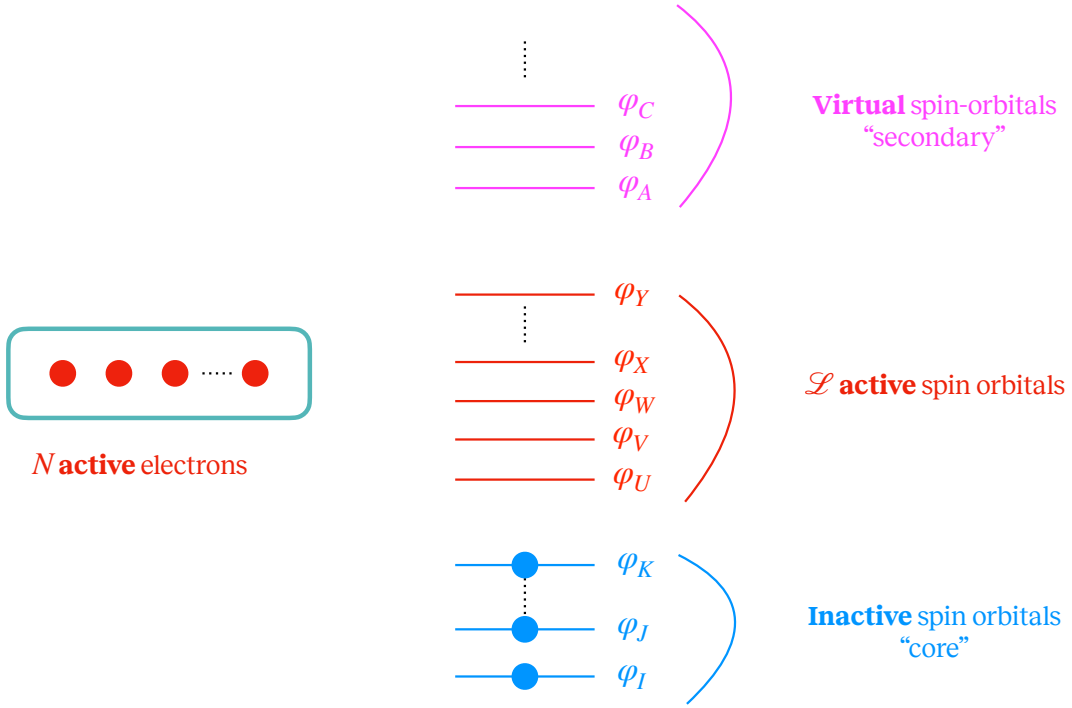
Multi-configurational wave function

$$\Phi_{(1\sigma_g)^2} \qquad \Phi_{(1\sigma_u)^2}$$

$$|\Psi\rangle \equiv \frac{1}{\sqrt{2}}\left(\varphi_{1\sigma_g}(\boldsymbol{r}_1)\varphi_{1\sigma_g}(\boldsymbol{r}_2) - \varphi_{1\sigma_u}(\boldsymbol{r}_1)\varphi_{1\sigma_u}(\boldsymbol{r}_2)\right)$$

The Past …

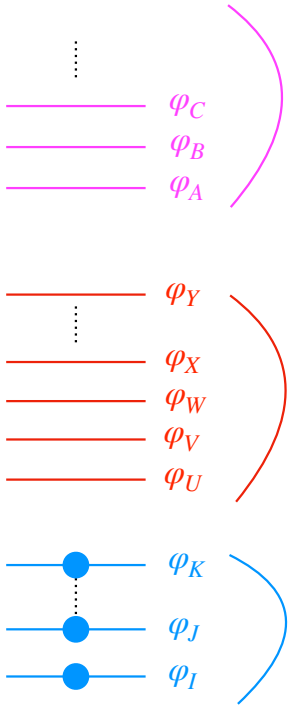
Complete Active Space CI (CASCI)



Complete Active Space CI (CASCI)

"CAS(N, L)"

N **active** electrons



Virtual spin-orbitals
"secondary"

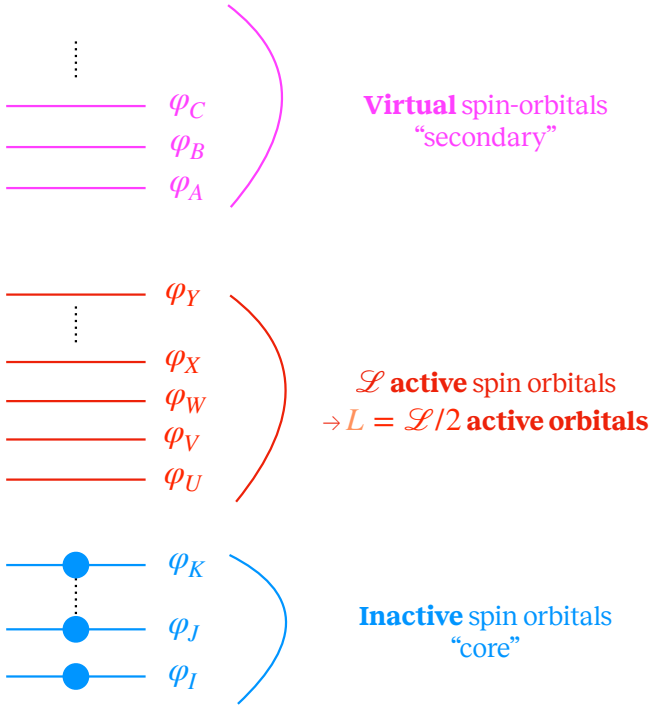
$\mathscr{L}$ **active** spin orbitals
$\rightarrow L = \mathscr{L}/2$ **active orbitals**

Inactive spin orbitals
"core"

Complete Active Space CI (CASCI)

"CAS(N, L)"

N **active** electrons



Virtual spin-orbitals
"secondary"

$\mathscr{L}$ **active** spin orbitals
$\rightarrow L = \mathscr{L}/2$ **active orbitals**

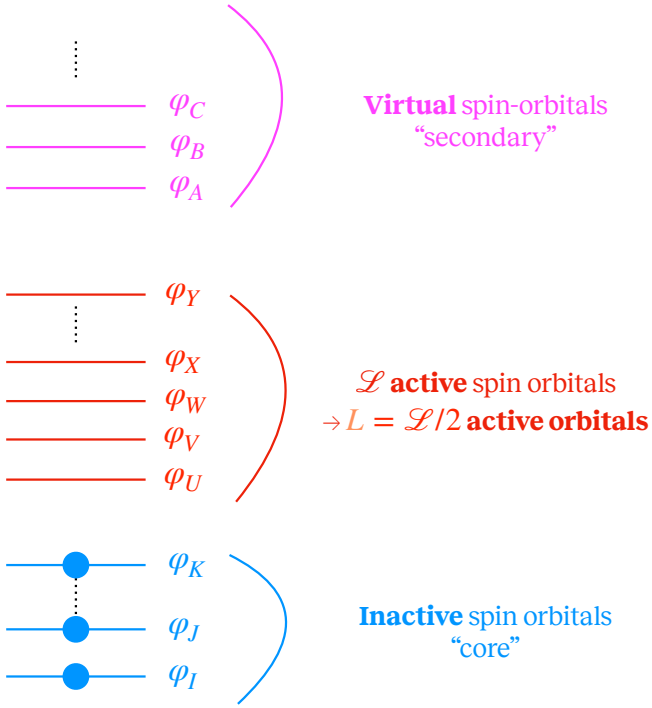
Inactive spin orbitals
"core"

CASCI is FCI in the active space!

