Neural network assisted calculations of electronic wave functions for molecules

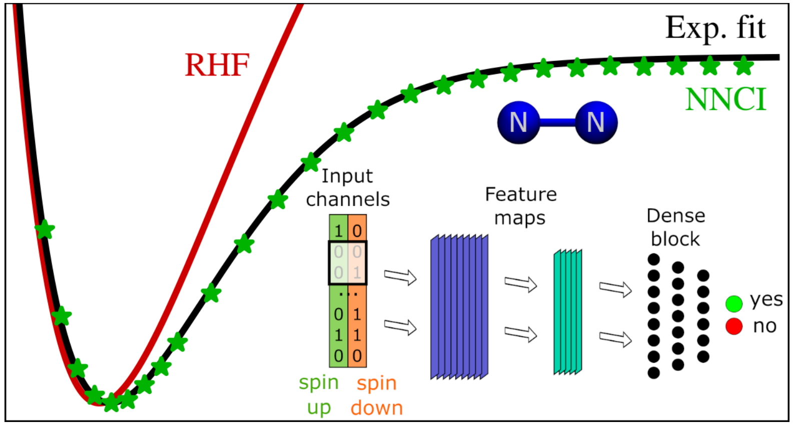
Overview:

A neural network is used to select Slater determinants (SDs) in configuration interaction (CI) calculations, **NNCI**.

A convolutional neural network is trained by carrying out CI calculations on trial sets of SD.

Calculations of N_2, H_2O, NH_3, CO and propane give larger correlation energy than full CI calculations using 4 to 6 *orders of magnitude* more SDs.

The dissociation curve of N_2 is in quite good agreement with a fit to experimental values, but the bond energy is still underestimated by 3% when the calculations are carried out with 52 molecular orbitals.



Y.L.A. Schmerwitz, L. Thirion, G. Levi *et al.*,
J. Chem. Theory Comput. (2025)

Reykjavík, Iceland

HÁSKÓLI ÍSLANDS


Dr. Yorick L. A. Schmerwitz


Assoc. Prof. Gianluca Levi


Assoc. Res. Prof. Elvar Ö. Jónsson


Prof. Hannes Jónsson

Erlangen, Germany

FAU


Louis Thirion


Prof. Philipp Hansmann

MAX PLANCK INSTITUTE
FOR THE SCIENCE OF LIGHT

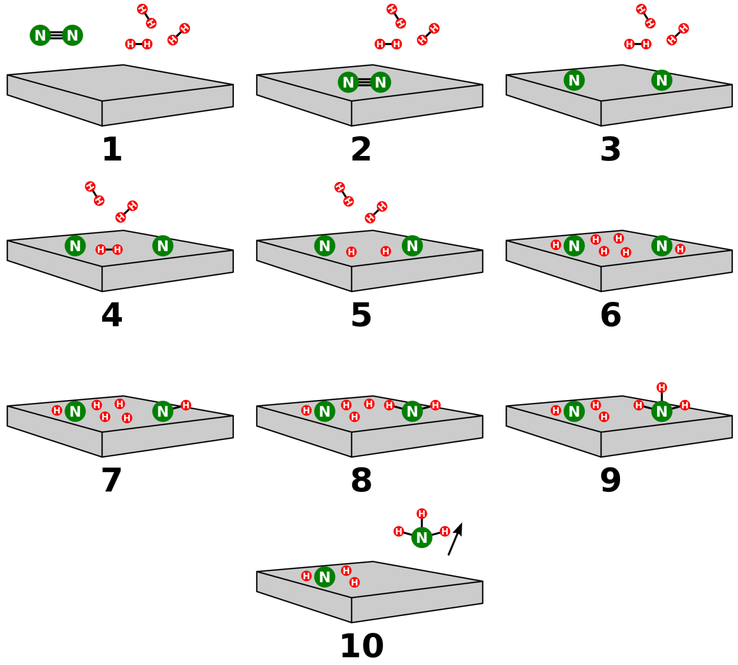

Max Kroesbergen


Dr. Pavlo Bilous

Where We Are Heading

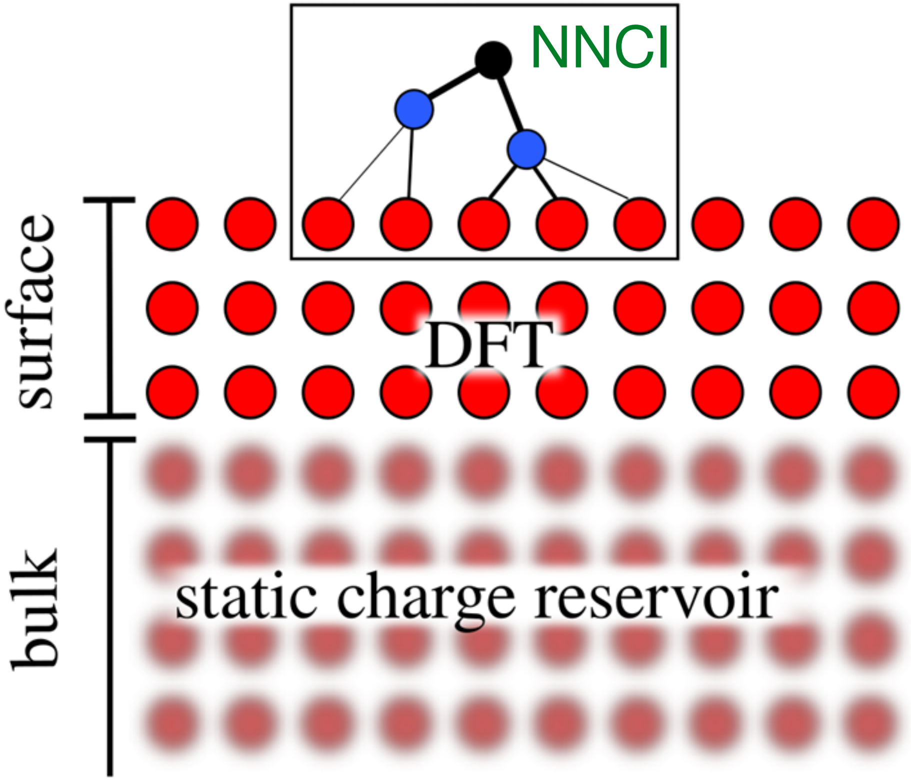
Challenge 1:

Heterogeneous catalysis calculations, improved accuracy greatly needed



Find mechanism and estimate rate of elementary steps in catalytic rxns.

PBE functional overestimates N-N bond energy by 0.6 eV



Embedding approach:

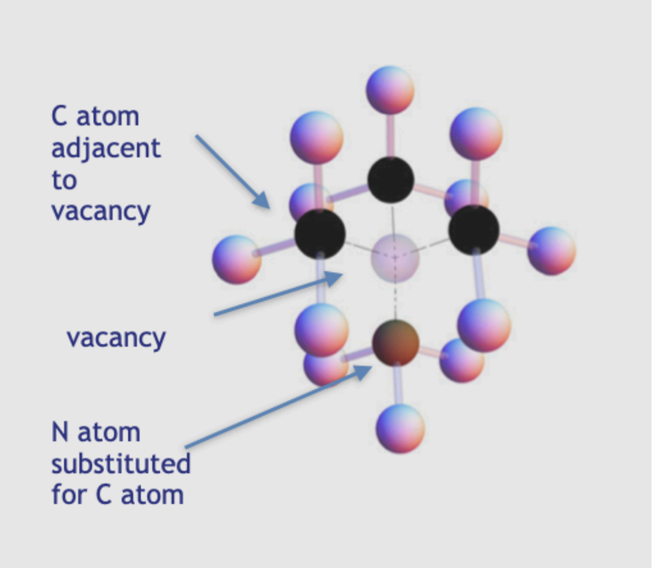
NNCI calculation of reactive site and reactant(s) embedded in a slab described with DFT which in turn is coupled to an extended solid with a Green function approach.

[1]

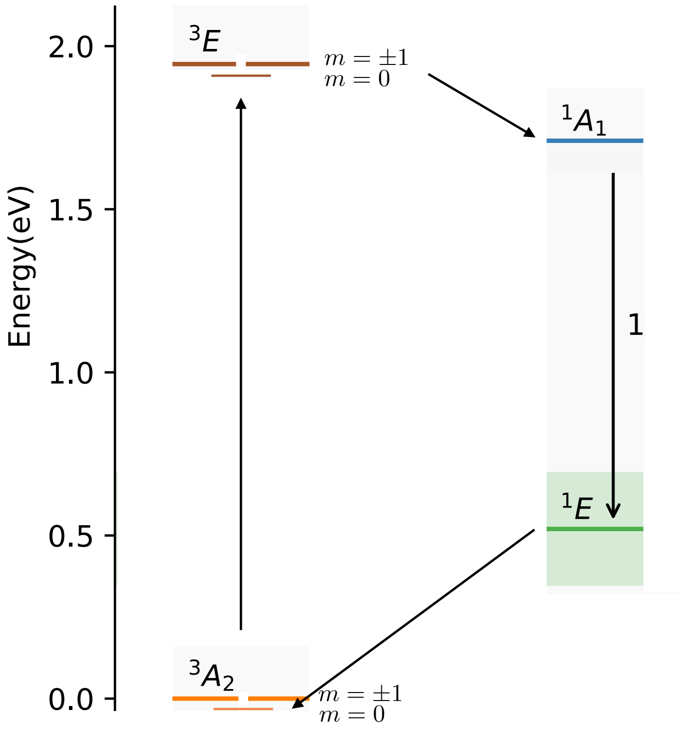
Where We Are Heading

Challenge 2:

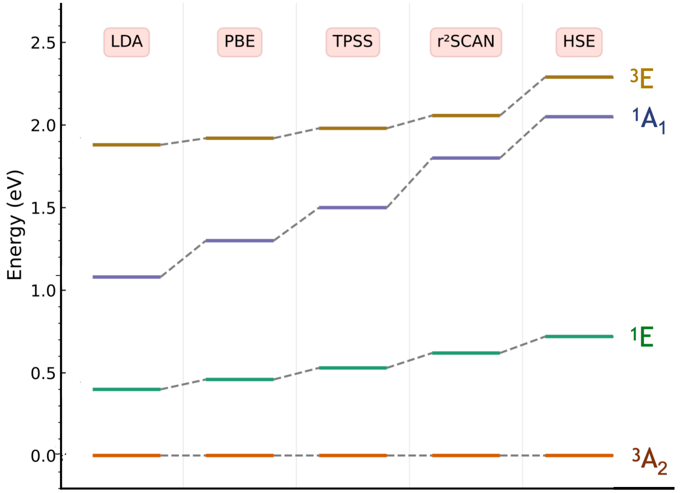
Ground and excited states of defects in solids, e.g. negatively charged N/V defect in diamond



Optical cycle for preparing a pure spin state, for quantum technology



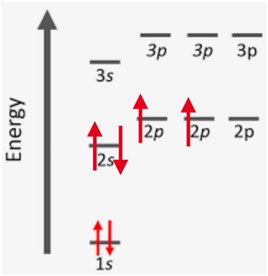
Density functional theory (DFT) gives results that depend strongly on the functional used



A. Ivanov, Y.L.A. Schmerwitz, G. Levi, H. Jónsson, *SciPost Physics* **15**, 009 (2023)

Use instead embedding approach: Quantum chemistry calculation of the defect in a solid described by DFT.

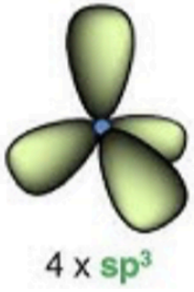
Start with a mean field approximation



C atom

Mean field approximation for electronic systems:
Each electron is only influenced by the average interaction with the other electrons.

Each electron is described by a function, an *orbital*, that only depends on its coordinates, not coordinates of the other electrons.
Spatial orbitals are functions of 3 variables, e.g. (x, y, z) or (r, θ, ϕ).
Spin is added as a 4th ‘variable’ to give *spin-orbitals*
(but, skip writing spin here explicitly for simplicity).



In a system of *N* electrons, get a product of *N* occupied orbitals
(Hartree product) $\phi_a(\mathbf{r}_1)\phi_b(\mathbf{r}_2)\ldots\phi_i(\mathbf{r}_i)\ldots\phi_j(\mathbf{r}_j)\ldots\phi_o(\mathbf{r}_N)$

The orbitals need to be orthonormal $\int \phi_i(\mathbf{r})\phi_j(\mathbf{r})\, d^3\mathbf{r} = \delta_{i,j}$

Electrons are fermions, so need to apply Fermi-Dirac statistics, antisymmetry.

[2]

Hartree-Fock theory

One way to implement Fermi-Dirac is to construct a Slater determinant of the orbitals (or, rather, spin-orbitals), *Fock exchange*

$$\psi(1,2,\ldots,N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_a(1) & \phi_b(1) & \ldots & \phi_0(1) \\ \phi_a(2) & \phi_b(2) & \ldots & \phi_0(2) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_a(n) & \phi_b(n) & \ldots & \phi_0(N) \end{vmatrix}$$

where *(1)* indicates spatial coordinates and spin of electron *1*, etc.

and calculate the expectation value of the Hamiltonian

$$< \psi | H | \psi >$$

where the Hamiltonian is $H = \sum_i^N h(i) + \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}}$

with $h(i) \equiv -\frac{1}{2} \nabla_i^2 - \sum_A^M \frac{Z_A}{r_{iA}}.$

[3]

Variational calculation of the ground electronic state

Recall, variational calculation of the ground electronic state:

Given a time-independent Hamiltonian, H, with eigenstates $|\phi_n >$

$$H|\phi_n >= E_n|\phi_n >$$

where *n=0* for ground state, *n=1* for first excited state, etc.

then for any arbitrary state vector $|\psi >$ in the space spanned by the eigenstates, i.e.

$$|\psi >= \sum_n c_n|\phi_n > \ ,$$

we have

$$< H >= \frac{< \psi|H|\psi >}{< \psi|\psi >} \ \geq \ E_0 \ .$$

The best estimate of the energy of the ground state is obtained by finding the minimum in the expectation value of the energy, and the corresponding state vector is the best estimate of the ground state.

Post-HF: Exact solutions to the non-relativistic Schrödinger equation

Build an exact N-electron wave function on top of the Hartree-Fock approximation $|\psi_0>$.

Generate a complete set of slater determinants (SD) by promoting electrons from occupied to unoccupied orbitals in the HF slater determinant.

Exact solution can be obtained by taking a linear combination of all SDs that can be formed with these orbitals and minimize $<\psi|H|\psi>$,

configuration interaction (CI)

$$|\Phi>= \underbrace{C_0|\psi_0>}_{\text{Hartree-Fock SD}} + \underbrace{\sum_r \sum_a C_a^r|\psi_a^r>}_{\text{single excitation SDs}} + \underbrace{\sum_a \sum_{b>a} \sum_r \sum_{s>r} C_{ab}^{rs}|\psi_{ab}^{rs}>}_{\text{double excitation SDs}} + \ldots.$$

By minimizing the expectation value of the energy

$$<\Psi|H|\Psi>_{(c_0,\{c_a^r\},\{c_{ab}^{rs}\}\ldots)}$$

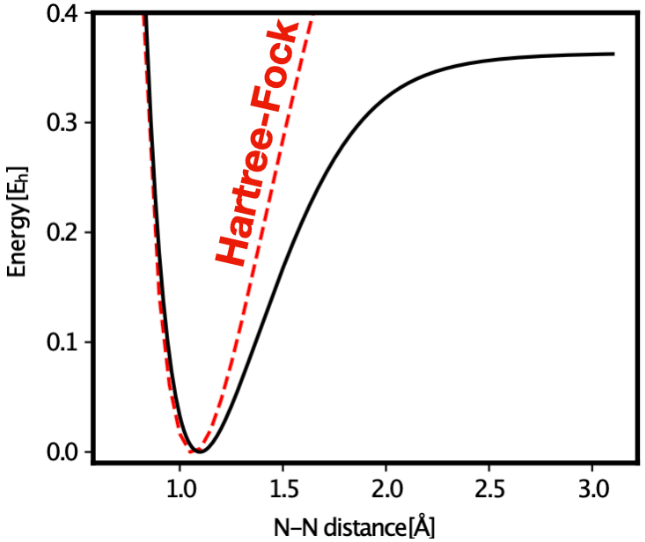
with respect to the linear expansion coefficients, a matrix inversion problem

an estimate of the ground state wave function is obtained
and higher order solutions give estimates of excited states.

If the basis set is complete, an exact value of energy and wave function is obtained.

Can find coefficients for lowest energy solution by using iterative Davidson (preconditioned Lanczos) [5]

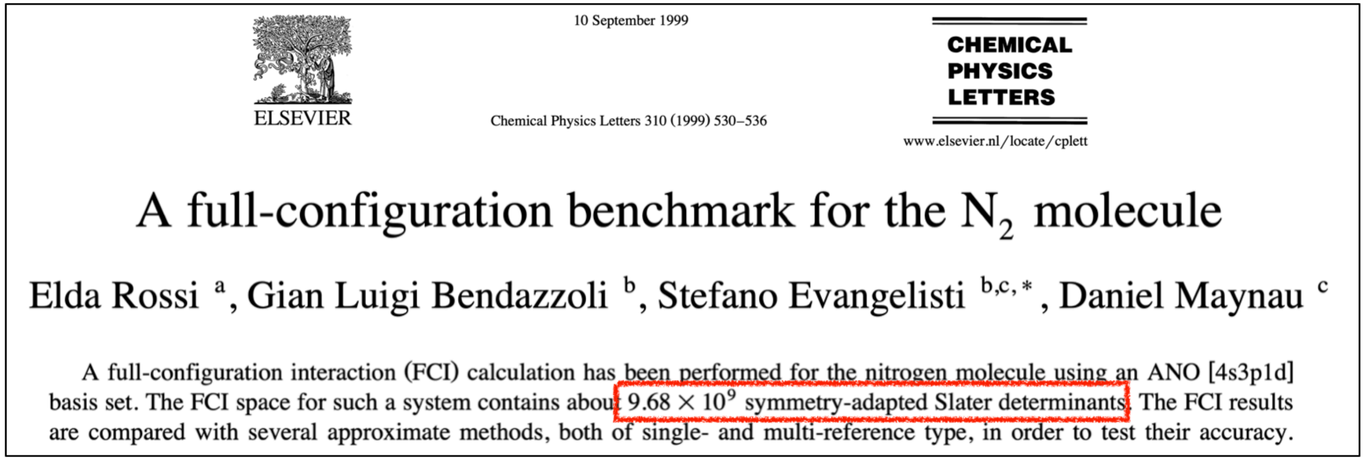
Example: N_2 calculated using Configuration Interaction



N_2 Dissociation Curve

34 orbital CI ~10^{10} Slater determinants

$$|\psi^N_{gs}\rangle = \sum_{\alpha=1}^{\alpha=10^{10}} c^{gs}_{\alpha} |\phi^N_{\alpha}\rangle$$



