

# Cheap and near exact calculations with large active spaces.

New applications for the SHCI algorithm

**Bastien Mussard**

bastien.mussard@colorado.edu

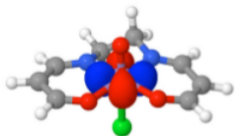
<https://mussard.github.io/>

**Sandeep Sharma**

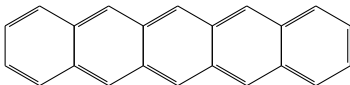
sandeep.sharma@colorado.edu

<http://www.colorado.edu/lab/sharmagroup/>

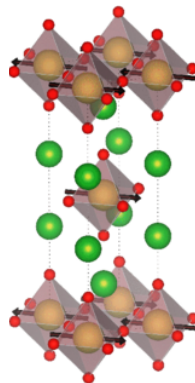
University of Colorado at Boulder



oxo-Mn(Salen) cluster



Acenes and derivatives



Transition metal oxydes

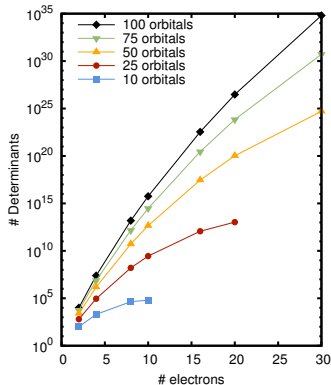
The systems of interest in **biological processes**, in **energy conversion** and **energy storage** are challenging because of static and dynamical correlation.

The wavefunction is written with **determinants**:

$$|\Psi_{\text{FCI}}\rangle = \sum_{D_i \in \text{all}} c_i |D_i\rangle \quad \boxed{\text{all orbitals}} \quad \binom{N_{\text{orb}}}{N_{\text{elec}}}$$

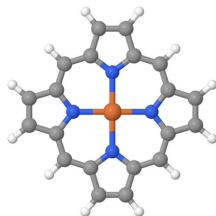
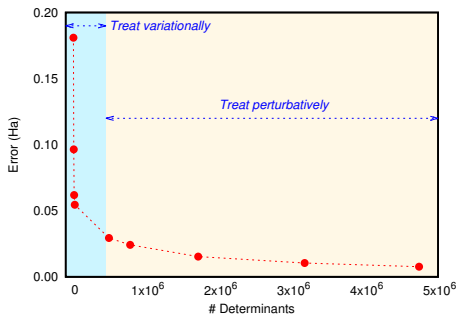
In **multireference methods**, one chooses a subset of orbitals and a subset of electrons, defining an **active space** which is treated exactly:

$$|\Psi_{\text{CAS}}\rangle = \sum_{D_n \in \text{active}} c_n |D_n\rangle \quad \boxed{\text{Core}} \quad \boxed{\text{Active}} \quad \boxed{\text{Virtual}} \quad \binom{N_{\text{orb}}}{N_{\text{elec}}}$$



For larger, more challenging systems, it is impossible to perform exact calculations for reasonable active spaces, and approximate methods are devised.

→ The size of the tractable active space is a crucial parameter.



Fe(Porphyrin)

Active space (32e,29o) Fe 3d, C 2p<sub>z</sub>, N 2p<sub>z</sub>  
 The Hilbert space is  $> 10^{14}$  determinants

## Selected CI + Perturbation theory

[Huron JCP (1973)]

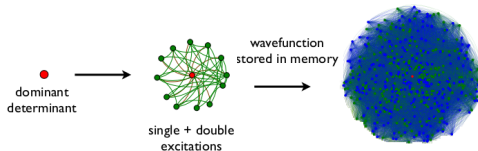
- 1) Variational stage: calculate  $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$  in a space of important det.

→ How to quickly select the “important” determinants?

- 2) Perturbational stage: calculate  $E_2 = \langle \Psi_0 | \hat{V} \frac{1}{\hat{H} - E_0} \hat{V} | \Psi_0 \rangle$

→ How to avoid the memory bottleneck?