Complete Active Space CI (CASCI)

"CAS(N, L)"

N **active** electrons



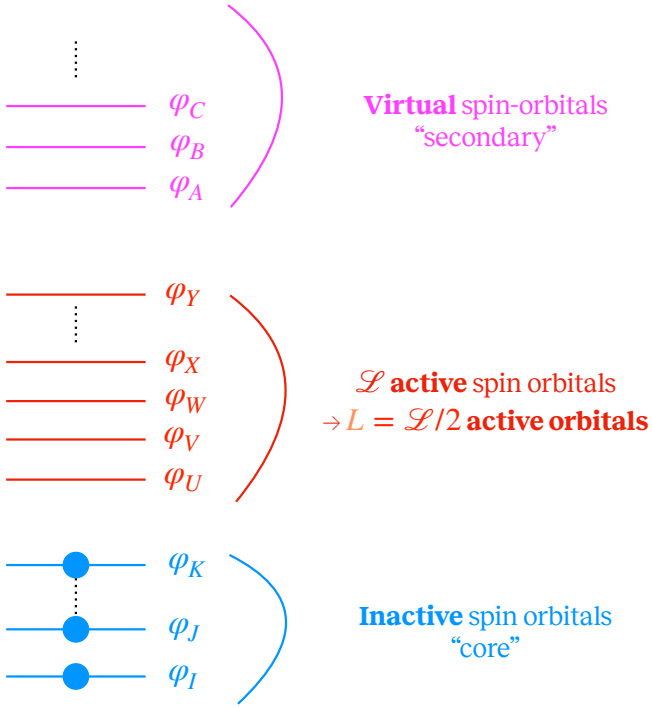
CASCI is FCI in the active space!

$$\left|\Psi^{\mathrm{CASCI}}\right\rangle = \sum_{\zeta \in \mathrm{CAS}} c_\zeta \left|\Phi_\zeta\right\rangle$$

Complete Active Space CI (CASCI)

"CAS(N, L)"

N **active** electrons

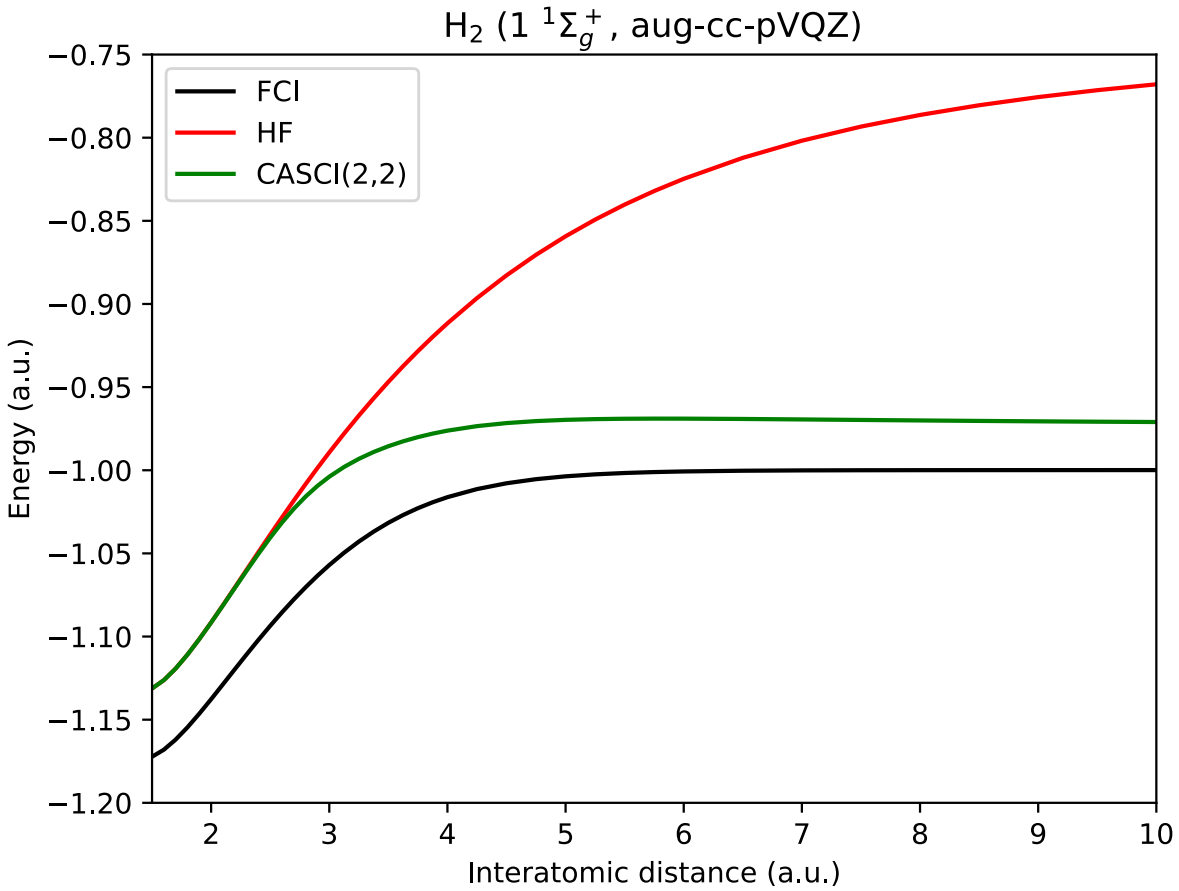


CASCI is FCI in the active space!

