



Laboratoire de Chimie et Physique Quantiques



LCT

Laboratoire de Chimie Théorique



Targeting Real chemical accuracy at the EXascale

Iterative CI wavefunction optimization using a similarity-transformed Hamiltonian and VMC

Abdallah Ammar¹, Emmanuel Giner² & Anthony Scemama¹

¹LCPQ, IRSAMC, UPS/CNRS, Toulouse, France

²LCT, Sorbonne Université/CNRS, Paris, France

January 13, 2021



1 Motivation

2 Variational scheme

3 Transcorrelated approach

4 Iterative symmetric dressing

5 Reduce statistical noise

6 Conclusion



CI wavefunction

→ Configuration Interaction (CI) wavefunction: linear combination of Slater determinants $\{D_I\}$ constructed from sets of orthonormal spin orbitals

$$\Phi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{I=1}^{N_{\text{det}}} C_I D_I(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

→ For large basis sets, full CI computation is not realizable. We improve the truncated wavefunction by adding a Jastrow factor:

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \underbrace{\sum_{I=1}^{N_{\text{det}}} C_I D_I(\mathbf{r}_1, \dots, \mathbf{r}_N)}_{\Phi(\mathbf{r}_1, \dots, \mathbf{r}_N)} \exp[J]$$

→ goal: optimize the determinantal part $\{C_I\}$ in the presence of Jastrow factor for large N_{det}

- 1 Motivation
- 2 Variational scheme
- 3 Transcorrelated approach
- 4 Iterative symmetric dressing
- 5 Reduce statistical noise
- 6 Conclusion



Generalized eigenvalue problem

→ The best CI coefficients can be obtained by minimizing the energy. This leads to a generalized matrix eigenvalue equation:

$$\mathbf{H}\mathbf{C} = \mathbf{E}\mathbf{S}\mathbf{C} \quad \text{where} \quad \begin{cases} \mathbf{H}_{IK} = \langle D_I \exp[J] | \hat{H} | D_K \exp[J] \rangle \\ \mathbf{S}_{IK} = \langle D_I \exp[J] | D_K \exp[J] \rangle \end{cases}$$

- ! Variational problem (\mathbf{H} is symmetric & \mathbf{S} is positive semidefinite)
- ! $3N$ -dimensional integrals

→ Quantum Monte Carlo methods

Generalized eigenvalue problem

→ Variational Monte Carlo (VMC)

$$\begin{cases} \mathbf{H}_{IK} \approx \left\langle \frac{D_I e^J}{\Psi} \hat{H} \frac{D_K e^J}{\Psi} \right\rangle_{\Psi^2} \\ \mathbf{S}_{IK} \approx \left\langle \frac{D_I e^J}{\Psi} \frac{D_K e^J}{\Psi} \right\rangle_{\Psi^2} \end{cases}$$

$$\langle X \rangle_{\Psi^2} = \frac{1}{N_{\text{config}}} \sum_{m=1}^{N_{\text{config}}} X(\mathbf{R}_m) \quad \text{drawn with } \Psi^2$$

■ large matrices to sample $\sim N_{\text{det}}^2$

■ statistical noise

↳ impractical/poor optimization for large N_{det}

- 1 Motivation
- 2 Variational scheme
- 3 Transcorrelated approach**
- 4 Iterative symmetric dressing
- 5 Reduce statistical noise
- 6 Conclusion



Transcorrelated formalism

→ A symmetric pair correlation factor τ is incorporated in the Hamiltonian

$$\hat{H}_{\text{TC}} \equiv e^{-\hat{\tau}} \hat{H} e^{\hat{\tau}}$$

→ The similarity-transformed Hamiltonian \hat{H}_{TC} and \hat{H} share the same spectra:

$$\hat{H}\Psi = E\Psi \iff \boxed{\hat{H}_{\text{TC}}\Phi = E\Phi} \quad \text{where} \quad \begin{cases} \Psi \equiv \Phi e^{\tau} \\ \Phi = \sum_{i=1}^{\infty} \phi_i \end{cases}$$