ELSEVIER

10 September 1999

Chemical Physics Letters 310 (1999) 530–536

CHEMICAL
PHYSICS
LETTERS

www.elsevier.nl/locate/cplett

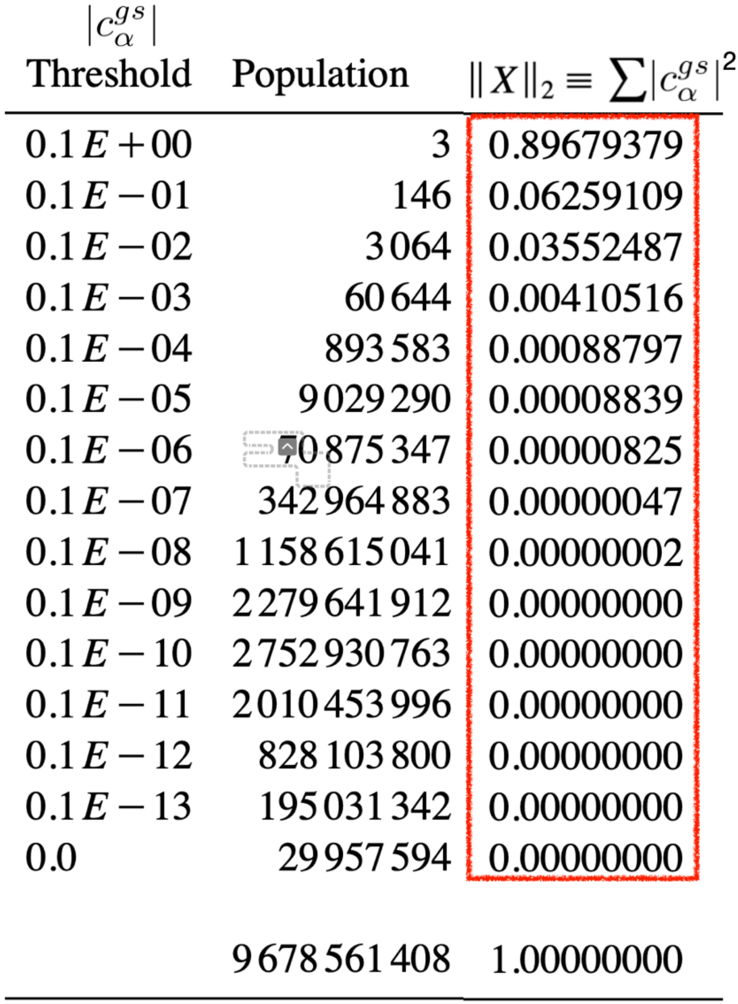
A full-configuration benchmark for the N_2 molecule

Elda Rossi [a], Gian Luigi Bendazzoli [b], Stefano Evangelisti [b,c,*], Daniel Maynau [c]

A full-configuration interaction (FCI) calculation has been performed for the nitrogen molecule using an ANO [4s3p1d] basis set. The FCI space for such a system contains about 9.68×10^9 symmetry-adapted Slater determinants. The FCI results are compared with several approximate methods, both of single- and multi-reference type, in order to test their accuracy.

[6]

Example: N_2 calculated using Configuration Interaction



$\lvert c_\alpha^{gs}\rvert$ Threshold	Population	$\lVert X\rVert_2 \equiv \sum \lvert c_\alpha^{gs}\rvert^2$
$0.1E+00$	3	0.89679379
$0.1E-01$	146	0.06259109
$0.1E-02$	3 064	0.03552487
$0.1E-03$	60 644	0.00410516
$0.1E-04$	893 583	0.00088797
$0.1E-05$	9 029 290	0.00008839
$0.1E-06$	70 875 347	0.00000825
$0.1E-07$	342 964 883	0.00000047
$0.1E-08$	1 158 615 041	0.00000002
$0.1E-09$	2 279 641 912	0.00000000
$0.1E-10$	2 752 930 763	0.00000000
$0.1E-11$	2 010 453 996	0.00000000
$0.1E-12$	828 103 800	0.00000000
$0.1E-13$	195 031 342	0.00000000
0.0	29 957 594	0.00000000
	9 678 561 408	1.00000000

34 molecular orbitals,
full configuration interaction (FCI)

$$|\psi_{gs}^{N}\rangle = \sum_{\alpha=1}^{\alpha=10^{10}} c_\alpha^{gs} |\phi_\alpha^{N}\rangle$$

Most Slater determinants irrelevant!

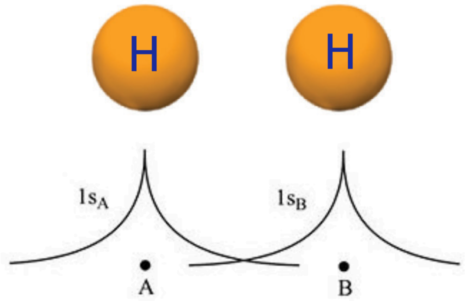
[7]

In calculations of molecules: use molecular orbitals instead of atomic

An electronic energy surface describes how a system's energy varies as a function of the electronic degrees of freedom (i.e. expansion coefficients)