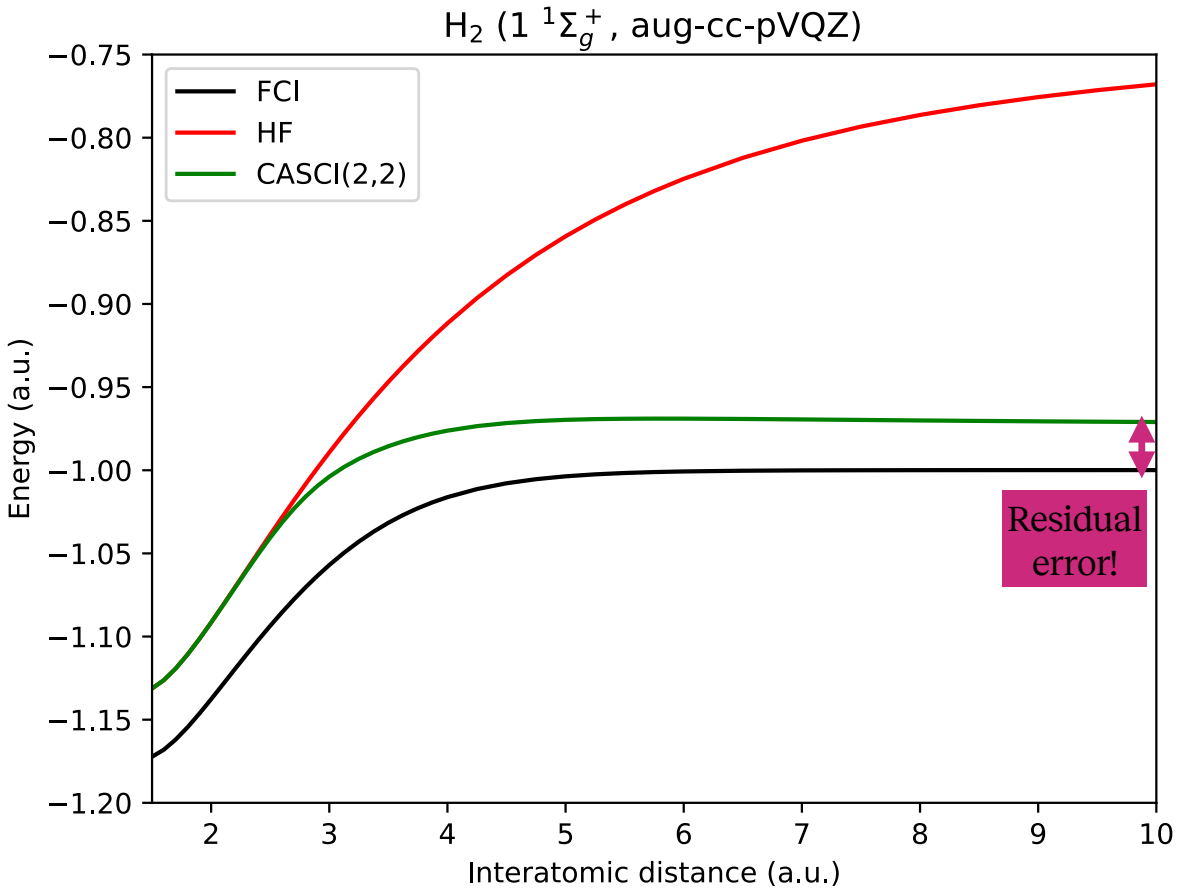
$$\left|\Psi^{\mathrm{CASCI}}\right\rangle = \sum_{\zeta \in \mathrm{CAS}} c_\zeta \left|\Phi_\zeta\right\rangle$$

How does it work for our H_2 example?

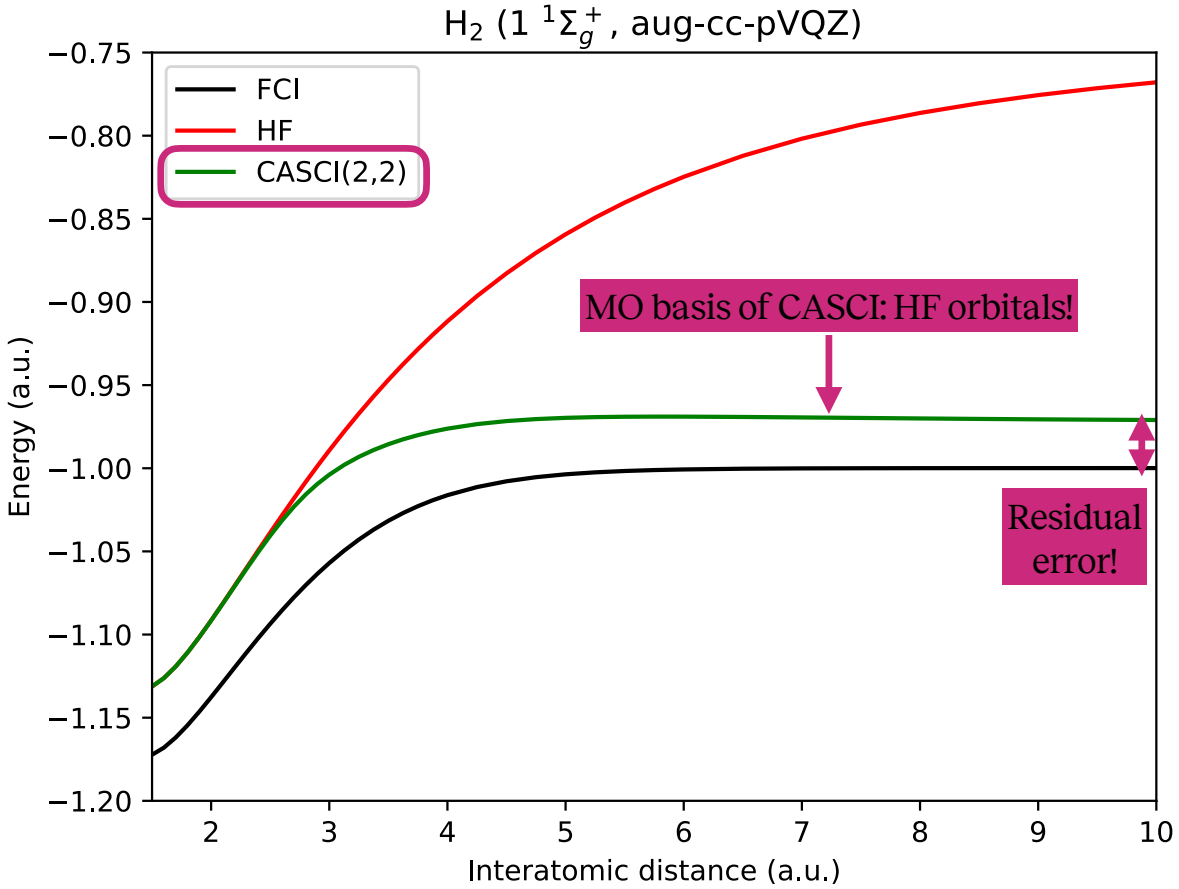
CASCI for H_2


H_2 ($1\ ^1\Sigma_g^+$, aug-cc-pVQZ)
FCI
HF
CASCI(2,2)
Energy (a.u.)
−0.75
−0.80
−0.85
−0.90
−0.95
−1.00
−1.05
−1.10
−1.15
−1.20
2
3
4
5
6
7
8
9
10
Interatomic distance (a.u.)

