

Optimization of quantum Monte Carlo wave functions and calculation of pair densities

Julien Toulouse^{*,†}, Cyrus J. Umrigar^{*}, Roland Assaraf[†]

* Theory Center and Physics Department, **Cornell University**, Ithaca, New York, USA.

† Laboratoire de Chimie Théorique, **Université Pierre et Marie Curie**, Paris, France.

Email: toulouse@lct.jussieu.fr

Web page: www.lct.jussieu.fr/pagesperso/toulouse/

Collaborators: **Claudia Filippi** (Lieden University), **Richard G. Hennig** (Cornell University), **Sandro Sorella** (SISSA, Trieste).

August 2007

- 1 Wave function optimization
- 2 Calculation of pair densities
- 3 Conclusions

- 1 Wave function optimization
- 2 Calculation of pair densities
- 3 Conclusions

Jastrow-Slater wave function

$$|\Psi(\mathbf{p})\rangle = \hat{J}(\boldsymbol{\alpha}) \sum_{i=1}^{N_{\text{CSF}}} c_i |C_i\rangle$$

- $\hat{J}(\boldsymbol{\alpha})$ = **Jastrow factor** (with e-e, e-n, e-e-n terms)
- $|C_i\rangle$ = **Configuration state function (CSF)** = linear combination of Slater determinants of given symmetry.

Jastrow-Slater wave function

$$|\Psi(\mathbf{p})\rangle = \hat{J}(\boldsymbol{\alpha}) \sum_{i=1}^{N_{\text{CSF}}} c_i |C_i\rangle$$

- $\hat{J}(\boldsymbol{\alpha}) =$ **Jastrow factor** (with e-e, e-n, e-e-n terms)
- $|C_i\rangle =$ **Configuration state function (CSF)** = linear combination of Slater determinants of given symmetry.

The Slater determinants are made of orbitals expanded on a Slater basis:

$$\phi_k(\mathbf{r}) = \sum_{\mu=1}^{N_{\text{basis}}} \lambda_{k\mu} \chi_{\mu}(\mathbf{r})$$

$$\chi(\mathbf{r}) = N(\zeta) r^{n-1} e^{-\zeta r} S_{l,m}(\theta, \phi)$$

Jastrow-Slater wave function

$$|\Psi(\mathbf{p})\rangle = \hat{J}(\boldsymbol{\alpha}) \sum_{i=1}^{N_{\text{CSF}}} c_i |C_i\rangle$$

- $\hat{J}(\boldsymbol{\alpha}) =$ **Jastrow factor** (with e-e, e-n, e-e-n terms)
- $|C_i\rangle =$ **Configuration state function (CSF)** = linear combination of Slater determinants of given symmetry.

The Slater determinants are made of orbitals expanded on a Slater basis:

$$\phi_k(\mathbf{r}) = \sum_{\mu=1}^{N_{\text{basis}}} \lambda_{k\mu} \chi_{\mu}(\mathbf{r})$$

$$\chi(\mathbf{r}) = N(\zeta) r^{n-1} e^{-\zeta r} S_{l,m}(\theta, \phi)$$

Parameters to optimize $\mathbf{p} = \{\boldsymbol{\alpha}, \mathbf{c}, \boldsymbol{\lambda}, \zeta\}$: **Jastrow parameters** $\boldsymbol{\alpha}$, **CSF coefficients** \mathbf{c} , **orbital coefficients** $\boldsymbol{\lambda}$ and **basis exponents** ζ

Wave function optimization: why and how?

Important for both VMC and DMC in order to

- reduce the systematic error
- reduce the statistical uncertainty

Wave function optimization: why and how?

Important for both VMC and DMC in order to

- reduce the systematic error
- reduce the statistical uncertainty

How to optimize?

Until recently: minimization of the variance of the energy

- OK for the few Jastrow parameters
- but does not work well for the many CSF and orbital parameters

Since recently: **minimization of the energy** (+ possibly a small fraction of variance)

- in order to optimize well all the parameters
- and because the energy is a better criterion

Wave function parametrization

- Jastrow parameters α , CSF coefficients \mathbf{c} , basis exponents ζ :
no difficulty
- **orbital coefficients λ** are redundant \implies **bad parametrization**

Wave function parametrization

- Jastrow parameters α , CSF coefficients \mathbf{c} , basis exponents ζ : no difficulty
- **orbital coefficients λ** are redundant \implies **bad parametrization**

Reparametrization of orbital coefficients $\lambda \rightarrow \kappa$ (used in MCSCF)

$$|\Psi(\mathbf{p})\rangle = \hat{J}(\alpha) e^{\hat{\kappa}(\kappa)} \sum_{i=1}^{N_{\text{CSF}}} c_i |C_i\rangle$$

where $\hat{\kappa}(\kappa)$ is the generator of **rotations in orbital space** (occupied and virtual):

$$\hat{\kappa}(\kappa) = \sum_{k < l} \kappa_{kl} \left(\hat{E}_{k \rightarrow l} - \hat{E}_{l \rightarrow k} \right)$$

and $\hat{E}_{k \rightarrow l} = \hat{a}_{k\uparrow}^\dagger \hat{a}_{l\uparrow} + \hat{a}_{k\downarrow}^\dagger \hat{a}_{l\downarrow}$ is the singlet excitation operator.