→ For a two-body correlation factor, the effective TC Hamiltonian:

$$\begin{aligned} \hat{H}_{\text{TC}} &= \hat{H} + [\hat{H}, \hat{\tau}] + \frac{1}{2} [[\hat{H}, \hat{\tau}], \hat{\tau}] \\ &= \hat{H} + \underbrace{\hat{K}_{12}}_{\text{non-Hermitian 2-body}} + \underbrace{\hat{L}_{123}}_{\text{3-body}} \end{aligned}$$

Transcorrelated formalism

→ Advantages:

- | N -body integrals → 3-body integrals
- | The Coulomb singularity $1/r_{ij}$ can be explicitly removed → improve convergence
- | Post-Hartree-Fock (CI, CC, ...) methods can be combined with TC approach

→ Numerical difficulties:

- | **three-body** term
- | The TC Hamiltonian is **non-Hermitian**

→ We propose an iterative scheme to overcome these difficulties

- 1 Motivation
- 2 Variational scheme
- 3 Transcorrelated approach
- 4 Iterative symmetric dressing**
- 5 Reduce statistical noise
- 6 Conclusion



Diagonal dressing

We start with a CI WF $\Phi^{(0)} = \sum_{I=1}^{N_{\text{det}}} c_I^{(0)} D_I$, and a fixed Jastrow factor $\exp[\tau]$

→ **Goal:** optimize the coefficients by solving the TC eigen-problem in the $\{D_I\}$ basis:

$$\hat{H}_{\text{TC}} \Phi = E_{\text{TC}} \Phi \Rightarrow \sum_{K=1}^{N_{\text{det}}} \langle D_I | \hat{H}_{\text{TC}} - \hat{H} + \hat{H} | D_K \rangle c_K^{(i)} = E_{\text{TC}} c_I^{(i)}$$

Diagonal dressing

We start with a CI WF $\Phi^{(0)} = \sum_{I=1}^{N_{\text{det}}} C_I^{(0)} D_I$, and a fixed Jastrow factor $\exp[\tau]$

→ **Goal:** optimize the coefficients by solving the TC eigen-problem in the $\{D_I\}$ basis:

$$\hat{H}_{\text{TC}} \Phi = E_{\text{TC}} \Phi \Rightarrow \sum_{K=1}^{N_{\text{det}}} \langle D_I | \hat{H}_{\text{TC}} - \hat{H} + \hat{H} | D_K \rangle C_K^{(i)} = E_{\text{TC}} C_I^{(i)}$$

$$\sum_{K \neq I} \mathbf{H}_{IK} C_K^{(i)} + \left(\mathbf{H}_{II} + \frac{1}{C_I^{(i-1)}} \langle D_I | \hat{H}_{\text{TC}} - \hat{H} | \Phi^{(i-1)} \rangle \right) C_I^{(i)} \approx E_{\text{TC}} C_I^{(i)}$$

dressing elements

$$\begin{pmatrix} x & & \\ & \ddots & \\ & & x \end{pmatrix}$$

Diagonal dressing

We start with a CI WF $\Phi^{(0)} = \sum_{I=1}^{N_{\text{det}}} C_I^{(0)} D_I$, and a fixed Jastrow factor $\exp[\tau]$

→ **Goal:** optimize the coefficients by solving the TC eigen-problem in the $\{D_I\}$ basis:

$$\hat{H}_{\text{TC}} \Phi = E_{\text{TC}} \Phi \Rightarrow \sum_{K=1}^{N_{\text{det}}} \langle D_I | \hat{H}_{\text{TC}} - \hat{H} + \hat{H} | D_K \rangle C_K^{(i)} = E_{\text{TC}} C_I^{(i)}$$

$$\sum_{K \neq I} \mathbf{H}_{IK} C_K^{(i)} + \left(\mathbf{H}_{II} + \frac{1}{C_I^{(i-1)}} \langle D_I | \hat{H}_{\text{TC}} - \hat{H} | \Phi^{(i-1)} \rangle \right) C_I^{(i)} \approx E_{\text{TC}} C_I^{(i)}$$

dressing elements

$$\begin{pmatrix} x & & \\ & \ddots & \\ & & x \end{pmatrix}$$

→ We have to build the diagonal dressing matrix $\Delta^{(i-1)}$:

$$\Delta_{IK}^{(i-1)} = \begin{cases} \frac{1}{C_I^{(i-1)}} \langle D_I | \hat{H}_{\text{TC}} - \hat{H} | \Phi^{(i-1)} \rangle & \text{if } I = K \\ 0 & \text{otherwise} \end{cases}$$