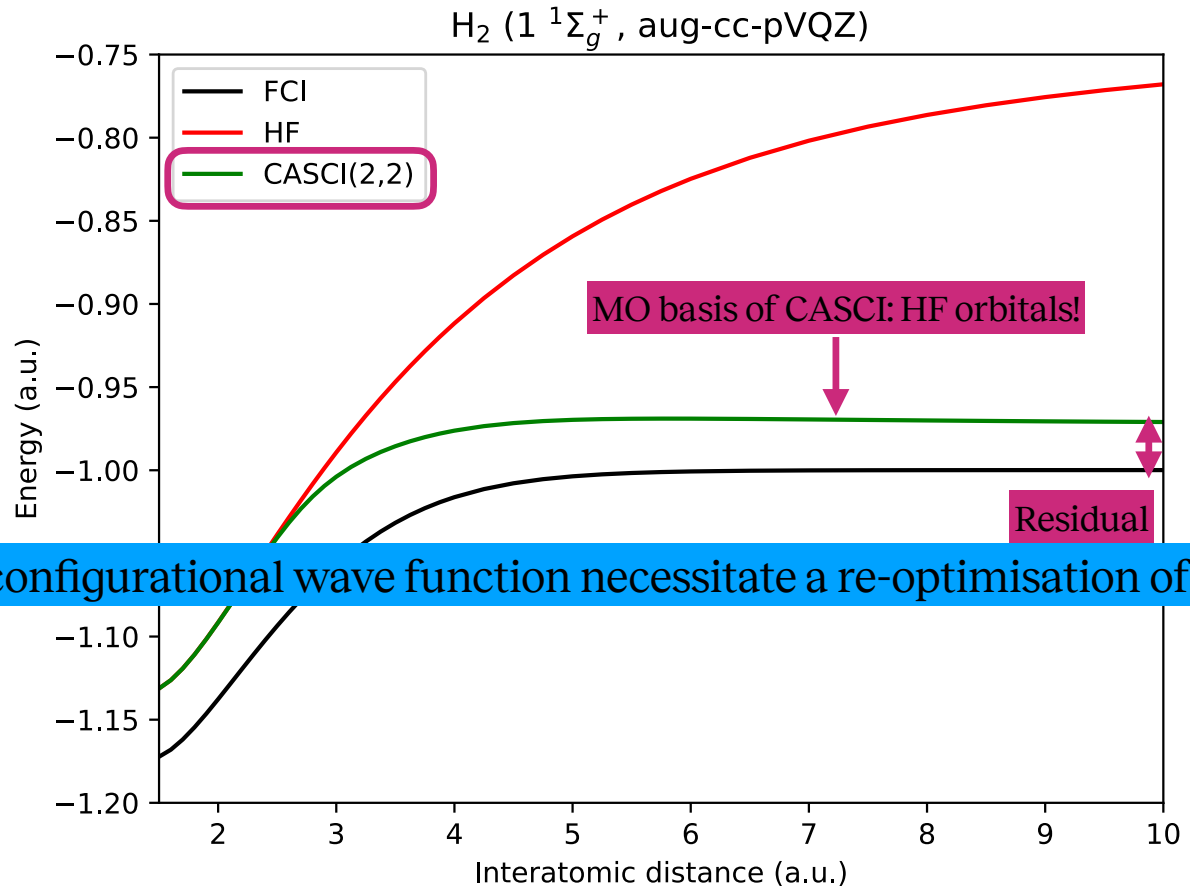
CASCI for H_2


H_2 ($1\ ^1\Sigma_g^+$, aug-cc-pVQZ)
FCI
HF
CASCI(2,2)
Residual error!
Energy (a.u.)
Interatomic distance (a.u.)
−0.75
−0.80
−0.85
−0.90
−0.95
−1.00
−1.05
−1.10
−1.15
−1.20
2
3
4
5
6
7
8
9
10

CASCI for H_2

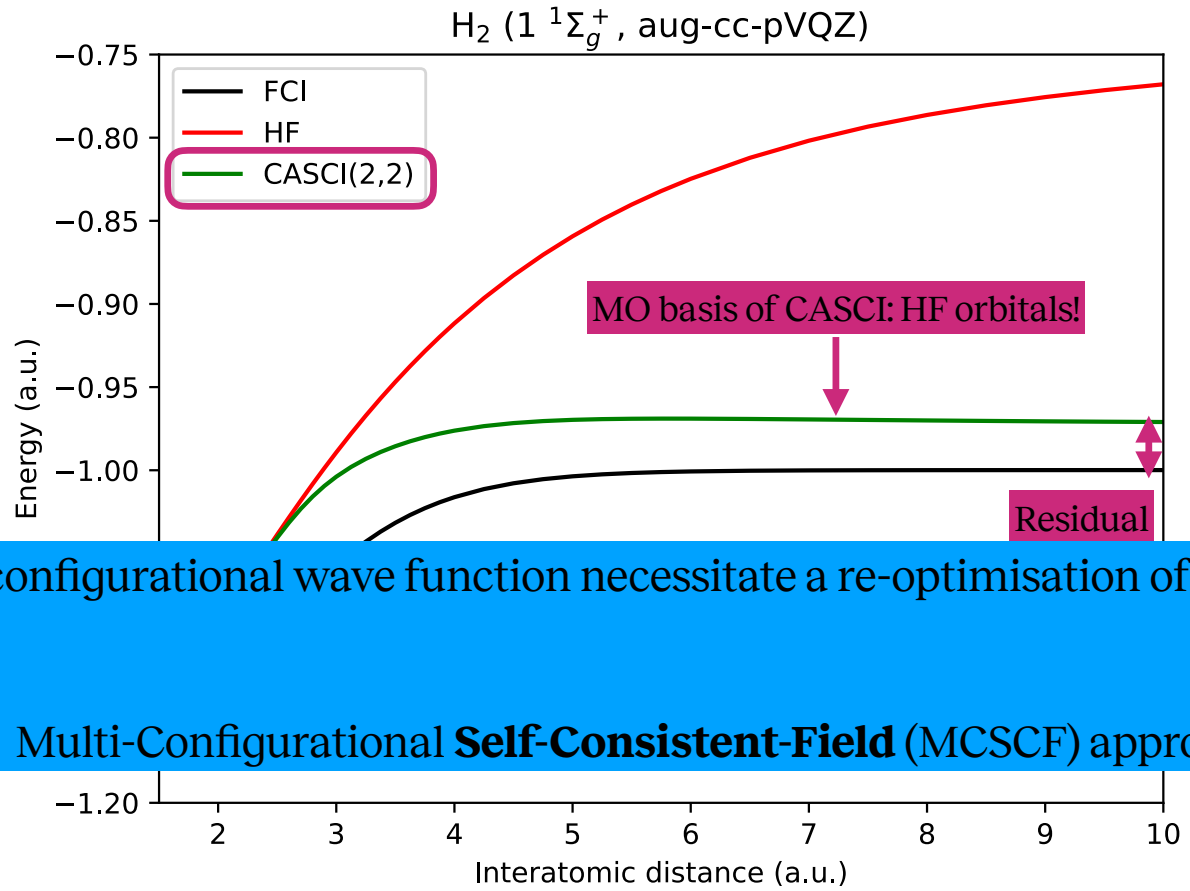

H_2 ($1\ ^1\Sigma_g^+$, aug-cc-pVQZ)
FCI
HF
CASCI(2,2)
MO basis of CASCI: HF orbitals!
Residual error!
Energy (a.u.)
Interatomic distance (a.u.)
−0.75
−0.80
−0.85
−0.90
−0.95
−1.00
−1.05
−1.10
−1.15
−1.20
2
3
4
5
6
7
8
9
10

CASCI for H_2



Multi-configurational wave function necessitate a re-optimisation of the orbitals!

