

Accurate computation of g. exc. states
w/ NN.

Pfau et al. 2308: 16848

Key idea

Expand Hilbert space \mathcal{H} to $\mathcal{A}\mathcal{H}^{\otimes k}$
and extend Hamiltonian to act
on this space, to yield the
'total energy'

$\mathcal{A}[|\psi_1\rangle \otimes \dots \otimes |\psi_k\rangle] = \emptyset$ if any of the
component states are the same

\Rightarrow Lowest achievable energy w/ nonzero
state is $E_1 + \dots + E_k$

$\text{① composite state} = \mathcal{A}[|E_1\rangle \otimes \dots |E_k\rangle]$

In practice

Need an ansatz $\Psi(\vec{x}_1, \dots, \vec{x}_k)$ that is
antisymmetric under exchanging the \vec{x}_i .

\Rightarrow Determinants!

$$\Psi(\vec{x}_1, \dots, \vec{x}_k) = \begin{vmatrix} \psi_1(\vec{x}_1) & \dots & \psi_1(\vec{x}_k) \\ \vdots & \ddots & \vdots \\ \psi_k(\vec{x}_1) & \dots & \psi_k(\vec{x}_k) \end{vmatrix}$$

where ψ_i are single- \mathcal{H} ansäze.

Computing the total energy

Let $H_{\text{tot}} = \underbrace{H_1 \oplus \dots \oplus H_K}_{\substack{\text{Hamiltonian} \\ (\text{or any other op.}) \\ \text{acting on } i^{\text{th}} \\ \text{component of } \mathcal{H}^{\otimes n}}}$

- obvious choice if the $|\psi_i\rangle$ are orthogonal
- symmetric under exchanging HS copies
 \Rightarrow maps $\mathcal{H}^{\otimes n} \rightarrow \mathcal{H}^{\otimes n}$.

Need to evaluate Rayleigh quotient

$$E_{\text{tot}} = \frac{\langle \Psi | H_{\text{tot}} | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

- again, if $|\psi_i\rangle$ orthogonal, $E_{\text{tot}} = E_1 + \dots + E_K$ where E_i are individual Rayleigh q.

$$\bullet |\Psi\rangle = \sum_{\sigma \in S_K} (-1)^\sigma |\psi_{\sigma_1}\rangle \otimes \dots \otimes |\psi_{\sigma_K}\rangle$$

$$\rightarrow \langle \Psi | \Psi \rangle = \sum_{\sigma, \pi} (-1)^\sigma (-1)^\pi \langle \psi_{\sigma_1} | \psi_{\pi_1} \rangle \dots \langle \psi_{\sigma_K} | \psi_{\pi_K} \rangle$$

$$\begin{aligned} &= K! \sum_S (-1)^S \langle \psi_1 | \psi_{S_1} \rangle \dots \langle \psi_K | \psi_{S_K} \rangle \\ &= K! \left| \begin{array}{c} \langle \psi_1 | \psi_1 \rangle \dots \langle \psi_1 | \psi_K \rangle \\ \vdots \\ \langle \psi_K | \psi_1 \rangle \dots \langle \psi_K | \psi_K \rangle \end{array} \right| \end{aligned}$$

$$\rightarrow \langle \Psi | H_i | \Psi \rangle = \sum_{\sigma, \pi} (-1)^\sigma (-1)^\pi \langle \psi_{\sigma_1} | H_i | \psi_{\pi_1} \rangle \dots \langle \psi_{\sigma_K} | \psi_{\pi_K} \rangle$$

$$\begin{aligned} &= \frac{K!}{K} \sum_S (-1)^S \left[\langle \psi_1 | H_i | \psi_{S_1} \rangle \dots \langle \psi_K | \psi_{S_K} \rangle + \right. \\ &\quad \langle \psi_1 | \psi_{S_1} \rangle \langle \psi_2 | H_i | \psi_{S_2} \rangle \dots \langle \psi_K | \psi_{S_K} \rangle + \dots \\ &\quad \left. \langle \psi_1 | \psi_{S_1} \rangle \dots \langle \psi_K | H_i | \psi_{S_K} \rangle \right] \end{aligned}$$

this is wrong in the paper

same for all $\langle \Psi | H_i | \Psi \rangle$

$$\therefore \langle \Psi | H | \Psi \rangle = k! \sum_i \underbrace{\left| \begin{array}{c} \langle \psi_1 | \psi_1 \rangle \cdots \langle \psi_1 | H | \psi_i \rangle \cdots \langle \psi_1 | \psi_k \rangle \\ \vdots \quad \vdots \quad \vdots \\ \langle \psi_k | \psi_1 \rangle \cdots \langle \psi_k | H | \psi_i \rangle \cdots \langle \psi_k | \psi_k \rangle \end{array} \right|}_{\underline{h_i}}$$

- Define the matrix

$$\underline{H} = \begin{bmatrix} \langle \psi_1 | H | \psi_1 \rangle & \cdots & \langle \psi_1 | H | \psi_k \rangle \\ \vdots & \ddots & \vdots \\ \langle \psi_k | H | \psi_1 \rangle & \cdots & \langle \psi_k | H | \psi_k \rangle \end{bmatrix}$$

$\Rightarrow \underline{h_i}$ is $\underline{\Sigma}$ with the i^{th} column replaced by the i^{th} column of \underline{H} .