→ Iteratively find important det. by connections to a current space.



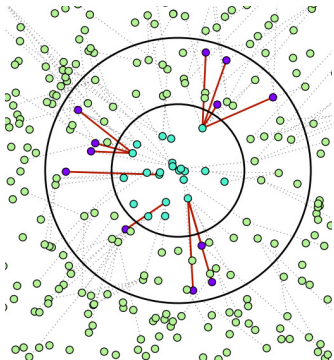
At iteration  $\mu$ , we know  $|\Psi_0^\mu\rangle$ , and  $c_i^\mu$ . The variational space  $\mathcal{V}^\mu$  is augmented by the determinants in:

$$\mathcal{C}^\mu(\epsilon_1) = \{ |D_a\rangle \mid \mathbf{f}^\mu(|D_a\rangle) > \epsilon_1 \}$$

HCI does that with a simple criterion  $\mathbf{f}^\mu$  that allows an **easy, and fast implementation**:

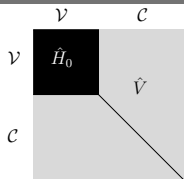
$$\mathbf{f}^\mu(|D_a\rangle) = \max_{D_i \in \mathcal{V}^\mu} |H_{ai} c_i^\mu|$$

$\hat{H}$  is diagonalized in  $\mathcal{V}^\mu + \mathcal{C}^\mu$  to get  $|\Psi_0^{\mu+1}\rangle$  and  $c_i^{\mu+1}$ .



$$E_2 = \langle \Psi_0 | \hat{V} \frac{1}{\hat{H}_0 - E_0} \hat{V} | \Psi_0 \rangle = \sum_{D_a \in \mathcal{C}} \frac{1}{E_a - E_0} \left( \sum_{D_i \in \mathcal{V}} H_{ai} c_i \right)^2$$

$\Psi_1$



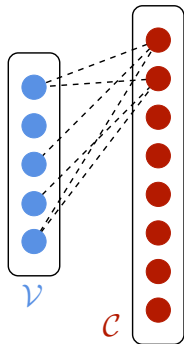
### 1) Deterministic evaluation (involves a truncation)

$$E_2^D = \sum_{D_a \in \mathcal{C}(\epsilon_2)} \frac{1}{E_a - E_0} \left( \sum_{D_i \in \mathcal{V}} H_{ai} c_i \right)^2$$

### 2) Stochastic estimate (unbiased)

$$E_2^S = \langle \Psi_0 | \hat{V} \frac{1}{\hat{H}_0 - E_0} \hat{V} | \Psi_0 \rangle$$

where about  $\approx 10^2$  dets of  $|\Psi_0\rangle$  are sampled



### 3) Semistochastic implementation (low stochastic error)

$$E_2 = E_2^D[\epsilon_2^{\text{large}}] + \left( E_2^S[\epsilon_2^{\text{small}}] - E_2^S[\epsilon_2^{\text{large}}] \right)$$

→ too memory intensive!

$$e^2 v^2 N_V \approx 12^2 50^2 10^7 = 10^{12}$$

SHCI = vHCI (optimal active space det) + Semistochastic PT

→ One-parameter ( $\epsilon_1$ ) method that allows efficient large active space calculations with perturbative corrections.

### Ground and Excited states energies

[Holmes,Tubman,Umrigar, JCTC (2016)]

[Sharma *et al.* JCTC (2017)]

[Holmes,Umrigar,Sharma (2017)]

-First row dimers,  $F_2$  (14e,108o)

- $Cr_2$  (12e,190o)

-Mn(Salen) (28e,22o)

-Acenes (singlet fission)

### Self-consistent field feature

[Smith,BM,Holmes,Sharma JCTC (2017) **Editor's Choice for 2017**]

### Spin-orbit coupling

[BM,Sharma JCTC (2017)]

### Dynamical correlation

[BM,Smith,Sharma (in prep.)]