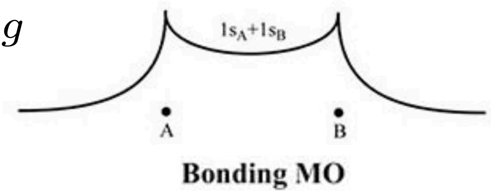
Example, H_2

use minimal atomic basis set,
1s atomic orbital on each atom

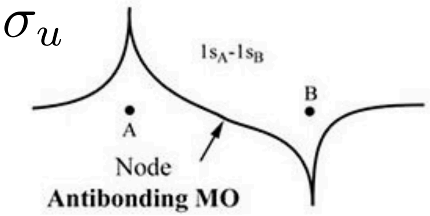


Molecular orbitals for spin-up, α, and spin-down, β

σ_g

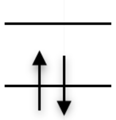


σ_u



The ground state wave function is

$$\psi = \sigma_g(1)\sigma_g(2)(\alpha(1)\beta(2) - \alpha(2)\beta(1))$$

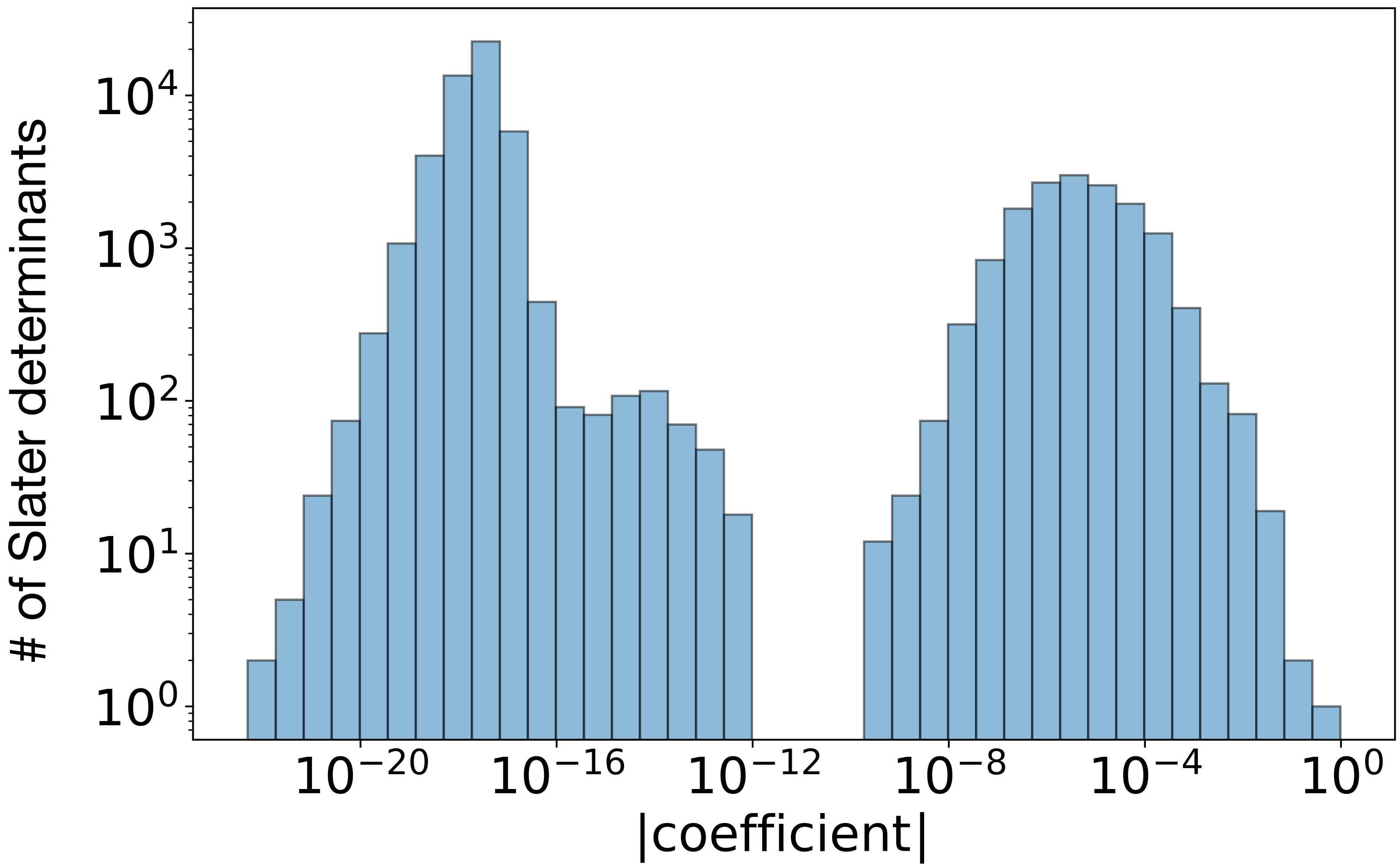


LUMO, σ_u

HOMO, σ_g

$= |11\ 00>$

Full CI calculation of N_2 using 10 molecular orbitals

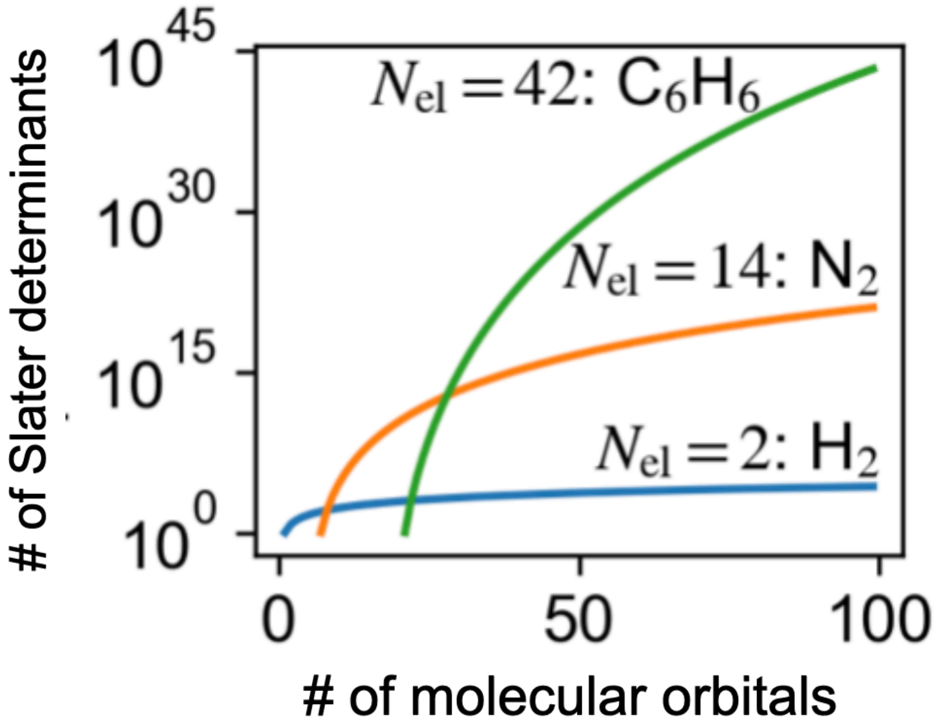


Most Slater determinants irrelevant!

[9]

# of Slater determinants
in full CI calculation

$$N = \dim(\mathcal{H}) = \binom{2N_{\mathrm{orb}}}{N_{\mathrm{el}}}$$



$$|\Psi_{\mathrm{gs}}^{\mathrm{ex}}\rangle = \sum_{i=1}^{N} c_i \, |\psi_i\rangle$$ full CI

$$\approx \sum_{i=1}^{N_s \ll N} \tilde{c}_i \, |\psi_i\rangle$$ selective CI

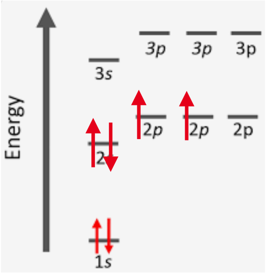
The challenge is to identify the
important Slater determinants

[10]

Representation of the Slater Determinants

Example: C atom

ground state SD



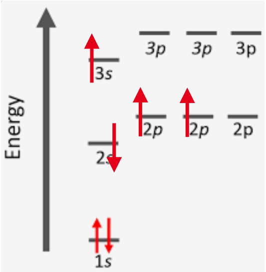
2s(up) 2s(down)

|11 11 10 10 00 00 00 00 00>

1s(down)

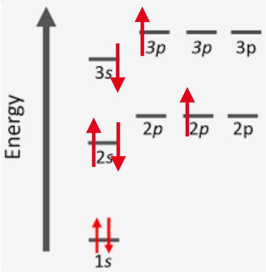
1s(up)

single excitation SD



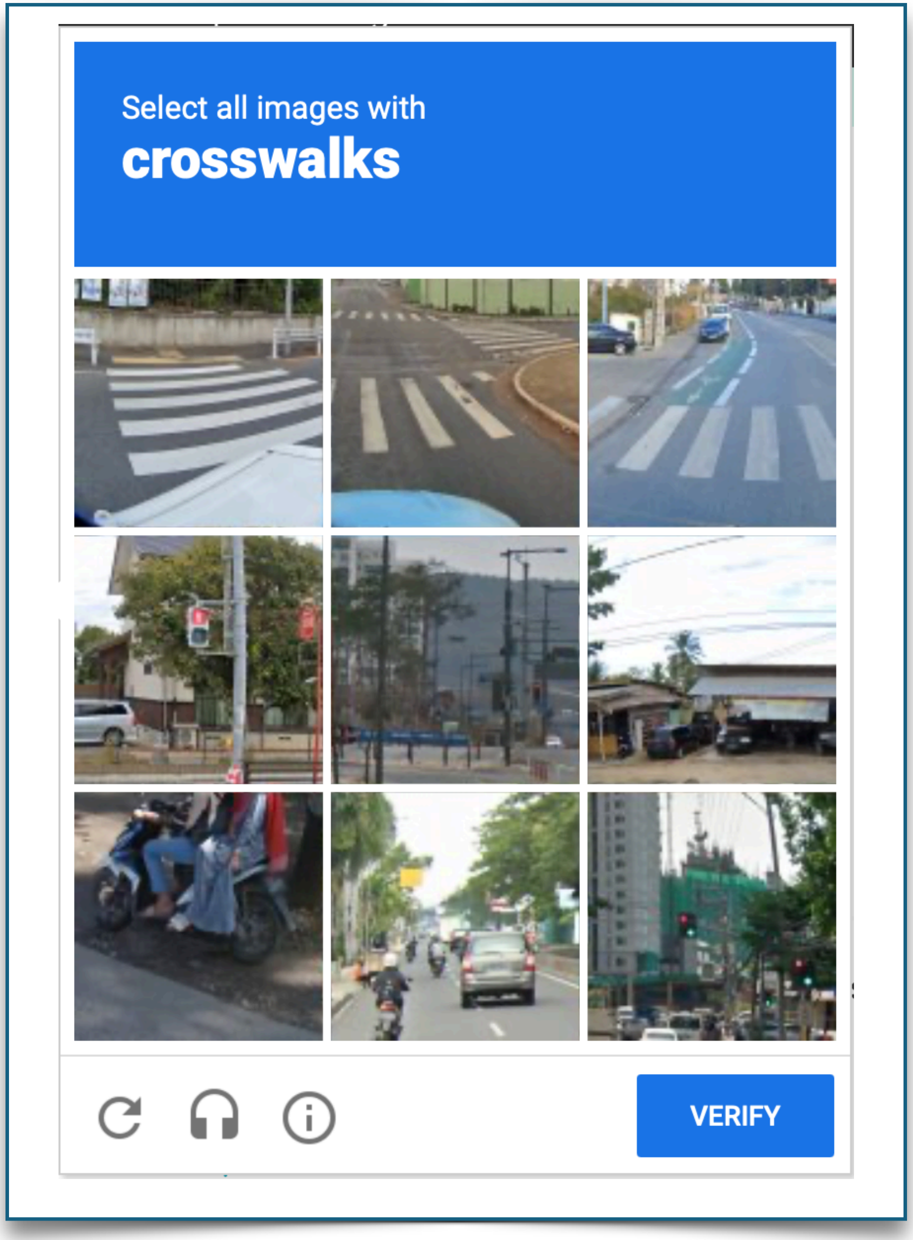
|11 01 10 10 10 00 00 00 00>

double excitation SD



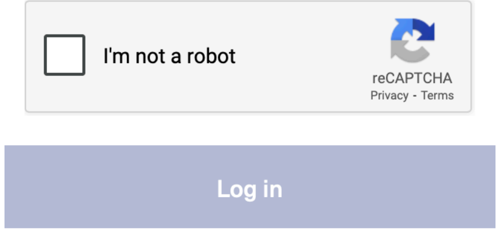
|11 11 00 10 00 01 10 00 00>

[11]