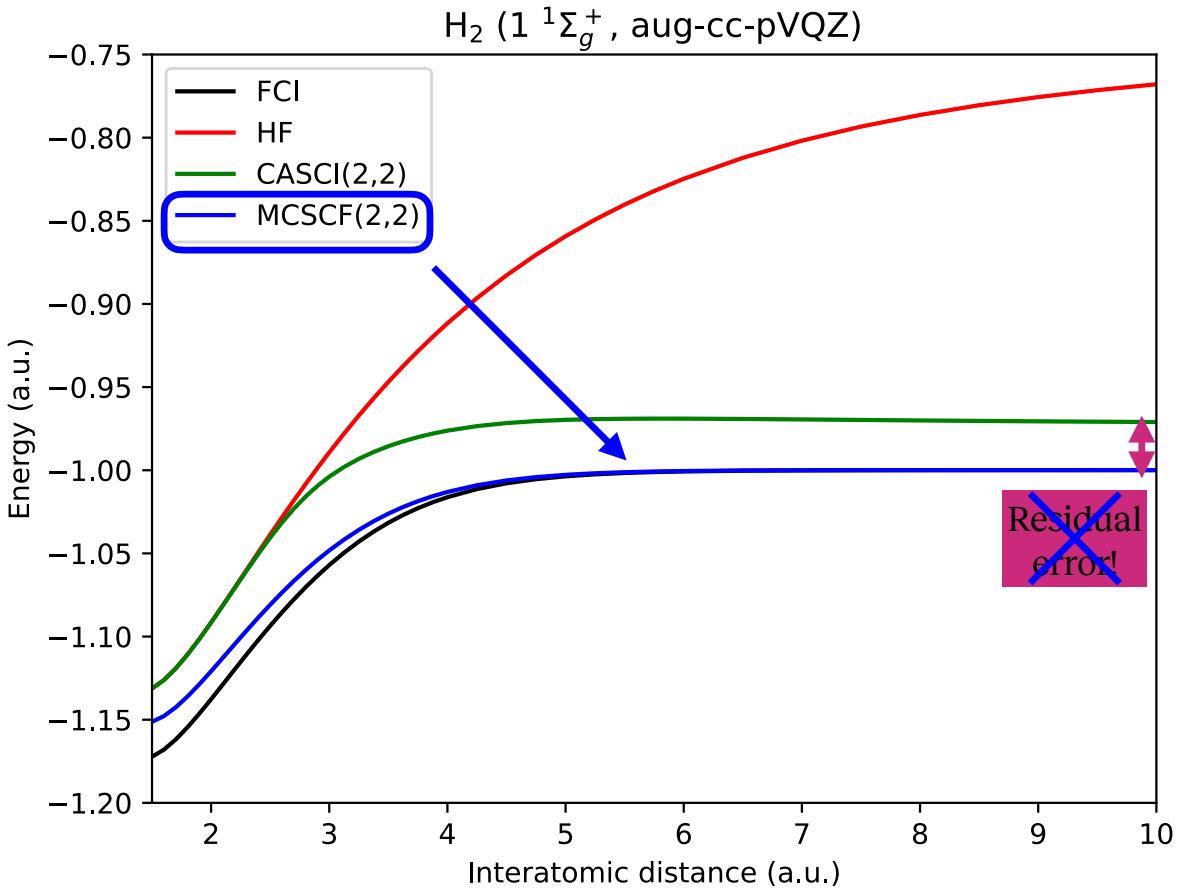
CASCI for H_2



Multi-configurational wave function necessitate a re-optimisation of the orbitals!

Multi-Configurational **Self-Consistent-Field** (MCSCF) approach

CASCI / MCSCF for H_2



MCSCF: concepts and purpose

MCSCF - concepts I

- Introduce as small number of (active) orbitals: $\approx 10 - 20$ with occupation numbers η allowed to vary
- Active orbitals with occupations: $0 \ll \eta \ll 2$
- Select configurations (many-particle basis states) to include
- Form of the wave function ($\left| \tilde{\Phi}_\zeta \right\rangle \rightarrow$ SD, ONVs or CSFs):

$$\left| \tilde{\Psi} \right\rangle = \sum_\zeta c_\zeta \left| \tilde{\Phi}_\zeta \right\rangle$$

- OPTIMISE the orbitals and the CI coefficients c_ζ

MCSCF - concepts II

- Simplify general MCSCF ansatz by including **all** configurations generated by allocating all **active electrons** to these **active orbitals**:
  FCI in the active space (↔ **CASCI**!)

- Picking a proper set of active orbitals is still required (can be automated too) but configuration generation is automated

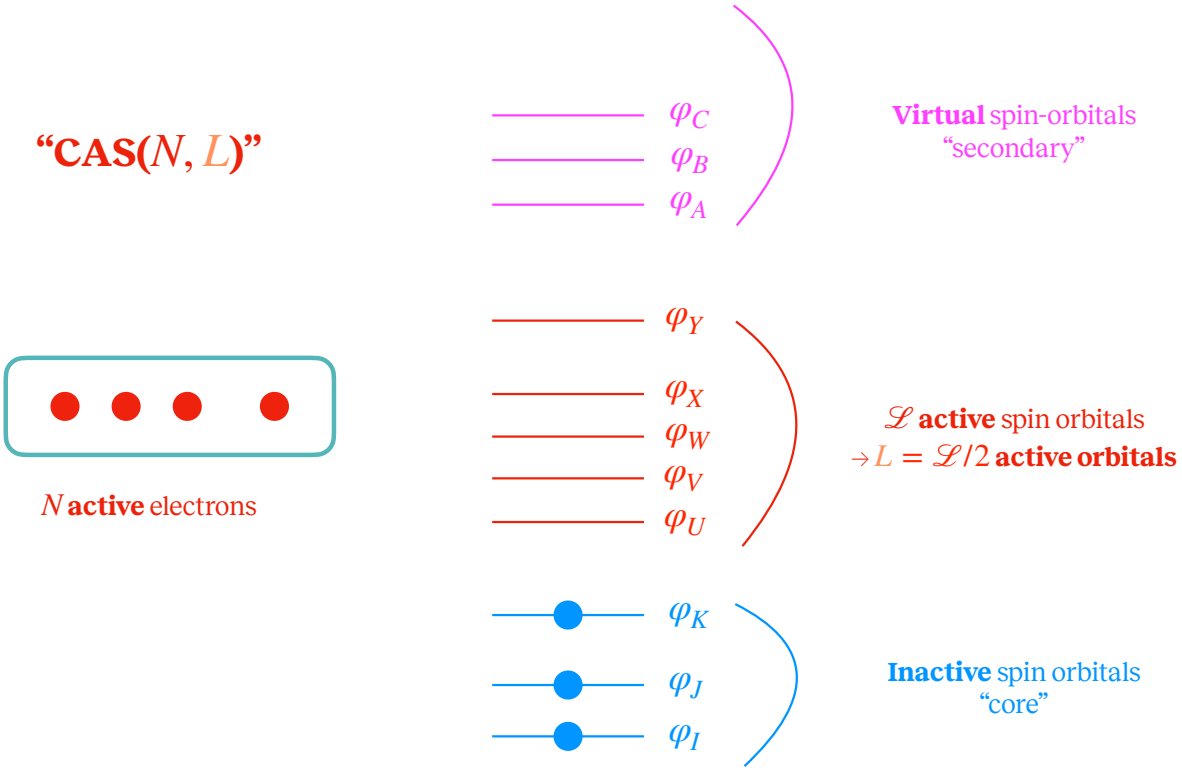
- MCSCF optimisation based on this simplification is coined as
  complete active space self-consistent field — or simply **CASSCF**