\Rightarrow Cramer's rule:

$$\frac{\det \underline{h_i}}{\det \underline{\Sigma}} = i^{\text{th}} \text{ component of solution of } \underline{\Sigma} \underline{b} = \underline{H} \cdot \underline{i}$$

$$= (\underline{\Sigma}^{-1} \underline{H} \cdot \underline{i})_i = (\underline{\Sigma}^{-1} \underline{H})_{ii}$$

* sum over i : $E_{\text{tot}} = \text{tr}(\underline{\Sigma}^{-1} \underline{H})$

- important: E_{tot} only depends on the space spanned by $|\psi_1\rangle \cdots |\psi_k\rangle$:

Let $[\psi'_1 \cdots \psi'_k] = [\psi_1 \cdots \psi_k] \underline{A}$
where A is any invertible matrix

$$\therefore S' = A^+ S A$$

$$H' = A^+ H A$$

$$\therefore E'_{\text{tot}} = \text{tr}(A^{-1} S^{-1} (A^+)^{-1} A^+ H A) = \text{tr}(S^{-1} H A A^{-1})$$

$$= E_{\text{tot}}$$

Computing individual energies + observables

So, E_{tot} is minimised by any K states that span the same space as $|E_1\rangle \dots |E_K\rangle$.

- but $|\psi_1\rangle \dots$ are not necessarily these states
- not necessarily orthogonal even!

Since $|E_i\rangle$ are orthonormal, if $|\psi_i\rangle = |E_i\rangle$,

$$\underline{\underline{S}} = \underline{\underline{1}}, \quad \underline{\underline{H}} = \underline{\underline{S}}^{-1} \underline{\underline{H}} = \text{diag}(E_1, \dots, E_K) =: \underline{\underline{E}}$$

Otherwise, $|\psi_j\rangle = |E_i\rangle A_{ij}$ ($\det \underline{\underline{A}} \neq 0$)

$$(\underline{\underline{S}}^{-1} \underline{\underline{H}})_{ij} = \underline{\underline{A}}^{-1} \underline{\underline{E}} \underline{\underline{A}}$$

∴ Eigenvalues of $S^{-1}H$ are E_1, \dots, E_K in any basis

However, eigenvectors not uniquely defined

• if $\underline{\underline{A}}$ is valid, so is $\underline{\underline{A}}' = \underline{\underline{S}} \underline{\underline{A}}$

∴ numerically computed eigenvectors $\tilde{\underline{\underline{A}}}$ recover $|E_i\rangle$ up to arbitrary prefactor σ_i :

$$|\psi_j\rangle = |E_i\rangle \sigma_i \tilde{A}_{ij}$$

Observable matrix elements

Can play the same game for any op. \mathcal{O} :

$$(\underline{\underline{S}}^{-1} \underline{\underline{\mathcal{O}}})_{ij} = \underline{\underline{A}}^{-1} [\langle E_i | \mathcal{O} | E_j \rangle] \underline{\underline{A}} \quad \tilde{\underline{\underline{\mathcal{A}}}}$$

Using the eigenvectors from before, can get

$$\tilde{\underline{\underline{\mathcal{O}}}} = \tilde{\underline{\underline{A}}} (\underline{\underline{S}}^{-1} \underline{\underline{\mathcal{O}}})_{ij} \tilde{\underline{\underline{A}}}^{-1} : \quad \tilde{O}_{ij} = \frac{\sigma_i}{\sigma_j} \langle E_i | \mathcal{O} | E_j \rangle$$

Diagonal matrix elements ✓

Offdiagonal: can get magnitudes at least:

$$\tilde{O}_{ij} \tilde{O}_{ji} = |\langle E_i | \mathcal{O} | E_j \rangle|^2$$

VMC estimators

We need a VMC estimator for $\underline{\underline{S}}^{-1} \underline{\underline{H}}$

For each sample $\vec{x} = (\vec{x}_1, \dots, \vec{x}_k)$, consider matrices

$$\underline{\underline{\Psi}}(\vec{x}) = \begin{pmatrix} \psi_1(\vec{x}_1) & \cdots & \psi_k(\vec{x}_1) \\ \vdots & \ddots & \vdots \\ \psi_1(\vec{x}_k) & \cdots & \psi_k(\vec{x}_k) \end{pmatrix}$$

$$\underline{\underline{H}\Psi}(\vec{x}) = \begin{pmatrix} H\psi_1(\vec{x}_1) & \cdots & H\psi_k(\vec{x}_1) \\ \vdots & \ddots & \vdots \\ H\psi_1(\vec{x}_k) & \cdots & H\psi_k(\vec{x}_k) \end{pmatrix}$$