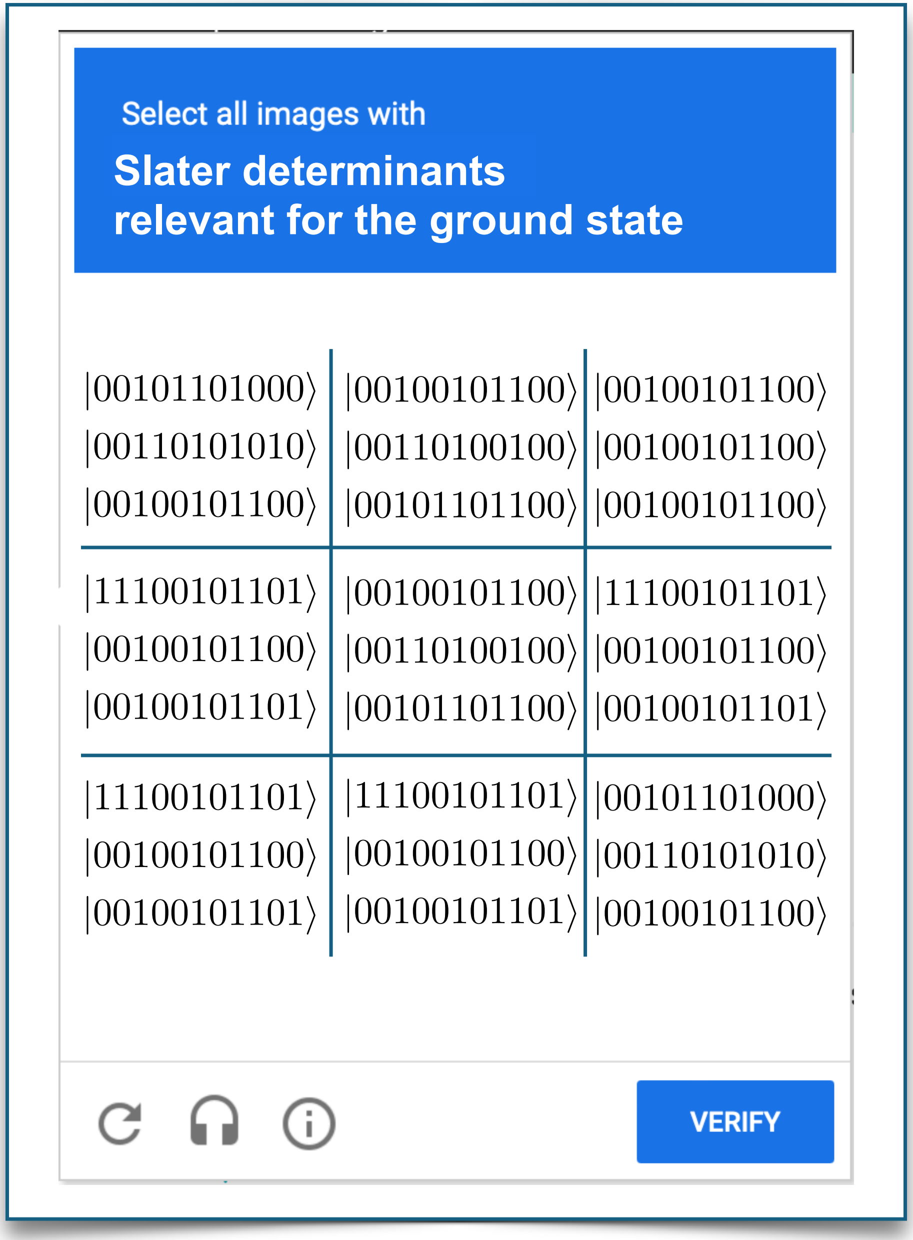


Select all images with

crosswalks



[12]



Select all images with

Slater determinants relevant for the ground state

$\vert 00101101000\rangle$ $\vert 00110101010\rangle$ $\vert 00100101100\rangle$	$\vert 00100101100\rangle$ $\vert 00110100100\rangle$ $\vert 00101101100\rangle$	$\vert 00100101100\rangle$ $\vert 00100101100\rangle$ $\vert 00100101100\rangle$
$\vert 11100101101\rangle$ $\vert 00100101100\rangle$ $\vert 00100101101\rangle$	$\vert 00100101100\rangle$ $\vert 00110100100\rangle$ $\vert 00101101100\rangle$	$\vert 11100101101\rangle$ $\vert 00100101100\rangle$ $\vert 00100101101\rangle$
$\vert 11100101101\rangle$ $\vert 00100101100\rangle$ $\vert 00100101101\rangle$	$\vert 11100101101\rangle$ $\vert 00100101100\rangle$ $\vert 00100101101\rangle$	$\vert 00101101000\rangle$ $\vert 00110101010\rangle$ $\vert 00100101100\rangle$

[13]

Convolutional Neural Networks

Used in particular for image analysis.

Good at detecting patterns and features.

Learns features by itself via filter (or kernel) optimization.

Slide along input features and provide translation-equivariant responses known as feature maps.

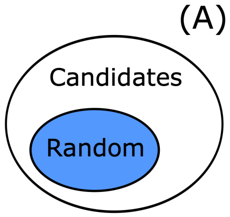
The network learns the filters through automated learning.

Chihuahua
or
muffin?

[14]

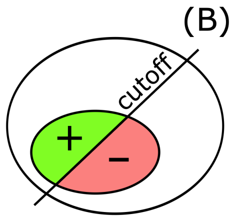
Is the Slater Determinant |111110010001100101000000000000000> relevant ?

Basic idea:



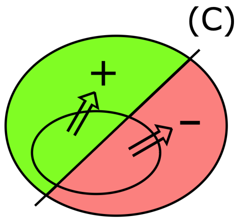
From a pool of SDs, choose a subset at random and calculate the CI expansion coefficients

⇓

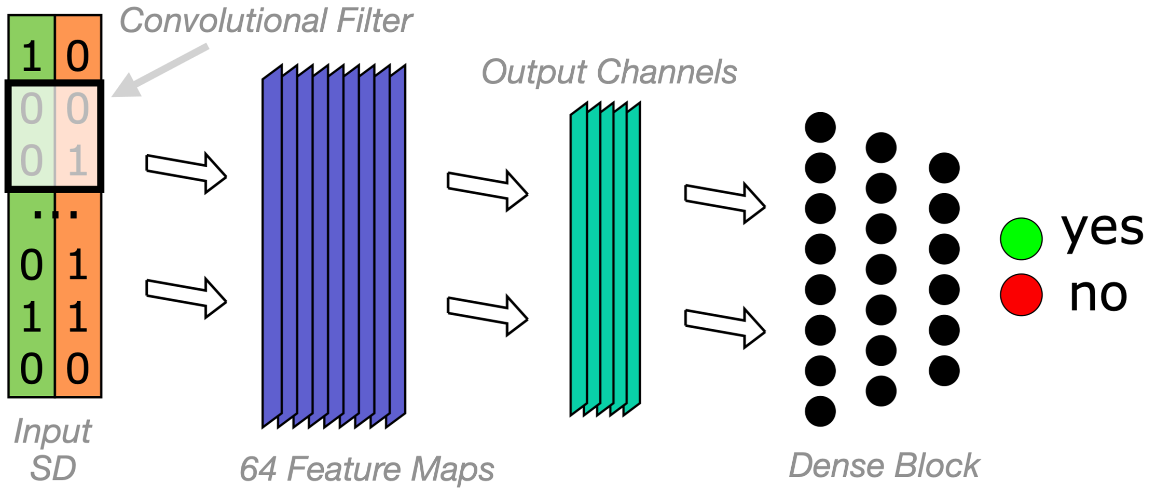


Select a certain user determined fraction of the SDs based on the exp. coeffs. and train NN

⇓



Apply the developed selection criteria to the full set of SDs.



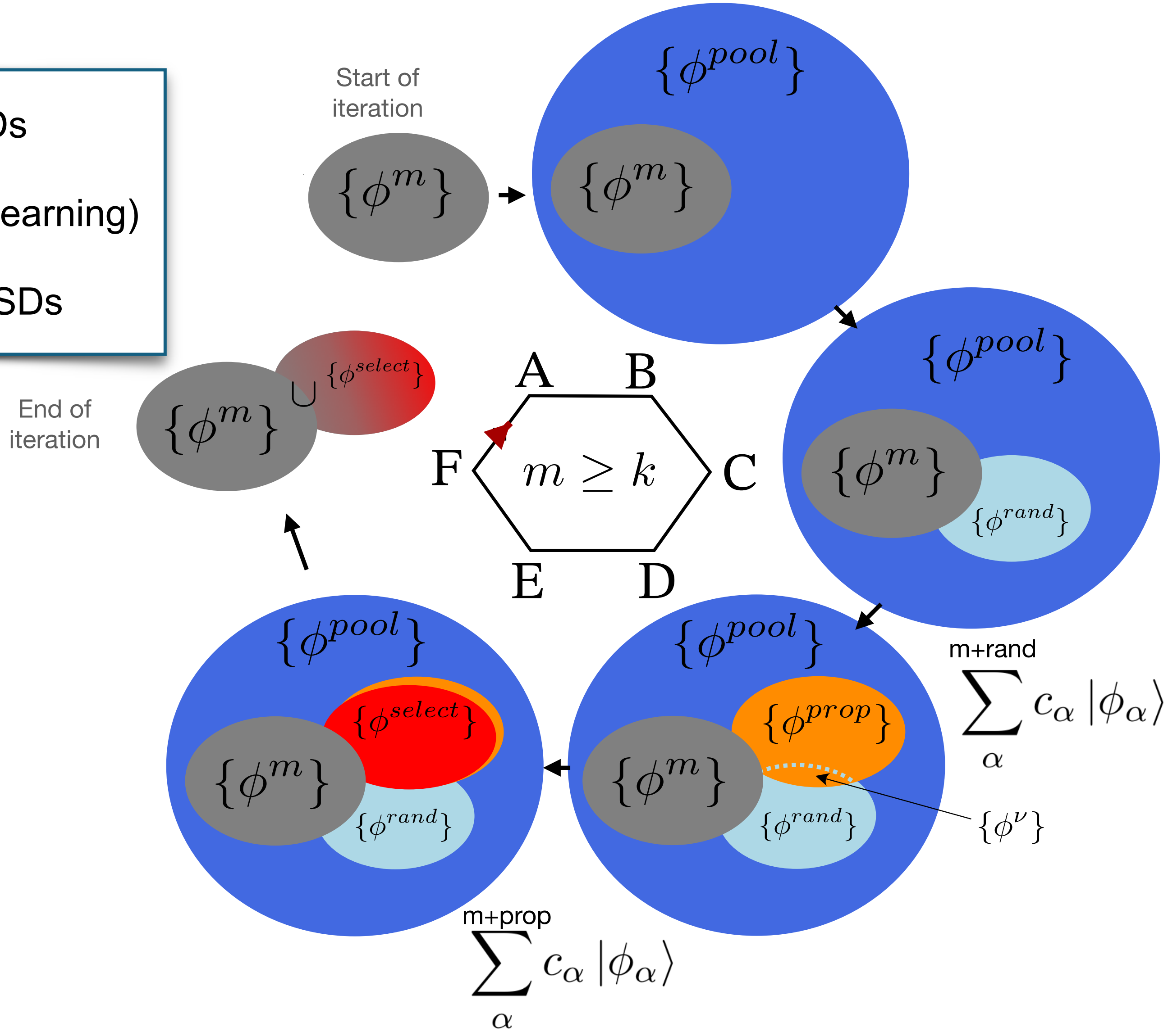
[15]