Column dressing

→ It is more stable to dress with the elements

$$\Gamma_{iL}^{(i-1)} = \frac{1}{C_L^{(i-1)}} \langle D_i | \hat{H}_{\text{TC}} - \hat{H} | \Phi^{(i-1)} \rangle$$

$$\begin{pmatrix} x & x & \cdots & x \end{pmatrix} \text{ row } L$$

where L corresponds to the largest coefficients $C_L^{(i-1)}$

Column dressing

→ It is more stable to dress with the elements

$$\Gamma_{iL}^{(i-1)} = \frac{1}{C_L^{(i-1)}} \langle D_i | \hat{H}_{TC} - \hat{H} | \Phi^{(i-1)} \rangle$$
$$\begin{pmatrix} & x & & \\ x & x & \cdots & x \\ & \vdots & & \\ & x & & \end{pmatrix} \text{ row } L$$

where L corresponds to the largest coefficients $C_L^{(i-1)}$

→ The dressing matrix $\Delta^{(i-1)}$ is chosen to be symmetric by construction:

$$\Delta_{LI}^{(i-1)} = \Delta_{iL}^{(i-1)} = \Gamma_{iL}^{(i-1)}$$
$$\Delta_{KI}^{(i-1)} = \Delta_{iK}^{(i-1)} = 0 \quad \text{for } K \neq L$$

Column dressing

→ It is more stable to dress with the elements

$$\Gamma_{iL}^{(i-1)} = \frac{1}{C_L^{(i-1)}} \langle D_i | \hat{H}_{\text{TC}} - \hat{H} | \Phi^{(i-1)} \rangle$$
$$\begin{pmatrix} & x & & \\ x & x & \cdots & x \\ & \vdots & & \\ & x & & \end{pmatrix} \text{ row } L$$

where L corresponds to the largest coefficients $C_L^{(i-1)}$

→ The dressing matrix $\Delta^{(i-1)}$ is chosen to be symmetric by construction:

$$\Delta_{Li}^{(i-1)} = \Delta_{iL}^{(i-1)} = \Gamma_{iL}^{(i-1)}$$
$$\Delta_{Ki}^{(i-1)} = \Delta_{iK}^{(i-1)} = 0 \quad \text{for } K \neq L$$

→ An extra term is introduced in the diagonal element to cancel the double counting:

$$\Delta_{LL}^{(i-1)} = 2\Gamma_{LL}^{(i-1)} - \frac{1}{C_L^{(i-1)}} \sum_{K=1}^{N_{\text{det}}} \Gamma_{KL}^{(i-1)} C_K^{(i-1)}$$

Iterative dressing algorithm

① choose a CI WF: $\Phi^{(0)} = \sum_{I=1}^{N_{\text{det}}} C_I^{(0)} D_I$ & a Jastrow factor $\exp(\tau)$

② **with VMC**, evaluate:

➔ the dressing elements $\langle D_I | \hat{H}_{\text{TC}} - \hat{H} | \Phi^{(i-1)} \rangle$

➔ the **variational** energy $E^{(i-1)} = \frac{\langle \Phi^{(i-1)} e^\tau | \hat{H} | e^\tau \Phi^{(i-1)} \rangle}{\langle \Phi^{(i-1)} e^\tau | e^\tau \Phi^{(i-1)} \rangle}$

③ dress the Hamiltonian matrix \mathbf{H} with the symmetric matrix $\Delta^{(i-1)}$

④ apply Davidson to obtain the new ground state $\Phi^{(i)}$ and $E_{\text{TC}}^{(i)}$