SHCI = vHCI (optimal active space det) + Semistochastic PT

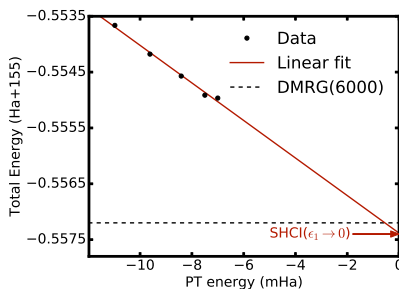
→ One-parameter ( $\epsilon_1$ ) method that allows efficient large active space calculations with perturbative corrections.

Butadiene



with (22e,82o) (all electrons except 1s)

$\epsilon_1$	$N_{\text{var}}$	vHCI	SHCI
$3.10^{-5}$	$1.10^7$	-0.5411	-0.5534(1)
$2.10^{-5}$	$2.10^7$	-0.5441	-0.5540(1)
$1.10^{-5}$	$6.10^7$	-0.5481	-0.5550(1)
SHCI( $\epsilon_1 \rightarrow 0$ )		-0.5574(8)	
CCSD(T)		-0.5550	
CCSDT		-0.5560	
DMRG(M=6000)		-0.5572	



→ The extrapolation procedure is automated in our implementation.



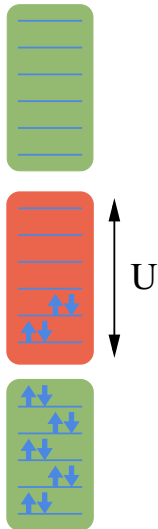
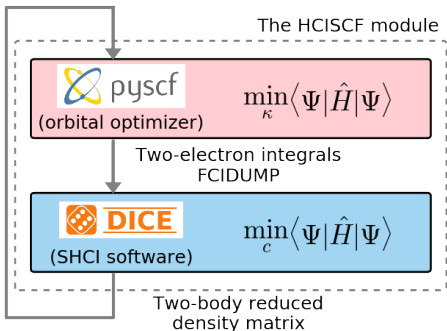
## Orbital Optimization

$E_{\text{SHCI}} = E_0 + E_2$  is not variational wrt. to its parameters:

$$\frac{\partial E}{\partial \kappa} \neq \langle \Psi_0 | \frac{\partial \hat{H}_0}{\partial \kappa} | \Psi_0 \rangle$$

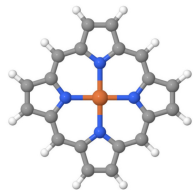
We use the **Lagrangian formalism** to get the analytical gradient:

$$\frac{\partial E}{\partial \kappa} = \langle \Psi_0 | \frac{\partial \hat{H}_0}{\partial \kappa} | \Psi_0 \rangle + \dots (\text{lagrangian terms}) \dots$$



## Fe(Porphyrin)

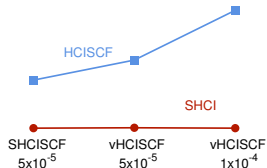
Model cluster for the active site of the heme group  
(hemoglobin, myoglobin, catalase)



Groundstate: experiments suggest a **triplet** ( $^3A_{2g}$  or  $^3E_g$ ) while almost all theoretical works suggest a **quintet** ( $^5A_{1g}$ )

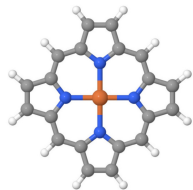
### 1) Effect of the orbitals

	$\epsilon_1$	$E_{\text{HCISCF}}$	$E_{\text{SHCI}}$
SHCISCF orbitals	$5 \times 10^{-5}$	-2245.0178	-2245.0314
vHCISCF orbitals	$5 \times 10^{-5}$	-2245.0121	-2245.0313
vHCISCF orbitals	$1 \times 10^{-4}$	-2244.9980	-2245.0314



→ An easy and quick vHCISCF calculation is enough to obtain converged active space orbitals to perform the SHCI tight calculation.