NNCI: Neural Network Configuration Interaction Algorithm

i) generate candidate SDs

ii) train on the fly (active learning)

iii) select most important SDs



[16]

HF and various matrix elements calculated using **GPAW** open software

GPAW stands for “grid, projector augmented wave”

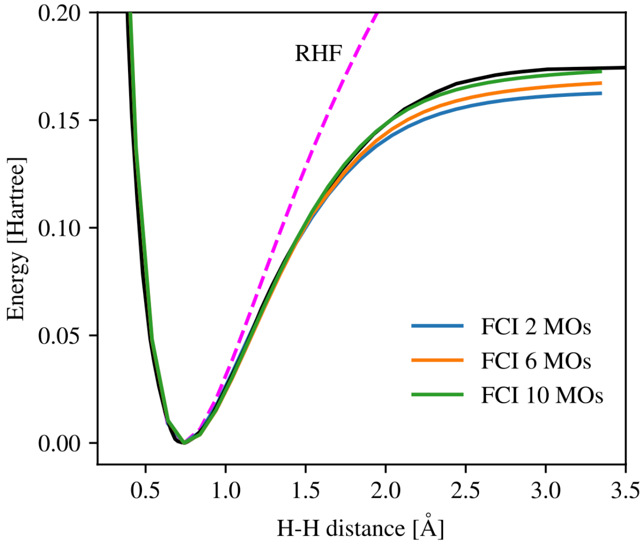
The PAW formalism (by P. E. Blöchl) is used to split orbitals:

- a smooth function in the outer region and
- a less smooth (with cusp) inner region function equal to that of an isolated atom.

Here, the **Hartree-Fock** calculations are carried out by first reading in an LCAO basis set represented on a grid but then the orbitals are optimised using a plane wave representation,

(with 1000 eV energy cutoff, 0.18 Å real space grid spacing, > 12.5 Å of vacuum between any atom and box boundary).

CI calculation of H2



GPAW: Open Python Package for Electronic-structure Calculations', J. J. Mortensen, A. H. Larsen, M. Kuisma, A. V. Ivanov, *et al.*, *J. Chem. Phys.* 160, 092503 (2024).

NNCI calculations then carried out using **SOLAX** open software

Code for (neural-network-supported)
CI computations

- Python library: `import solax as sx`
- Actors: Basis, State, Operator, Matrix
- Neural-network-support out of the box
- Tools: JAX + NumPy + Pandas (+SciPy)
- GPU-accelerated — multi-GPU under dev.

SOLAX: A Python solver for fermionic quantum systems with neural network support,
Louis Thirion, Philipp Hansmann and Pavlo Bilous, *SciPost Phys. Codebases 51*(2025).

https://github.com/pavlobilous/SOLAX.

Some previous applications of neural networks in CI calculations

J. P. Coe:
“Machine Learning Configuration Interaction”,
JCTC 14, 5739 (2018).
“Machine Learning Configuration Interaction for ab Initio Potential Energy Curves”,
JCTC 15, 6179 (2019).

An iterative active-learning approach.
NN used as a regressor for the prediction of transformed CI coefficients.

Jeong, Gaggioli and Gagliardi:
“Active Learning Configuration Interaction for Excited-State Calculations of Polycyclic Aromatic Hydrocarbons”,
JCTC 17, 7518 (2021).
and
S. D. Flores:
“Chembot: A Machine Learning Approach to Selective Configuration Interaction”,
JCTC 17, 4028 (2021).

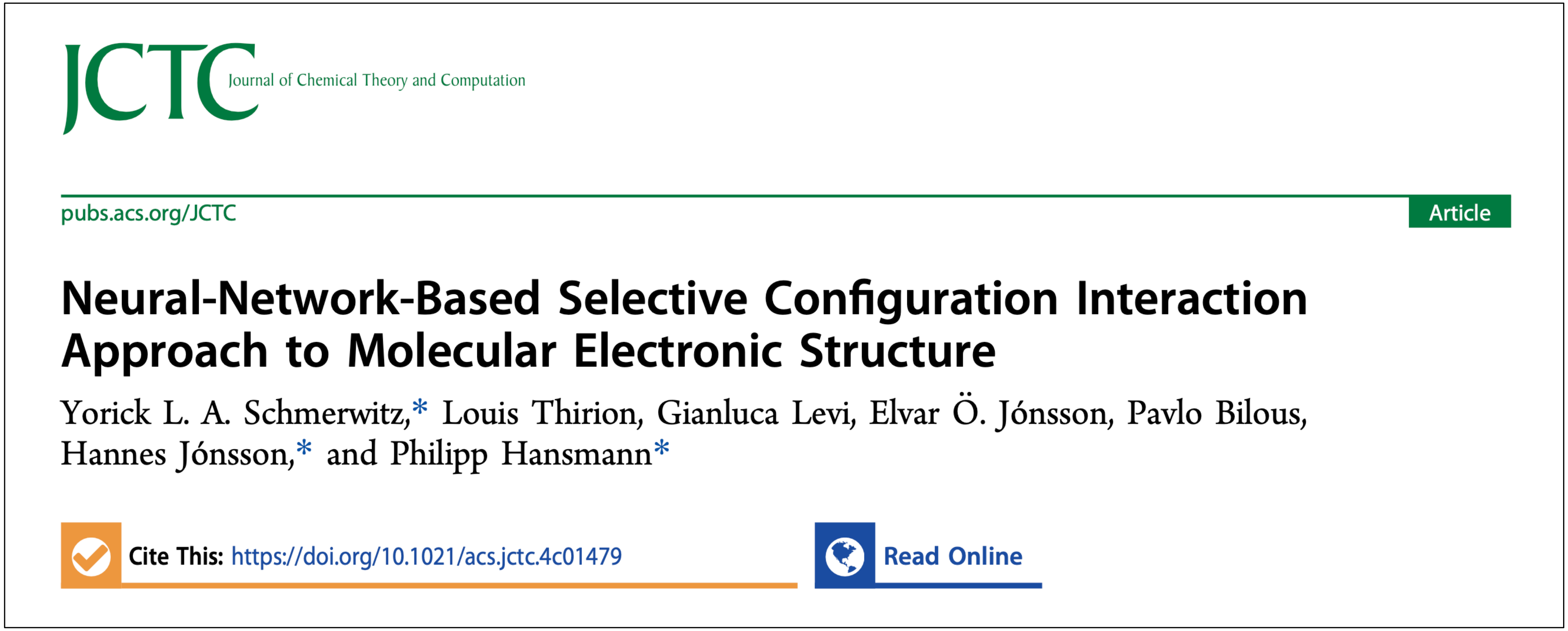
Direct classification approaches were determinants are sorted into “important” and “unimportant” based on a cutoff parameter.

Herzog, Casier, Lebegue, Rocca:
“Solving the Schr. Eqn. in the Configuration Space with Generative Machine Learning”,
JCTC 19 2484 (2023).

CIgen method, a generative NN used to expand the configuration space iteratively by proposing new Slater determinants that are expected to contribute significantly to the wave function.

Present work:

Use a *convolutional* NN classifier, and a *running cutoff* which is automatically computed in active learning based on a desired fraction of important determinants to be added in each iteration.



JCTC Journal of Chemical Theory and Computation

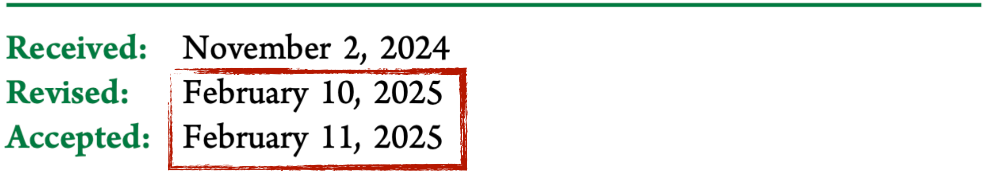
pubs.acs.org/JCTC Article

Neural-Network-Based Selective Configuration Interaction Approach to Molecular Electronic Structure

Yorick L. A. Schmerwitz,* Louis Thirion, Gianluca Levi, Elvar Ö. Jónsson, Pavlo Bilous, Hannes Jónsson,* and Philipp Hansmann*