MCSCF - purpose

- MCSCF **<u>does not</u> describe** the short-range correlation contributions that arise as $r_{12} \rightarrow 0$, that is *dynamical correlation*
  - MCSCF aims at including *non-dynamical correlation* that arises from
    (i) configurational near-degeneracies and/or
    (ii) gross deficiencies in the RHF wave function
- Includes near-degenerate orbitals to account for **static correlation**
- **Will in general not describe the complete correlation energy!**
- **NB: There is no rigorous physical decomposition of** $E_{corr} = E_{static} + E_{dynamic}$

CASSCF: concepts, optimisation and limitations

CASSCF - concepts

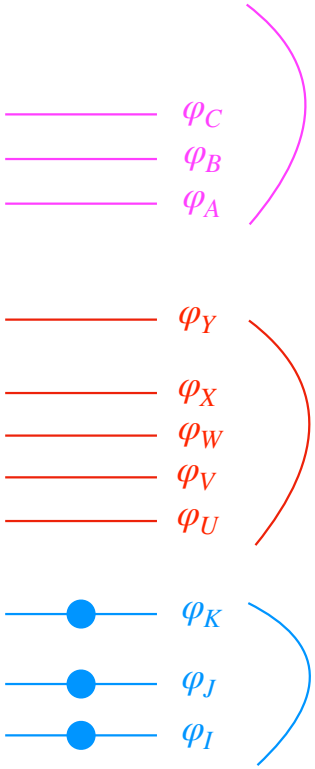


- Definition of the orbital spaces follows from CASCI: inactive active and secondary
- The active orbital space should include:
  - all orbitals where the occupation number η changes significantly during a process (reaction, excitation, ionisation, ...)
  - orbitals where $0 \ll \eta \ll 2$
  - can be **automated** (overlap, orbital entropy, perturbative estimates, ...) but do not underestimate "chemical intuition"

CASSCF - concepts

"CAS(N, L)"

N **active** electrons



Virtual spin-orbitals
"secondary"

$\mathscr{L}$ **active** spin orbitals
$\rightarrow L = \mathscr{L}/2$ **active orbitals**

Inactive spin orbitals
"core"

AUTOCAS: A Program for Fully Automated Multiconfigurational Calculations

Christopher J. Stein, Markus Reiher

First published: 07 June 2019 | https://doi.org/10.1002/jcc.25869 | VIEW METRICS

HQS QUANTUM SIMULATIONS

Simplifying the Selection of Active Spaces in Quantum Chemistry with the ASF

Active Space Finder

The Active Space Finder (ASF) is a set of functions for the (semi-)automatic selection of active spaces in molecules to be employed with methods such as CASSCF. Choosing an appropriate set of active orbitals can be a complicated task, which requires a significant amount of expertise. We have set out to develop a tool to make such calculations easier and more accessible to both expert and non-expert users.

Employing early-stage quantum devices, in the so-called Noisy-Intermediate-Scale-Quantum (NISQ) era, for applications in quantum chemistry requires the majority of problems to be reduced to their most important degrees of freedom. The ASF sets out to determine those parts of a molecular system that need to be treated at the most advanced quantum mechanical level.

Get the ASF: GitHub

[Submitted on 14 Aug 2025 (v1), last revised 15 Jul 2026 (this version, v3)]

AEGISS -- Atomic orbital and Entropy-based Guided Inference for Space Selection -- A novel semi-automated active space selection workflow for quantum chemistry and quantum computing applications

Fabio Tarocco, Pi A. B. Haase, Fabijan Pavošević, Viiav Krishna. Leonardo Guidoni. Stefan Knecht. Martina Stella

Subjects: **Chemical Physics**
Cite as: arXiv:2508.10671

RLEASE: Reinforcement Learning Efficient Active Space Engine

Etinosa Osaro, Abhishek Mitra, Andrew J. Jenkins, Kelsey A. Parker, Robert H. Lavroff, Verena A. Neufeld, Arpan Kundu, Arvin Kakekhani, Dario Rocca

Subjects: **Chemical Physics (physics.chem-ph)**
Cite as: arXiv:2606.07879 **[physics.chem-ph]**

CASSCF - scaling

2	4
4	36
6	400
8	4.900
10	63.504
12	853.776
14	11.778.896
16	165.636.896
18	2.363.904.260
20	34.134.777.856
24	7.312.459.672.336

- The number of Slater determinant (or ONVs) $\{\Phi\}$ scales for $2k$ electrons in $2k$ orbitals as

$$\{\Phi\} = \binom{2k}{k}^2$$

- Largest calculation with standard CI: CAS(24,24), practical limit about (20,20)!

- Larger CAS spaces require more advanced approaches (→ see The Present!)

CASSCF - active orbital spaces

Simple diatomics

- Sometimes simple and intuitive like for H2: $(\sigma_g, \sigma_u)^2 \rightarrow$ CAS(2,2)
- Ground state of N2 requires 2p orbitals: $(\sigma_g, \pi_u, \pi_g, \sigma_u)^6 \rightarrow$ CAS(6,6)
- C2/Be2: require inclusion of 2s/2p shells because of near-degeneracies!
- Picking CAS for main-group dimers can be straightforward... but does not necessarily have to be the case
- Transition metal dimers are far from trivial! ($\rightarrow$ see “The Present”)

CASSCF - active orbital spaces

Polyatomic molecules

- In general, including the full valence space is not an option (too many e^-/o)
- Simple guidelines:
  - Breaking a C-H or C-C bond in a hydrocarbon → include $(\sigma, \sigma^*)^2$
  - Spectroscopy/reaction of aromatic/conjugated π-systems → include (π, π^*)
- If even the minimal CAS reaches beyond (20,20), consider alternative approaches like DMRG, HCI, FCIQMC, selective CI, ...

CASSCF - targeting individual states

State-specific approach

- Goal: target individual excited states which are not the lowest states wrt spin and/or spatial symmetry
- Challenges:
  - Requires convergence of optimisation algorithm to a saddle point
  - Root flipping: excited state may become the lowest state in CI along a path
  - Converged MCSCF/CASSCF wave functions for two roots of the same symmetry (spatial/spin) are in general **NOT** orthogonal!
    → use *state-interaction* to calculate properties: SOC, NAC, ...