## Fe(Porphyrin)



Model cluster for the active site of the heme group  
(hemoglobin, myoglobin, catalase)

Groundstate: experiments suggest a **triplet** ( $^3A_{2g}$  or  $^3E_g$ )  
while almost all theoretical works suggest a **quintet** ( $^5A_{1g}$ )

### 2) Large active spaces

(32e,29o) Fe 3d, conjugated  $\pi$  orbitals

(44e,44o)+Fe 4d, 4p<sub>x/y</sub>, N 2p<sub>x/y</sub> orbitals

→ **Accurate theoretical prediction!**

Need to include Fe-N bonding/anti-bonding  
and a second d shell

	$E_{\text{SHCI}}$	$E_{\text{ex}}$
cc-pVDZ (32e,29o)		
$^5A_g$	-0.0314(5)	16.7
$^3B_{1g}$	-0.0049(6)	
cc-pVTZ (32e,29o)		
$^5A_g$	-0.2549(5)	16.4
$^3B_{1g}$	-0.2288(6)	
cc-pVDZ (44e,44o)		
$^5A_g$	-0.1964(9)	-2.0
$^3B_{1g}$	-0.1995(6)	

→ **HCI is fast,  
the limiting step is  
the orbital optimization**

	$^5A_g$	$^3B_{1g}$
$T_{\text{OO}}$	277	264
$T_{\text{Cl}}$	185	147

**Relativistic effects** are important to capture phosphorescence, magnetic spectroscopy, hyperfine constant, g-tensors, ZFS,...

In heavy atoms, it is also comparable to electron correlation.

**Issue:**  $\hat{H} = \hat{H}_{\text{SR}} + \hat{H}_{\text{SOC}}$  where  $\langle N_\alpha N_\beta | \hat{H}_{\text{SOC}} | N'_\alpha N'_\beta \rangle \neq 0$

→ **The size of the Hilbert space of the problem sharply increases**

$$\binom{n/2}{e/2}^2 \text{ versus } \binom{n}{e}$$

**State-of-the-art** is mainly done using a two-step process with QDPT:

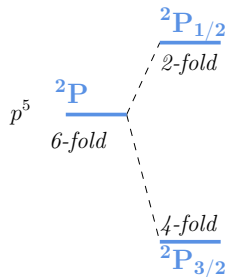
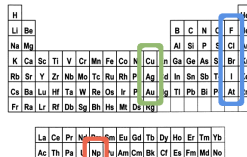
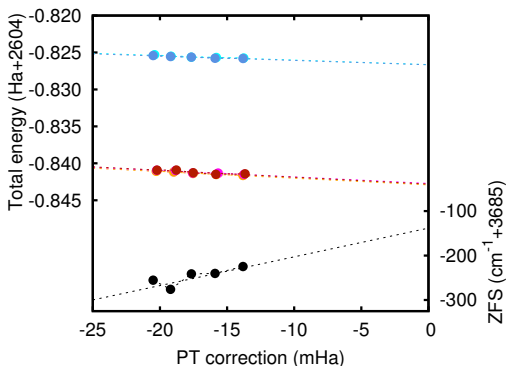
- 1) obtain the eigenstates  $|\Psi_1\rangle, |\Psi_2\rangle, \dots, |\Psi_n\rangle$  of  $\hat{H}_{\text{SR}}$
- 2) let them interact by diagonalizing  $\langle \Psi_i | \hat{H}_{\text{SR}} + \hat{H}_{\text{SOC}} | \Psi_j \rangle$

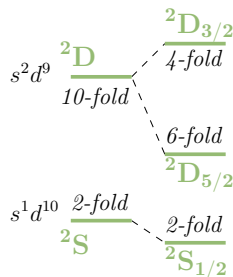
→ **SHCI is used with an X2C Hamiltonian to treat relativistic effects with large active spaces in a one-step procedure on an equal footing with correlation**

HCI retains its ability to effectively discard large parts of the low importance increased-size Hilbert space.