Cite This: https://doi.org/10.1021/acs.jctc.4c01479 Read Online

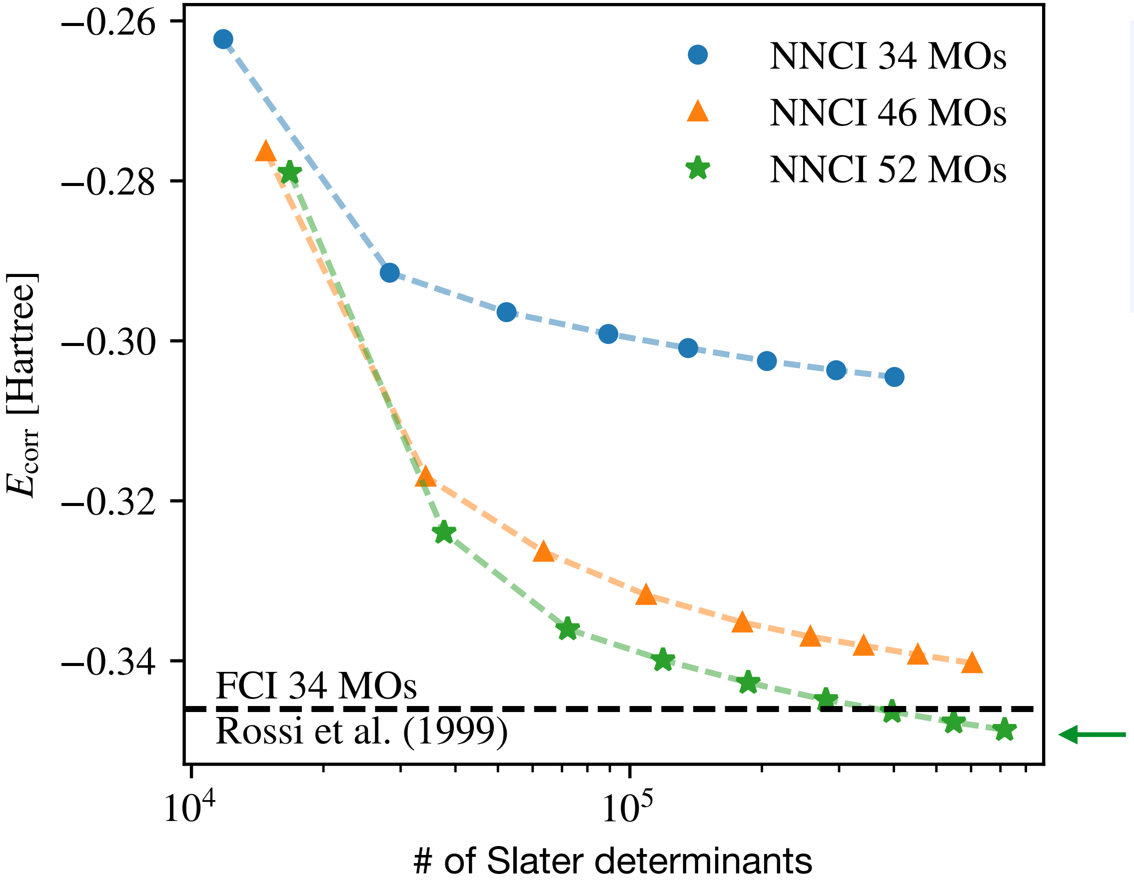
J. Chem. Theory Comput. **21**, 2301 (2025)



Received: November 2, 2024
Revised: February 10, 2025
Accepted: February 11, 2025

NNCI calculations of N_2

Calculated correlation energy at N-N distance of 1.1 Å, *ca*. the optimal bond length.



NNCI

52 orbitals   ~10^5 selected SDs

Full CI of Rossi *et al.*

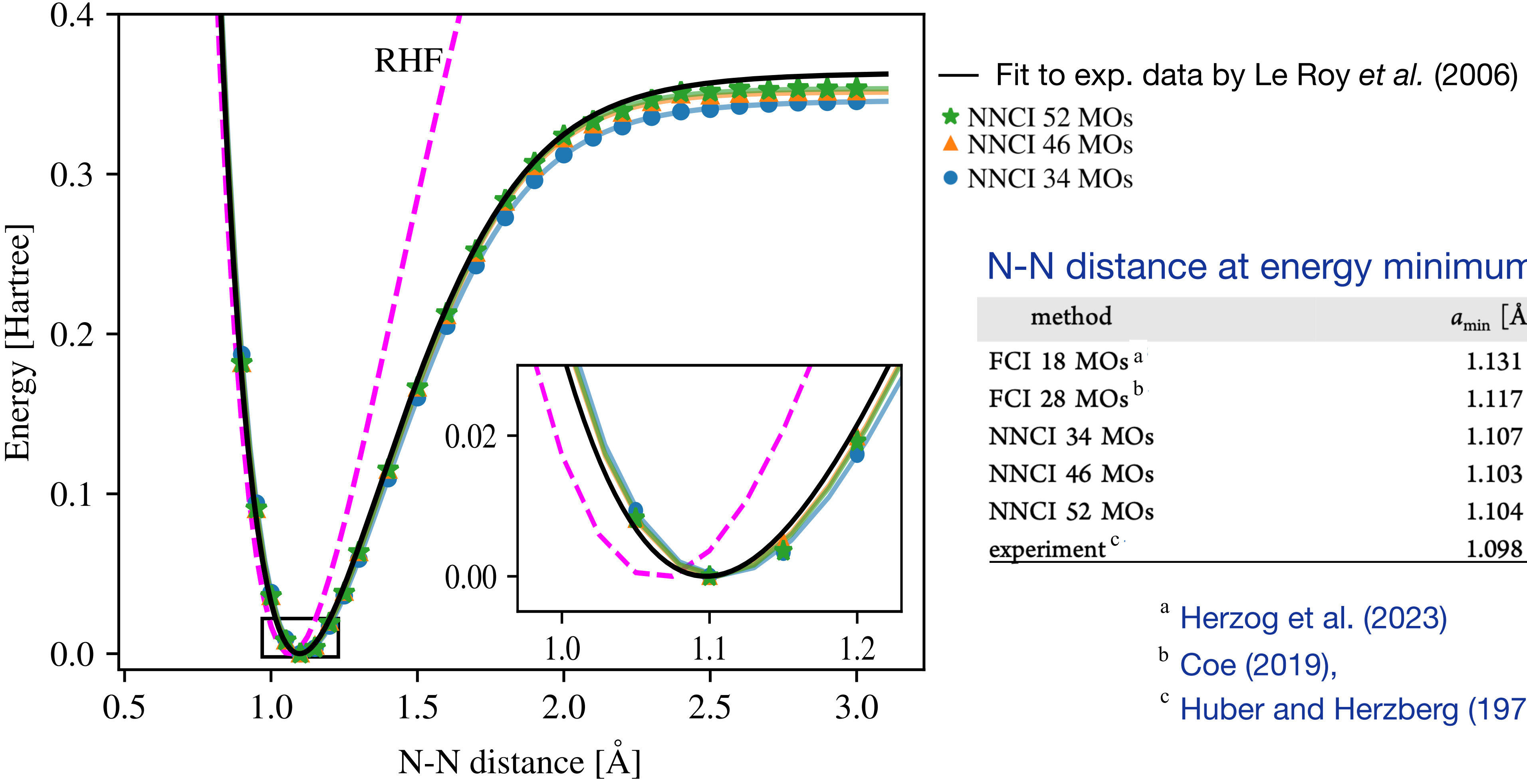
34 orbitals   ~10^{10} SDs

NN selected CI gives larger absolute correlation energy with 4 orders of magnitude fever Slater determinants.

More correlation gained by increasing the number of MOs than by adding more Slater determinants.

[21]

NNCI calculations of N_2



N-N distance at energy minimum

method	$a_{\min}$ [Å]
FCI 18 MOs [a]	1.131
FCI 28 MOs [b]	1.117
NNCI 34 MOs	1.107
NNCI 46 MOs	1.103
NNCI 52 MOs	1.104
experiment [c]	1.098

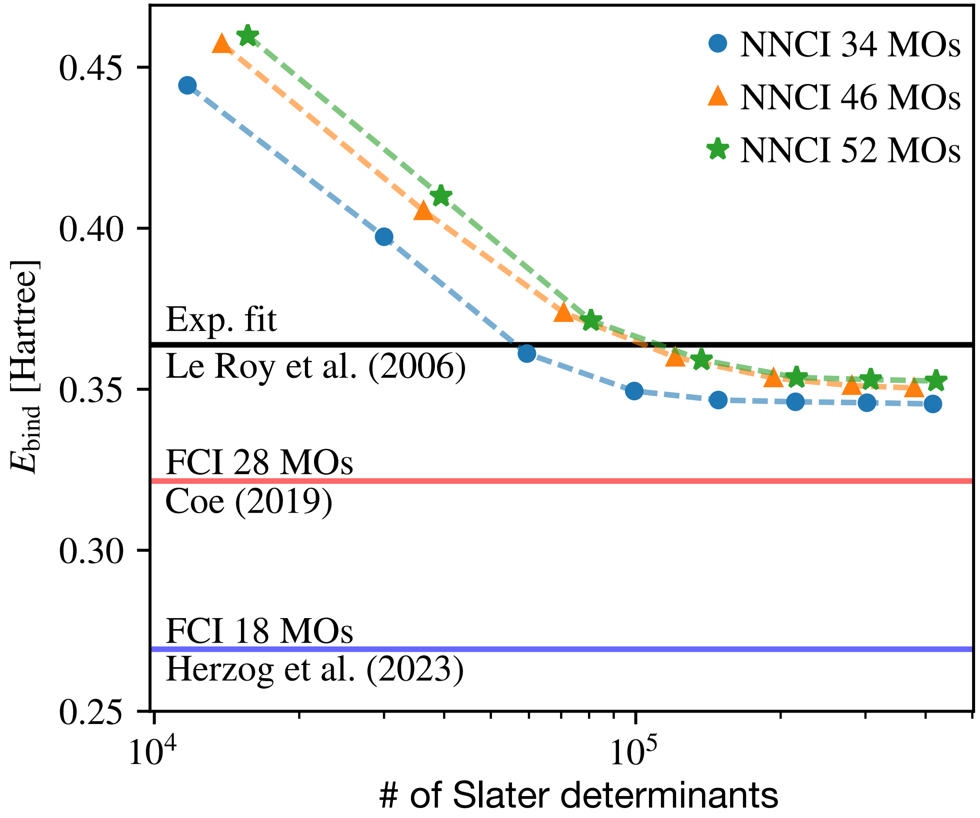
[a] Herzog et al. (2023)