CASSCF - targeting an ensemble of states

State-average approach

- Goal: target an ensemble of states **simultaneously** wrt spin and/or spatial symmetry
  - Introduce a weighted ensemble $\{M\}_\omega$ of the energies of M states:

$$E^{SA} = \sum_{i=1}^{M} \omega_i E_i$$

  - Each state in $\{M\}_\omega$ will have identical MOs but different CI coefficients
- NB: Wavefunction is no longer fully variational: properties and gradients are more complicated to compute as compared to for SS-CASSCF.
- Challenge: MOs in different states may be very different
  → may require larger CAS to ensure smooth convergence

CASSCF - targeting an ensemble of states

State-average approach

- Example: consider an SA CAS(2,2)SCF calculation of ethylene (C_2H_4) for the singlet (S=0) ground ("*N* state") and lowest excited state ("*V* state"):
  - CAS: $(\pi_u, \pi_g)^2$
  - Character of *N* state: $|\Psi_0> = c_1(\pi_u)^2 + c_2(\pi_g)^2 + \ldots$
  - Character of *V* state: $|\Psi_1> = (\pi_u \pi_g) + \ldots$
- *N* and *V* state → different spatial extents ($< z^2 >$ = 1.7 / 9.1)
- Cannot be described by a single set of π orbitals!

CASSCF - Optimisation of the wave function

Concepts

- Wave function *ansatz* (note change of notation $\left|\tilde{0}\right\rangle \equiv \left|\tilde{\Psi}\right\rangle$:

$$\left|\tilde{0}\right\rangle = \sum_{\zeta} c_{\zeta} \left|\tilde{\Phi}_{\zeta}\right\rangle$$

Determine the MO and CI coefficients using the variational principle (normalisation!):

$$\delta E = \delta \left(\frac{\left\langle \tilde{0} \right| \hat{H} \left| \tilde{0} \right\rangle}{\left\langle \tilde{0} \middle| \tilde{0} \right\rangle} \right) = 0$$

Indices: $i, j, k, \ldots$ $v, w, x, \ldots$ $a, b, c, \ldots$

CASSCF - Optimisation of the wave function

Energy

- Total CASSCF energy: $$E = \left\langle \tilde{0} \middle| \hat{H} \middle| \tilde{0} \right\rangle$$
$$= \sum_{pq} h_{pq} D_{pq} + \sum_{pqrs} g_{pqrs} P_{pqrs} + h_{nuc}$$

  with

  $D_{pq} = \left\langle \tilde{0} \middle| \hat{E}_{pq} \middle| \tilde{0} \right\rangle$ (first-order RDM)

  $P_{pqrs} = \frac{1}{2} \left\langle \tilde{0} \middle| \hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps} \middle| \tilde{0} \right\rangle$ (second-order RDM)

- MO coefficients $\{\tilde{\varphi}\}$ appear in h_{pq} and g_{pqrs}, CI coefficients $\{\tilde{\Phi}\}$ in $\mathbf{D}$ and $\mathbf{P}$

CASSCF - Optimisation of the wave function

Unitary transformations of MOs

- MOs $\{\tilde{\varphi}\}$ are orthonormal → unitary transformation $\boldsymbol{U}$ ensures orthonormality

$$\tilde{\varphi} = \varphi \boldsymbol{U} \text{ with } \boldsymbol{U}^\dagger \boldsymbol{U} = \mathbf{1}$$

- Write $\boldsymbol{U}$ as $\boldsymbol{U} = \exp(\boldsymbol{T})$ with $\boldsymbol{T}^\dagger = -\boldsymbol{T}$, that is $\boldsymbol{T}$ is anti-hermitian

$$\hat{T} = \sum_{pq} T_{pq} \hat{E}_{pq} = \sum_{p>q} T_{pq} (\hat{E}_{pq} - \hat{E}_{qp})$$

- Transformation of creation operators:

$$\hat{a}_{\tilde{p}} = \exp(\hat{T}) \hat{a}_p^\dagger \exp(-\hat{T})$$

- CAS → we only need the following rotation parameters: $\{pq\}$, $\{pq\}$, $\{pq\}$

CASSCF - Optimisation of the wave function

Unitary transformations of MOs

- MOs $\{\tilde{\varphi}\}$ are orthonormal → unitary transformation $\boldsymbol{U}$ ensures orthonormality

$$\tilde{\varphi} = \varphi \boldsymbol{U} \text{ with } \boldsymbol{U}^\dagger \boldsymbol{U} = \mathbf{1}$$

- Write $\boldsymbol{U}$ as $\boldsymbol{U} = \exp(\boldsymbol{T})$ with $\boldsymbol{T}^\dagger = -\boldsymbol{T}$, that is $\boldsymbol{T}$ is anti-hermitian

$$\hat{T} = \sum_{pq} T_{pq} \hat{E}_{pq} = \sum_{p>q} T_{pq} (\hat{E}_{pq} - \hat{E}_{qp})$$

- Transformation of creation operators:

$$\hat{a}_{\tilde{p}} = \exp(\hat{T}) \hat{a}_p^\dagger \exp(-\hat{T})$$

- CAS → we only need the following rotation parameters: $\{pq\}$, $\{pq\}$, $\{pq\}$

NB:
no intraspace
rotations

CASSCF - Optimisation of the wave function

Unitary transformations of CI vectors

- CI vector(s) $\{\Phi\}$ are normalised:

$$\left|0\right\rangle = \sum_{\zeta} c_{\zeta} \left|\Phi_{\zeta}\right\rangle \text{ with } \sum_{\zeta} |c_{\zeta}|^2 = 1$$

- Complementary space $\left|K\right\rangle$ orthogonal to $\left|0\right\rangle$: $\left\langle 0 \middle| K \right\rangle = 0$

- Define anti-hermitian operator $\hat{S}^{\dagger} = -\hat{S}$:

$$\hat{S} = \sum_{K \neq 0} S_{K0} \left(\left|K\right\rangle \left\langle 0\right| - \left|0\right\rangle \left\langle K\right| \right)$$

- Unitary transformation of $\left|0\right\rangle$ such that $\left|\tilde{0}\right\rangle$ remains normalised

$$\left|\tilde{0}\right\rangle = \exp(\hat{S}) \left|0\right\rangle$$

CASSCF - Optimisation of the wave function

Unitary transformations of CI vectors

- Note that the $|K\rangle$'s are mainly a **formal device.** The apparently huge sum over all complementary CI states only defines a single tangent direction in CI space. In fact, the state optimization is just a two-dimensional unitary rotation between the current CI vector and an orthogonal update vector $|\Delta\rangle$.
- Define $|\Delta\rangle = \sum_{K\neq 0} S_{K0} |K\rangle$
- then, by construction, we have $\langle 0|\Delta\rangle = 0 \rightarrow \hat{S} = |\Delta\rangle\langle 0| - |0\rangle\langle\Delta|$
- Hence: $|\tilde{0}\rangle = \exp(\hat{S})|0\rangle = \cos\|\Delta\| \, |0\rangle + \frac{\sin\|\Delta\|}{\|\Delta\|} |\Delta\rangle$ with $\|\Delta\| = \sqrt{\langle\Delta|\Delta\rangle}$
- The exponential of an anti-Hermitian CI variation produces an ordinary rotation on the unit sphere of CI vectors → preservation of normalisation: $\langle\tilde{0}|\tilde{0}\rangle = \cos^2\|\Delta\| + \sin^2\|\Delta\| = 1$

CASSCF - Optimisation of the wave function

Concepts

- Resulting wave function ansatz:

$$\left|\tilde{0}\right\rangle = \exp(\hat{T})\exp(\hat{S})\left|0\right\rangle$$