⑤ go back to ②

Iterative dressing algorithm

① choose a CI WF: $\Phi^{(0)} = \sum_{I=1}^{N_{\text{det}}} C_I^{(0)} D_I$ & a Jastrow factor $\exp(\tau)$

② **with VMC**, evaluate:

➔ the dressing elements $\langle D_I | \hat{H}_{\text{TC}} - \hat{H} | \Phi^{(i-1)} \rangle$

➔ the **variational** energy $E^{(i-1)} = \frac{\langle \Phi^{(i-1)} e^\tau | \hat{H} | e^\tau \Phi^{(i-1)} \rangle}{\langle \Phi^{(i-1)} e^\tau | e^\tau \Phi^{(i-1)} \rangle}$

③ dress the Hamiltonian matrix \mathbf{H} with the symmetric matrix $\Delta^{(i-1)}$

④ apply Davidson to obtain the new ground state $\Phi^{(i)}$ and $E_{\text{TC}}^{(i)}$

⑤ go back to ②

★ all steps are deterministic except ②

★ the TC energy E_{TC} is not variational. *per contra* the VMC energy E is

Iterative dressing algorithm: advantages

- after few iterations ($\sim 2 - 4$), we converge to the solution
- we need to sample only N_{det} elements instead of N_{det}^2
- handle the **non-Hermiticity** of the TC eigenproblem thanks to the **symmetric dressing**
- VMC calculation allows to handle **three-body integrals** and evaluate a **variational energy** (instead of the TC energy)
- reduced statistical errors:

$$\begin{aligned}\langle D_I | \hat{H}_{\text{TC}} - \hat{H} | D_K \rangle &= \langle D_I e^{-\tau} | \hat{T} + \hat{V} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} + \hat{V} | D_K \rangle \\ &= \langle D_I e^{-\tau} | \hat{T} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} | D_K \rangle\end{aligned}$$

for large distances $\exp(\tau) \sim \exp(-\tau) \sim 1 \Rightarrow$ large fluctuations occur only when electrons are close, which has a relatively low probability

Iterative dressing algorithm: advantages

- after few iterations ($\sim 2 - 4$), we converge to the solution
- we need to sample only N_{det} elements instead of N_{det}^2
- handle the **non-Hermiticity** of the TC eigenproblem thanks to the **symmetric dressing**
- VMC calculation allows to handle **three-body integrals** and evaluate a **variational energy** (instead of the TC energy)
- reduced statistical errors:

$$\begin{aligned}\langle D_I | \hat{H}_{\text{TC}} - \hat{H} | D_K \rangle &= \langle D_I e^{-\tau} | \hat{T} + \hat{V} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} + \hat{V} | D_K \rangle \\ &= \langle D_I e^{-\tau} | \hat{T} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} | D_K \rangle\end{aligned}$$

for large distances $\exp(\tau) \sim \exp(-\tau) \sim 1 \Rightarrow$ large fluctuations occur only when electrons are close, which has a relatively low probability

Iterative dressing algorithm: advantages

- after few iterations ($\sim 2 - 4$), we converge to the solution
- we need to sample only N_{det} elements instead of N_{det}^2
- handle the **non-Hermiticity** of the TC eigenproblem thanks to the **symmetric dressing**
- VMC calculation allows to handle **three-body integrals** and evaluate a **variational energy** (instead of the TC energy)
- reduced statistical errors:

$$\begin{aligned}\langle D_I | \hat{H}_{\text{TC}} - \hat{H} | D_K \rangle &= \langle D_I e^{-\tau} | \hat{T} + \hat{V} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} + \hat{V} | D_K \rangle \\ &= \langle D_I e^{-\tau} | \hat{T} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} | D_K \rangle\end{aligned}$$

for large distances $\exp(\tau) \sim \exp(-\tau) \sim 1 \Rightarrow$ large fluctuations occur only when electrons are close, which has a relatively low probability