	(7e,4o)	$\approx$ (7e,100o)	$\approx$ (17e,100o)	Ref.
F	405	399		398
Cl	827	806	866	876
Br	3428	3403	3454	3649
I	7021	7277	7487	7755

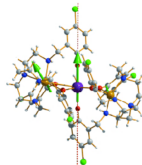
→ Results are converged to better than 1%



[illegible]

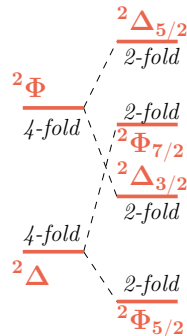
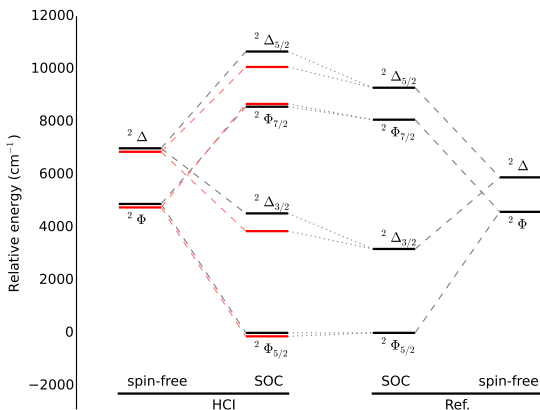
[BM,Sharma JCTC (2017)]

- Insight into complexes containing lanthanide and actinide (**Single Molecular Magnets**)
- SOC splitting is larger than the crystal field



H																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Y
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	N



## - Dynamical correlation

The addition of out-of-active space dynamical correlation can be necessary for some systems. We are working on interfacing SHCI with the recently-introduced MRLCC.

## - Spin-orbit coupling

Four-component DCB Hamiltonian exactly, Quantitative calculations for realistic systems, Ab-initio parameters for model spin-Hamiltonian (analytical response theory).

## - Seniority

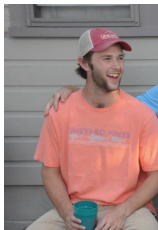
We are investigating the use of seniority-restricted wavefunction in DMRG and SHCI, for strongly-correlated systems.

## - Machine learning

The use of gaussian process techniques could allow us to grow the variational space even more intelligently.



Sandeep Sharma



James E. T. Smith



Adam A. Holmes



<https://sanshar.github.io/Dice/>

**Sandeep Sharma, James E. T. Smith, Adam A. Holmes, Bastien Mussard**

Dice is interfaced with PySCF, a widely-used *ab initio* computational chemistry program implemented in Python and available with pip

## Welcome to Dice's documentation!



*Dice* implements the semistochastic heat bath configuration interaction (SHCI) algorithm for *ab initio* Hamiltonian of a quantum chemical system. Unlike full configuration interaction (FCI), SHCI can be used to treat active spaces containing 30–100 orbitals. SHCI is able to accomplish this by taking advantage of the fact that although the full Hilbert space may be enormous, only a small fraction of the determinants in the space have appreciable coefficients. Compared to other methods in its class SHCI is often not only orders of magnitude faster, it also does not suffer from a serious memory bottleneck that plagues such methods. The resulting algorithm as implemented in *Dice* allows us to treat difficult benchmark systems such as the Chromium dimer and Mn–Salen (a challenging bioinorganic cluster) at a cost that is often an order of magnitude faster than either density matrix renormalization group (DMRG) or full configuration interaction quantum Monte Carlo (FCIQMC). Thus if you are interested in performing multireference calculations with active space containing several tens to hundreds of orbitals, SHCI might be an ideal choice for you.

Contents:

- [Overview](#)
  - [Algorithm](#)