- "Double"-exponential parametrisation with parameters T_{pq} $(p > q)$ and S_{K0}

- Energy function with parameter space $\{\mathbf{p}\} \equiv \{\{T_{pq}\}_{p>q}, \{S_{K0}\}\}$

$$E(\boldsymbol{p}) = E(\boldsymbol{T}, \boldsymbol{S}) = \left\langle\tilde{0}\right| \hat{H} \left|\tilde{0}\right\rangle = \langle 0| \exp(-\hat{S})\exp(-\hat{T})\hat{H}\exp(\hat{T})\exp(\hat{S})\left|0\right\rangle$$

CASSCF - Optimisation of the wave function

Optimal energy

- Vary parameters $\boldsymbol{T}$ and $\boldsymbol{S}$ such that energy $E(\boldsymbol{T},\boldsymbol{S}) = \langle 0 \left| e^{-\hat{S}} e^{-\hat{T}} \hat{H} e^{\hat{T}} e^{\hat{S}} \right| 0 \rangle$ becomes stationary:

$$\frac{\partial E}{\partial T_{pq}} = 0 \quad \frac{\partial E}{\partial S_{K0}} = 0$$

- Leads to a set of nonlinear equations that must be solved iteratively
- In the following, we examine the Newton-Raphson method for CASSCF optimisation

CASSCF - Optimisation of the wave function

Newton-Raphson method - concepts I

- Find stationary point of E($\mathbf{p}$) where $\mathbf{p}$ is a set of parameters to be freely varied
- Requires start guess, for example with $\mathbf{p}_0 = 0$

  Expand E through second-order around this point

$$E(\mathbf{p}) \approx E^{(2)} = E(0) + \sum_i \left(\frac{\partial E}{\partial p_i}\right)_0 p_i + \frac{1}{2}\sum_{ij} p_i \left(\frac{\partial^2 E}{\partial p_i \partial p_j}\right)_0 p_j$$

$$= E(0) + \mathbf{g}^\dagger \mathbf{p} + \frac{1}{2}\mathbf{p}^\dagger \mathbf{H} \mathbf{p}$$

- $\mathbf{g}$ is the gradient vector & $\mathbf{H}$ is the Hessian matrix

CASSCF - Optimisation of the wave function

Newton-Raphson method - concepts II

- Approximation to stationary point by finding stationary point of $E^{(2)}$:

$$\frac{\partial E^{(2)}}{\partial p_i} = 0 \;\rightarrow\; \mathbf{g} + \mathbf{H}\mathbf{p} = \mathbf{0} \rightarrow \;\; \mathbf{p} = -\mathbf{H}^{-1}\mathbf{g}$$

- For the current $\mathbf{p}$, compute new gradient and Hessian
- Continue iterative procedure until $|\mathbf{g}| \approx 0$
- **Note:**
  - Approximating E with $E^{(2)}$ is only valid for small $\mathbf{p}$
  - Quadratic convergence for small $\mathbf{p}$

CASSCF - Optimisation of the wave function

Taylor expansion of the energy

- **Recall:** BCH expansion for operators

$$\exp(-\hat{A})\hat{B}\exp(\hat{A}) = \hat{B} + \left[\hat{B},\hat{A}\right] + \frac{1}{2}\left[\left[\hat{B},\hat{A}\right],\hat{A}\right] + \ldots$$

- Expand E through second order in $\hat{T}, \hat{S}$ with $E(0)$, $\mathbf{g}$ and $\mathbf{H}$ parts

$$E^{(2)}(\mathbf{T},\mathbf{S}) = \left\langle 0 \left| \hat{H} \right| 0 \right\rangle + \left\langle 0 \left| \left[\hat{H},\hat{T}\right] + \left[\hat{H},\hat{S}\right] \right| 0 \right\rangle$$

$$+ \left\langle 0 \left| \frac{1}{2}\left[\left[\hat{H},\hat{T}\right],\hat{T}\right] + \frac{1}{2}\left[\left[\hat{H},\hat{S}\right],\hat{S}\right] + \left[\left[\hat{H},\hat{T}\right],\hat{S}\right] \right| 0 \right\rangle$$

CASSCF - Optimisation of the wave function

Newton-Raphson equations

- After some straightforward math, Newton-Raphson equations assume the form

$$\begin{pmatrix} \mathbf{H}^{cc} & \mathbf{H}^{co} \\ \mathbf{H}^{oc} & \mathbf{H}^{oo} \end{pmatrix} \begin{pmatrix} \mathbf{S} \\ \mathbf{T} \end{pmatrix} = - \begin{pmatrix} \mathbf{g}^{c} \\ \mathbf{g}^{o} \end{pmatrix}$$

  comprising orbital (o) and configurational (c) parts as well as mixed terms

- Common approximations:
  - Neglect oc/co parts in H
  - Approximate $H^{oo} \rightarrow$ super-CI approach

CASSCF - Optimisation of the wave function

Code availability and capabilities

- DALTON: NR and other optimization techniques → state-specific MCSCF
- OpenMolcas: no *co*/*oc* couplings → state-specific + state-average MCSCF
- MOLPRO: second-order optimisation → state-specific + state-average MCSCF
- ORCA: second-order optimisation → state-specific + state-average MCSCF
- pySCF: quasi-second order → state-specific + state-average MCSCF
- VeloxChem/MultiPsi: second-order opt. → state-specific + state-average MCSCF
- ...

CASSCF - Final thoughts

CASSCF is just a start for applying dynamic correlation!

Relative energies and molecular properties at the CASSCF level **should not** be compared to experimental results.

Dynamical electron correlation combined with static correlation

Dynamical correlation combined with MCSCF

Wishlist

- GOAL: treat dynamical correlation in combination with MCSCF wave function
  - simultaneously (**diagonalize-and-perturb**)
  - *a posteriori* (**diagonalize-then-perturb**)
- Should preferably be both size-extensive and size-consistent
  - size-extensive: energy scales linearly with number of particles N
  - size-consistent: $E_{AB} = E_A + E_B$ at $r(A - B) \rightarrow \infty$
- Should allow to treat an ensemble of states on an equal footing

Dynamical correlation combined with MCSCF

A not-so-complete summary ...

- **Multi-reference CI:**
  - Based on excitations out of a MC state
  - Variational but not size-consistent/size-extensive
- **Multi-reference perturbation (MRPT2) theories**: CASPT2, NEVPT2, ...
  - Based on (internally contracted) excitations out of a MC state
  - Differ in form of $\hat{H}_0$ and form of wave function corrections
  - (Nearly) size-extensive
  - MS-CASPT2 is a (diagonalize-then-perturb-then-diagonalize) approach, a hybrid between the two approaches mentioned on the previous slide
- **MRCC**: less developed but most rigorous!
- **post-MCSCF on-top pair-DFT correction** (MC-pDFT)
  - Requires specialised DFAs
  - Works for SS and SA MC reference wave functions
- **srDFT-lrMCSCF**:
  - Requires specialised DFAs
  - Allows a simultaneous treatment of static and dynamic correlation