- If $|\psi_i\rangle \dots$ are eigenstates of H ,
 $\underline{\underline{\Psi}} = \underline{\underline{\Psi}} \text{ diag}(E_1, \dots, E_k) + \vec{x}$ } promising estimator
- $\underline{\underline{\Psi}}^{-1}(\underline{\underline{H}\Psi})$ transforms like $\underline{\underline{S}}^{-1} \underline{\underline{H}}$ under } candidate basis transformation

Let's check:

$$\bullet \mathbb{E}_{\vec{x} \sim \Psi^2} [\underline{\underline{\Psi}}^{-1}(\underline{\underline{H}\Psi})]_{ij} = \frac{1}{\langle \underline{\underline{\Psi}} | \underline{\underline{\Psi}} \rangle} \sum_{\vec{x}} \underline{\underline{\Psi}}^* \det \underline{\underline{\Psi}} \times (\underline{\underline{\Psi}}^{-1})_{i\epsilon} (\underline{\underline{H}\Psi})_{\epsilon j}$$

This derivation
is wrong in the paper

$$= \frac{1}{\langle \underline{\underline{\Psi}} | \underline{\underline{\Psi}} \rangle} \sum_{\vec{x}} \underline{\underline{\Psi}}^* (\text{adj } \underline{\underline{\Psi}})_{i\epsilon} (\underline{\underline{H}\Psi})_{\epsilon j}$$

$$= \frac{1}{\langle \underline{\underline{\Psi}} | \underline{\underline{\Psi}} \rangle} \sum_{\vec{x}} \underline{\underline{\Psi}}^* \begin{vmatrix} \psi_1(\vec{x}_1) & \cdots & H\psi_j(\vec{x}_1) & \cdots & \psi_k(\vec{x}_1) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \psi_1(\vec{x}_k) & \cdots & H\psi_j(\vec{x}_k) & \cdots & \psi_k(\vec{x}_k) \end{vmatrix}$$

$$= \frac{\langle \underline{\underline{\Psi}} | \tilde{\underline{\underline{\Psi}}} \rangle}{\langle \underline{\underline{\Psi}} | \underline{\underline{\Psi}} \rangle}$$

where $|\tilde{\underline{\underline{\Psi}}}\rangle = \mathcal{A}(|\psi_1\rangle \otimes \cdots \otimes |\psi_{i-1}\rangle \otimes H|\psi_i\rangle \otimes \cdots \otimes |\psi_k\rangle)$

- But also,

$$\frac{\langle \Psi | \tilde{\Psi} \rangle}{\langle \Psi | \Psi \rangle} = \frac{1}{\det \Sigma} \begin{vmatrix} \langle \psi_1 | \psi_1 \rangle & \cdots & \langle \psi_1 | \psi_k \rangle & \cdots & \langle \psi_1 | \psi_e \rangle \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \langle \psi_k | \psi_1 \rangle & \cdots & \langle \psi_k | \psi_j \rangle & \cdots & \langle \psi_k | \psi_e \rangle \end{vmatrix}$$

Cramer's

$$= \left(\begin{matrix} S^{-1} H \end{matrix} \right)_{ij} = (S^{-1} H)_{ij}.$$

$$\therefore \mathbb{E}_{X \sim \mathbb{H}^2} [\underline{\Psi}^{-1} (\underline{H} \underline{\Psi})] = \underline{\Sigma}^{-1} \underline{H} \quad \text{as a matrix.}$$

Variational energy for optimisation is the trace.

Application

Low-energy eigenstates of molecules

Relevant questions:

- which HF/atomic orbitals are filled in each state?
- compute DM in atomic/molecule orbital basis
- symmetry quantum numbers
- which states are dark? (Fig. 2)
 - relevant quantity: off-diagonal matrix element of dipole moment.

Each ψ_1, \dots, ψ_K is either Fermi/Peierls

↳ general idea of both:

- compute many-body deformed orbitals
- use them in "Slater" determinants
- Q: might it be possible to combine these + many-eigenstate determinants?
is it a good idea?
- Q: how much can ψ_1, \dots, ψ_K share?

Initialisation:

- for single FN/ΨF, important to start from a good guess (aka. HF)
- here, need K lin. independent initial states
 - "ordered": K combination of orbitals w/ lowest HF energy
 - "random": random selection of orbitals from some active space.

Overall verdict:

- works well for most problems
- much better than orthogonal penalties

Q: is it b/c subspace has a simpler structure than indiv. eigenstates?