Wave function parametrization

Reparametrization of orbital coefficients $\lambda \rightarrow \kappa$ (used in MCSCF)

$$|\Psi(\mathbf{p})\rangle = \hat{J}(\boldsymbol{\alpha}) e^{\hat{\kappa}(\boldsymbol{\kappa})} \sum_{i=1}^{N_{\text{CSF}}} c_i |C_i\rangle$$

where $\hat{\kappa}(\boldsymbol{\kappa})$ is the generator of **rotations in orbital space** (occupied and virtual):

$$\hat{\kappa}(\boldsymbol{\kappa}) = \sum_{k < l} \kappa_{kl} \left(\hat{E}_{k \rightarrow l} - \hat{E}_{l \rightarrow k} \right)$$

and $\hat{E}_{k \rightarrow l} = \hat{a}_{k\uparrow}^\dagger \hat{a}_{l\uparrow} + \hat{a}_{k\downarrow}^\dagger \hat{a}_{l\downarrow}$ is the singlet excitation operator.

Some points to note

- non-redundant parametrization
- orthonormalization of orbitals preserved if basis exponents are not varied
- can be generalized if basis exponents are varied

Energy minimization in VMC

We have worked on three methods

We have worked on three methods

- **Newton method** (Rappe *et al.*, Umrigar & Filippi, Sorella)

$$E(\mathbf{p}) \approx E(\mathbf{p}^0) + \sum_i \frac{\partial E(\mathbf{p}^0)}{\partial p_i} \Delta p_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 E(\mathbf{p}^0)}{\partial p_i \partial p_j} \Delta p_i \Delta p_j$$

We have worked on three methods

- **Newton method** (Rappe *et al.*, Umrigar & Filippi, Sorella)

$$E(\mathbf{p}) \approx E(\mathbf{p}^0) + \sum_i \frac{\partial E(\mathbf{p}^0)}{\partial p_i} \Delta p_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 E(\mathbf{p}^0)}{\partial p_i \partial p_j} \Delta p_i \Delta p_j$$

- **Linear method** (Nightingale *et al.*)

$$\Psi(\mathbf{p}) \approx \Psi(\mathbf{p}^0) + \sum_i \frac{\partial \Psi(\mathbf{p}^0)}{\partial p_i} \Delta p_i$$

\implies diagonalization of \hat{H} in the basis $\{\Psi(\mathbf{p}^0), \frac{\partial \Psi(\mathbf{p}^0)}{\partial p_i}\}$

$$\mathbf{H} \cdot \Delta \mathbf{p} = E \mathbf{S} \cdot \Delta \mathbf{p}$$

We have worked on three methods

- **Newton method** (Rappe *et al.*, Umrigar & Filippi, Sorella)

$$E(\mathbf{p}) \approx E(\mathbf{p}^0) + \sum_i \frac{\partial E(\mathbf{p}^0)}{\partial p_i} \Delta p_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 E(\mathbf{p}^0)}{\partial p_i \partial p_j} \Delta p_i \Delta p_j$$

- **Linear method** (Nightingale *et al.*)

$$\Psi(\mathbf{p}) \approx \Psi(\mathbf{p}^0) + \sum_i \frac{\partial \Psi(\mathbf{p}^0)}{\partial p_i} \Delta p_i$$

\implies diagonalization of \hat{H} in the basis $\{\Psi(\mathbf{p}^0), \frac{\partial \Psi(\mathbf{p}^0)}{\partial p_i}\}$

$$\mathbf{H} \cdot \Delta \mathbf{p} = E \mathbf{S} \cdot \Delta \mathbf{p}$$

- **Perturbative method** (Scemama & Filippi)

approximate resolution of $\mathbf{H} \cdot \Delta \mathbf{p} = E \mathbf{S} \cdot \Delta \mathbf{p}$ by nonorthogonal perturbation theory

Linear optimization method: principle

- **Expansion of the wave function** around \mathbf{p}^0 to **linear order** in $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}^0$:

$$|\Psi^{[1]}(\mathbf{p})\rangle = |\Psi_0\rangle + \sum_j \Delta p_j |\Psi_j\rangle$$

where $|\Psi_0\rangle = |\Psi(\mathbf{p}^0)\rangle$ and $|\Psi_j\rangle = \frac{\partial |\Psi(\mathbf{p}^0)\rangle}{\partial p_j}$.

Linear optimization method: principle

- **Expansion of the wave function** around \mathbf{p}^0 to **linear order** in $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}^0$:

$$|\Psi^{[1]}(\mathbf{p})\rangle = |\Psi_0\rangle + \sum_j \Delta p_j |\Psi_j\rangle$$

where $|\Psi_0\rangle = |\Psi(\mathbf{p}^0)\rangle$ and $|\Psi_j\rangle = \frac{\partial |\Psi(\mathbf{p}^0)\rangle}{\partial p_j}$.