Iterative dressing algorithm: advantages

- after few iterations ($\sim 2 - 4$), we converge to the solution
- we need to sample only N_{det} elements instead of N_{det}^2
- handle the **non-Hermiticity** of the TC eigenproblem thanks to the **symmetric dressing**
- VMC calculation allows to handle **three-body integrals** and evaluate a **variational energy** (instead of the TC energy)
- reduced statistical errors:

$$\begin{aligned}\langle D_I | \hat{H}_{\text{TC}} - \hat{H} | D_K \rangle &= \langle D_I e^{-\tau} | \hat{T} + \hat{V} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} + \hat{V} | D_K \rangle \\ &= \langle D_I e^{-\tau} | \hat{T} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} | D_K \rangle\end{aligned}$$

for large distances $\exp(\tau) \sim \exp(-\tau) \sim 1 \Rightarrow$ large fluctuations occur only when electrons are close, which has a relatively low probability

Iterative dressing algorithm: advantages

- after few iterations ($\sim 2 - 4$), we converge to the solution
- we need to sample only N_{det} elements instead of N_{det}^2
- handle the **non-Hermiticity** of the TC eigenproblem thanks to the **symmetric dressing**
- VMC calculation allows to handle **three-body integrals** and evaluate a **variational energy** (instead of the TC energy)
- reduced statistical errors:

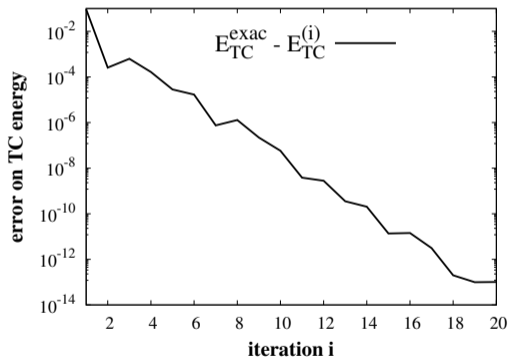
$$\begin{aligned}\langle D_I | \hat{H}_{\text{TC}} - \hat{H} | D_K \rangle &= \langle D_I e^{-\tau} | \hat{T} + \hat{V} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} + \hat{V} | D_K \rangle \\ &= \langle D_I e^{-\tau} | \hat{T} | e^{\tau} D_K \rangle - \langle D_I | \hat{T} | D_K \rangle\end{aligned}$$

for large distances $\exp(\tau) \sim \exp(-\tau) \sim 1 \Rightarrow$ large fluctuations occur only when electrons are close, which has a relatively low probability

Proof of concept

→ Jastrow inspired by Range-Separated DFT¹: $\tau_{\mu} = \sum_{i < j} \left[\frac{r_{ij} (1 - \text{erf}(\mu r_{ij}))}{2} - \frac{\exp[-(\mu r_{ij})^2]}{2\sqrt{\pi}\mu} \right]$

Benchmark: Be (cc-pcvdz)

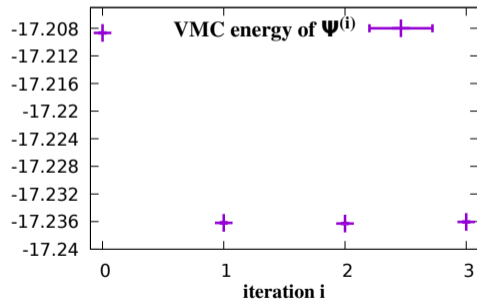


¹Emmanuel Giner, *J. Chem. Phys.*, **154**, 2021

Application: H₂O (cc-pvdz_ecp_bfd): $N_{\text{det}} \sim 71\,000$

→ A simple Jastrow factor $\tau = \sum_{i<j}^N \frac{a r_{ij}}{1 + b r_{ij}} - \sum_{A=1}^M \sum_{i=1}^N \left(\frac{\alpha_A r_{iA}}{1 + \alpha_A r_{iA}} \right)^2$

CI energy: $E_{\Phi} = -17.16096$ (no-Jastrow)