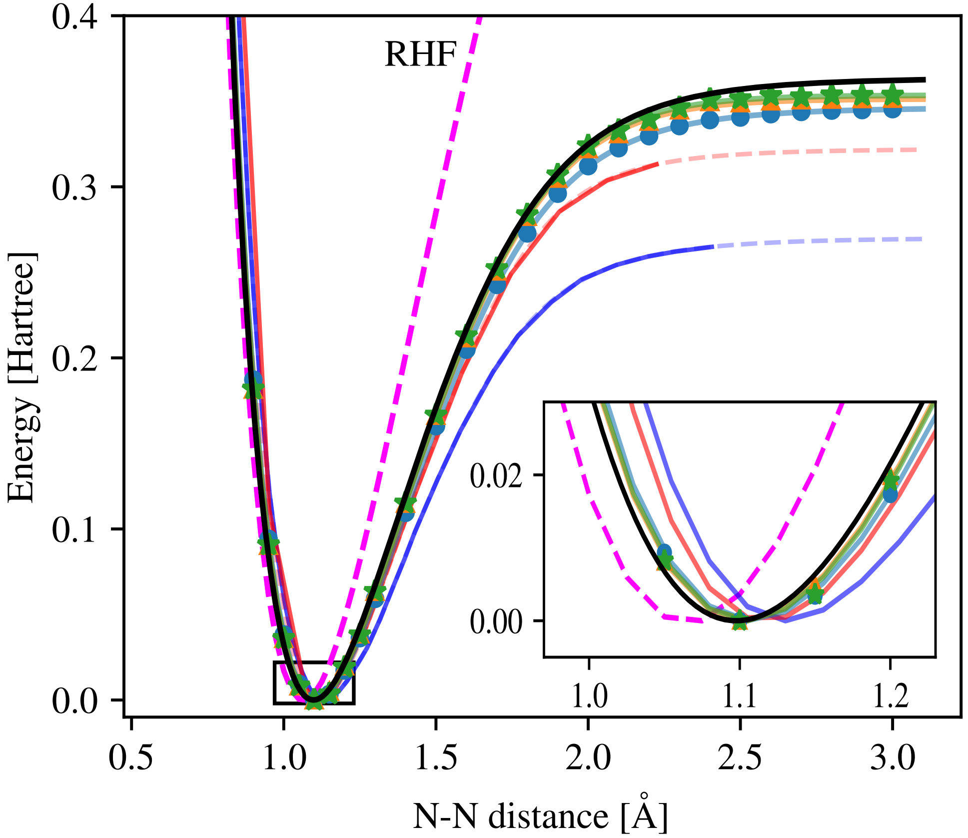
[b] Coe (2019),

[c] Huber and Herzberg (1979)

NNCI calculations of N_2

Compare with a fit to experimental data and previous FCI calculations





Calculated binding energy is converged for each set of orbitals, but still too low compared with experiment by about 3%.

[23]

NNCI calculations of more molecules

Calculated correlation energy at near optimal geometry.

	$(N_{\mathrm{el}}, N_{\mathrm{orb}})$	$\dim\mathcal{H}^{\mathrm{full}}$	$\dim\mathcal{H}^{\mathrm{NNCI}}$	E_{corr}	$E_{\mathrm{corr}}^{\mathrm{FCI}}$	$\dim\mathcal{H}_{\mathrm{ref.}}^{\mathrm{FCI}}$
N_2	(10,52)	$6.75 \cdot 10^{12}$	$7.16 \cdot 10^{5}$	-0.349	-0.346[b]	$9.68 \cdot 10^{9}$
CO	(10,46)	$1.88 \cdot 10^{12}$	$2.44 \cdot 10^{5}$	-0.306	-0.215[a]	$1.01 \cdot 10^{9}$
NH_3	(8,56)	$4.49 \cdot 10^{10}$	$1.93 \cdot 10^{5}$	-0.221	-0.208[a]	$1.41 \cdot 10^{10}$
H_2O	(8,43)	$1.52 \cdot 10^{10}$	$2.51 \cdot 10^{5}$	-0.218	-0.216[a]	$1.81 \cdot 10^{9}$

a) Gao, Imamura, Kasagi, Yoshida:
“Distributed Implementation of Full Configuration Interaction for One Trillion Determinants”,
JCTC 20, 1185 (2024).

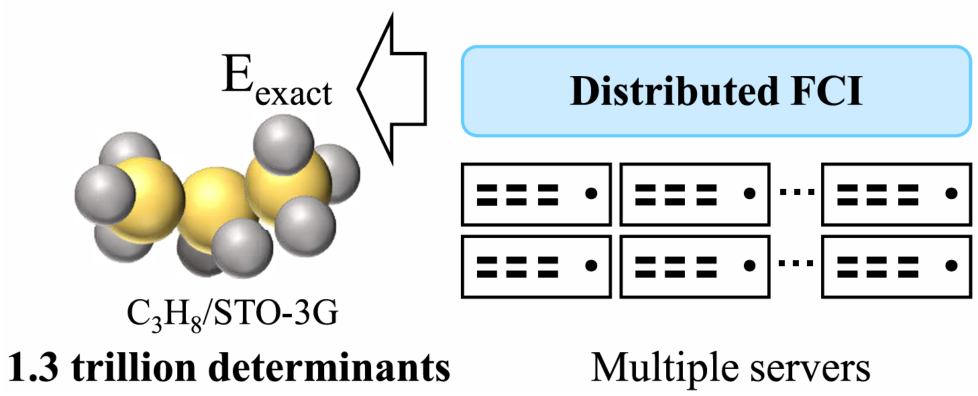
b) Rossi, Bendazzoli, Evangelisti, Maynau:
“A full-configuration benchmark for the N_2 molecule”,
Chem. Phys. Lett. 310 530 (1999).

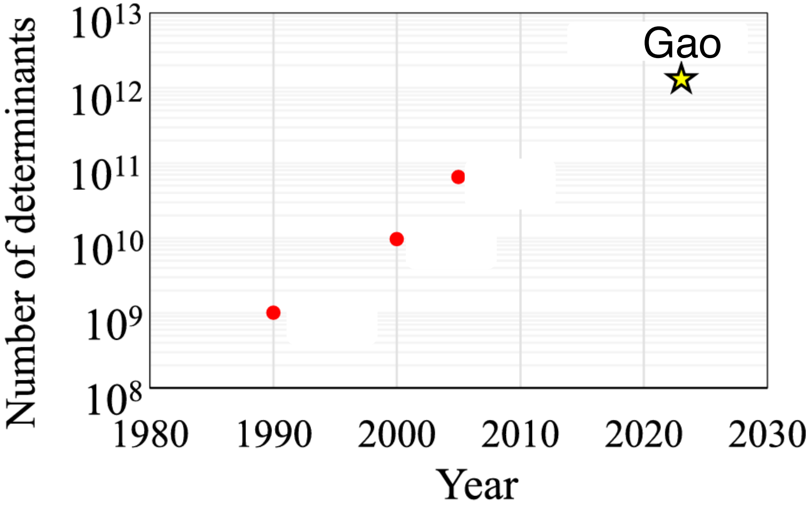
NNCI with 10^5 to 10^6 Slater determinants
gives larger absolute correlation energy
than FCI including 10^9 to 10^{10} SDs.

JCTC Journal of Chemical Theory and Computation

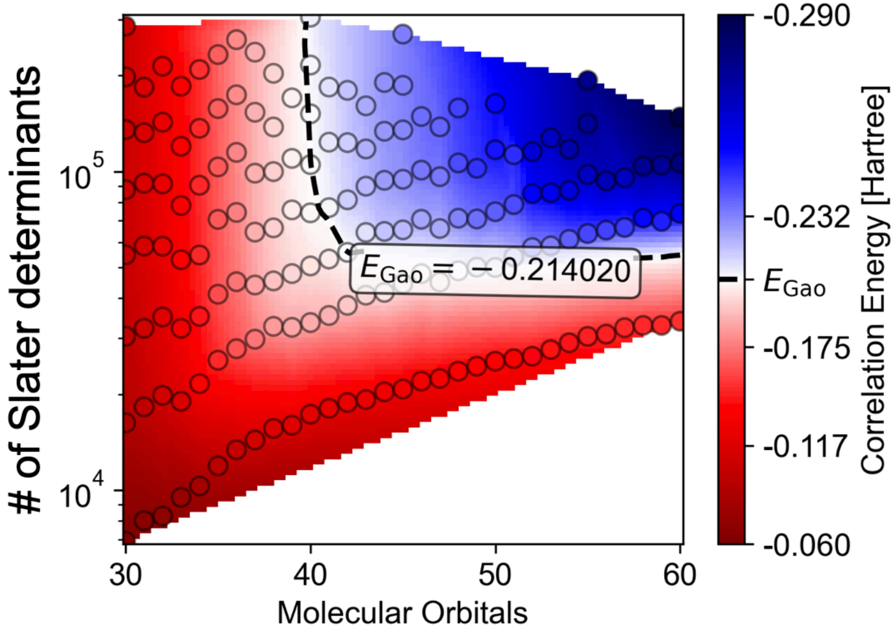
Distributed Implementation of Full Configuration Interaction for One Trillion Determinants

Hong Gao,* Satoshi Imamura, Akihiko Kasagi, and Eiji Yoshida





NNCI calculations



50% larger correlation energy obtained using NNCI w. 6 orders of magnitude fewer Slater determinants

L. Thirion *et al.* (unpublished)

[25]

Summary

Method:

A neural network algorithm based on convolutional filter approach has been applied in **selective configuration interaction** calculations of molecules.

This NNCI algorithm has been implemented in open software, **GPAW/SOLAX**, where either finite or periodic boundary conditions can be applied, setting the stage for embedding calculations of condensed matter.

Integration of the two codes is ongoing.

Results:

For the N_2 molecule: the full dissociation curve has been calculated.
At min. energy N-N distance, a calculation with 4 orders of magnitude fewer Slater determinants than FCI of Rossi *et al.* gives larger correlation energy.

For NH_3, H_2O, CO and propane: larger correlation energy is obtained using 4 to 6 orders of magnitude fewer Slater determinants than FCI calcs. (Gao *et al.* 2024).

The NNCI algorithm benefits greatly from a large set of MOs, going well beyond what can be used as a basis for full CI.