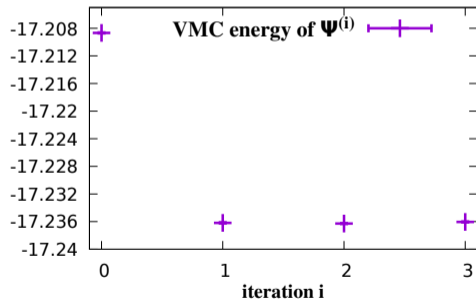


$$E^{(3)} - E^{(0)} \approx 27mE_h$$

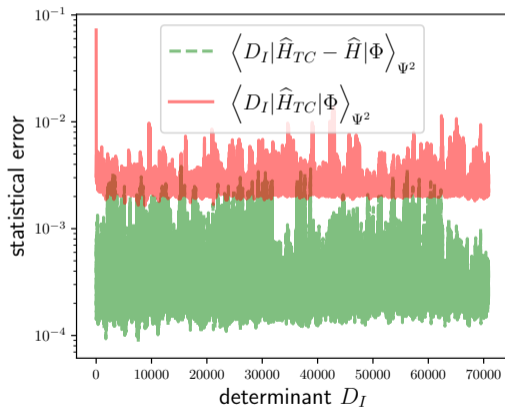
Application: H₂O (cc-pvdz_ecp_bfd): $N_{\text{det}} \sim 71\,000$

→ A simple Jastrow factor $\tau = \sum_{i<j}^N \frac{a r_{ij}}{1 + b r_{ij}} - \sum_{A=1}^M \sum_{i=1}^N \left(\frac{\alpha_A r_{iA}}{1 + \alpha_A r_{iA}} \right)^2$

CI energy: $E_\Phi = -17.16096$ (no-Jastrow)



$$E^{(3)} - E^{(0)} \approx 27mE_h$$



$$\langle \text{error} \rangle \approx 8 \langle \text{error} \rangle$$

- 1 Motivation
- 2 Variational scheme
- 3 Transcorrelated approach
- 4 Iterative symmetric dressing
- 5 Reduce statistical noise**
- 6 Conclusion



Modified dressing elements

→ For an arbitrary choice of Jastrow factor (arbitrary \hat{H}_{TC}), we have:

$$\begin{aligned}\langle D_I | \hat{H}_{TC} | D_J \rangle &= \langle D_I | \hat{H}_{TC} - \hat{H}_\mu + \hat{H}_\mu - \hat{H} + \hat{H} | D_J \rangle \\ &= \underbrace{\langle D_I | \hat{H} | D_J \rangle}_{\text{matrix to be dressed}} + \underbrace{\langle D_I | \hat{H}_\mu - \hat{H} | D_J \rangle}_{\sim \text{analytic elements}} + \underbrace{\langle D_I | \hat{H}_{TC} - \hat{H}_\mu | D_J \rangle}_{\text{VMC elements}}\end{aligned}$$

→ New dressing elements:

$$\langle D_I | \hat{H}_{TC} - \hat{H} | \Phi \rangle \rightarrow \boxed{\langle D_I | \hat{H}_{TC} - \hat{H}_\mu | \Phi \rangle} + \langle D_I | \hat{H}_\mu - \hat{H} | \Phi \rangle$$

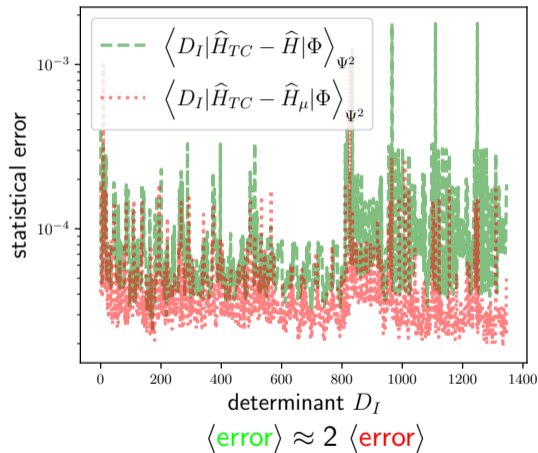
→ Local energies with Jastrow factor are more correlated:

$$\left| \frac{\hat{H}_{TC} \Phi}{\Phi} - \frac{\hat{H}_\mu \Phi}{\Phi} \right| < \left| \frac{\hat{H}_{TC} \Phi}{\Phi} - \frac{\hat{H} \Phi}{\Phi} \right| \Rightarrow \text{reduced statistical errors}$$

illustration: Be (cc-pcvdz)

$$\blacktriangleright \tau_{TC} = \sum_{i < j} \frac{a r_{ij}}{1 + b r_{ij}} - \sum_{A,i} \left(\frac{\alpha_A r_{iA}}{1 + \alpha_A r_{iA}} \right)^2$$

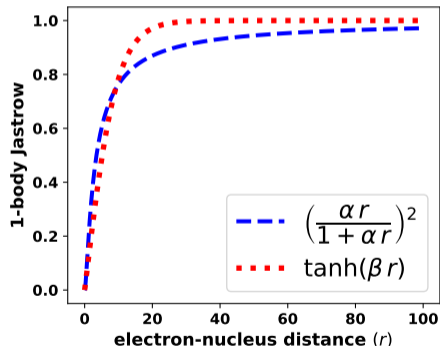
$$\blacktriangleright \tau_{\mu} = \sum_{i < j} \left[\frac{r_{ij} (1 - \text{erf}(\mu r_{ij}))}{2} - \frac{\exp[-(\mu r_{ij})^2]}{2\sqrt{\pi}\mu} \right]$$



1-body Jastrow

→ we add a 1-body Jastrow to avoid unfavorable effect of the two-body Jastrow factor τ_μ that changes the charge density:

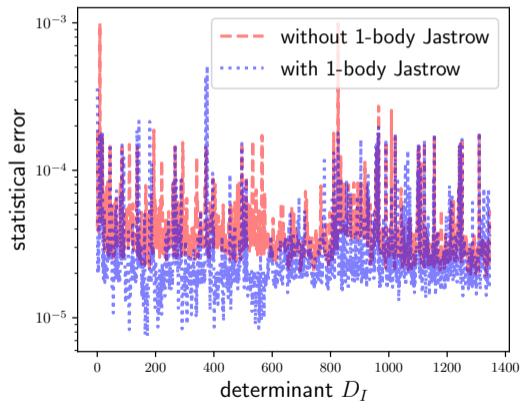
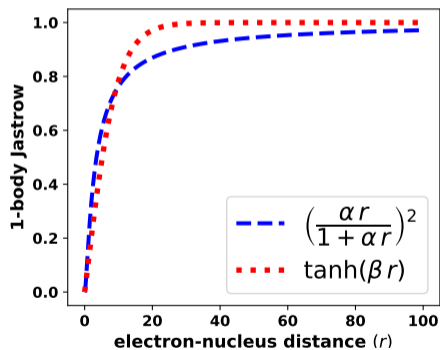
$$\tau_{mu} \rightarrow \tau_{mu} - \sum_{A,i} \tanh(\beta_A r_{iA})$$



1-body Jastrow

→ we add a 1-body Jastrow to avoid unfavorable effect of the two-body Jastrow factor τ_μ that changes the charge density:

$$\tau_{mu} \rightarrow \tau_{mu} - \sum_{A,i} \tanh(\beta_A r_{iA})$$



$$\langle \text{error} \rangle \approx 1.3 \langle \text{error} \rangle$$

Conclusion

→ iterative symmetric dressing within TC approach:

- number of elements to sample in VMC: $N_{\text{det}}^2 \rightarrow N_{\text{det}}$

- addresses the **non-Hermiticity** of the TC Hamiltonian & **three-body** terms

- fast convergence ($\sim 2 - 3$ iterations)

→ employ \hat{H}_μ in the dressing elements calculation allows to reduce the statistical noise

- we can improve the results by adding 1-body term to τ_μ

→ ongoing work:

- application for larger CI expansion ($N_{\text{det}} \sim 10^6 - 10^7$)

- develop a compact representation for dressing vector

Acknowledgments

Anthony Scemama & Emmanuel Giner



This project has received funding from the European Unions Horizon 2020 - Research and Innovation program - under grant agreement no. 952165.

Thank you for your attention