- **Normalization of wave function chosen so that the derivatives $|\Psi_j\rangle$ are orthogonal to $|\Psi_0\rangle$.**

Linear optimization method: principle

- **Expansion of the wave function** around \mathbf{p}^0 to **linear order** in $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}^0$:

$$|\Psi^{[1]}(\mathbf{p})\rangle = |\Psi_0\rangle + \sum_j \Delta p_j |\Psi_j\rangle$$

where $|\Psi_0\rangle = |\Psi(\mathbf{p}^0)\rangle$ and $|\Psi_j\rangle = \frac{\partial |\Psi(\mathbf{p}^0)\rangle}{\partial p_j}$.

- **Normalization of wave function chosen so that the derivatives $|\Psi_j\rangle$ are orthogonal to $|\Psi_0\rangle$.**
- Minimization of the energy \implies **generalized eigenvalue equation:**

$$\begin{pmatrix} E_0 & \mathbf{g}^T/2 \\ \mathbf{g}/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix}$$

where $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$, $g_i = \frac{\partial E(\mathbf{p}^0)}{\partial p_i}$, $H_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle$, $S_{ij} = \langle \Psi_i | \Psi_j \rangle$.

Linear optimization method: principle

- **Expansion of the wave function** around \mathbf{p}^0 to **linear order** in $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}^0$:

$$|\Psi^{[1]}(\mathbf{p})\rangle = |\Psi_0\rangle + \sum_j \Delta p_j |\Psi_j\rangle$$

where $|\Psi_0\rangle = |\Psi(\mathbf{p}^0)\rangle$ and $|\Psi_j\rangle = \frac{\partial |\Psi(\mathbf{p}^0)\rangle}{\partial p_j}$.

- **Normalization of wave function chosen so that the derivatives $|\Psi_j\rangle$ are orthogonal to $|\Psi_0\rangle$.**
- Minimization of the energy \implies **generalized eigenvalue equation:**

$$\begin{pmatrix} E_0 & \mathbf{g}^T/2 \\ \mathbf{g}/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\mathbf{p} \end{pmatrix}$$

where $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$, $g_i = \frac{\partial E(\mathbf{p}^0)}{\partial p_i}$, $H_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle$, $S_{ij} = \langle \Psi_i | \Psi_j \rangle$.

- **Update of the parameters:** $\mathbf{p}^0 \rightarrow \mathbf{p}^0 + \Delta\mathbf{p}$.

- The linear method is **equivalent to a stabilized Newton method**:

$$\begin{pmatrix} E_0 & \mathbf{g}^T/2 \\ \mathbf{g}/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}$$

$$\iff \begin{cases} (\mathbf{h} + 2\Delta E \mathbf{S}) \cdot \Delta \mathbf{p} = -\mathbf{g} \\ 2\Delta E = -\mathbf{g}^T \cdot \Delta \mathbf{p} \end{cases}$$

where $\mathbf{h} = 2(\mathbf{H} - E_0 \mathbf{S})$ is an approximate Hessian, and $\Delta E = E_0 - E_{\text{lin}} > 0$ is the energy stabilization.

\implies **more robust than Newton method**

- The linear method is **equivalent to a stabilized Newton method**:

$$\begin{pmatrix} E_0 & \mathbf{g}^T/2 \\ \mathbf{g}/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}$$

$$\iff \begin{cases} (\mathbf{h} + 2\Delta E \mathbf{S}) \cdot \Delta \mathbf{p} = -\mathbf{g} \\ 2\Delta E = -\mathbf{g}^T \cdot \Delta \mathbf{p} \end{cases}$$

where $\mathbf{h} = 2(\mathbf{H} - E_0 \mathbf{S})$ is an approximate Hessian, and $\Delta E = E_0 - E_{\text{lin}} > 0$ is the energy stabilization.

\implies **more robust than Newton method**

- In quantum chemistry, it is known as super-CI method or augmented Hessian method.

Linear optimization method: robustness

- The linear method is **equivalent to a stabilized Newton method**:

$$\begin{pmatrix} E_0 & \mathbf{g}^T/2 \\ \mathbf{g}/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}$$

$$\iff \begin{cases} (\mathbf{h} + 2\Delta E \mathbf{S}) \cdot \Delta \mathbf{p} = -\mathbf{g} \\ 2\Delta E = -\mathbf{g}^T \cdot \Delta \mathbf{p} \end{cases}$$

where $\mathbf{h} = 2(\mathbf{H} - E_0 \mathbf{S})$ is an approximate Hessian, and $\Delta E = E_0 - E_{\text{lin}} > 0$ is the energy stabilization.

\implies **more robust than Newton method**

- In quantum chemistry, it is known as super-CI method or augmented Hessian method.
- **Additional stabilization:** $H_{ij} \rightarrow H_{ij} + a \delta_{ij}$ where $a \geq 0$.

Linear optimization method: on a finite VMC sample

The generalized eigenvalue equation is estimated as

$$\begin{pmatrix} E_0 & \mathbf{g}_R^T/2 \\ \mathbf{g}_L/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}$$

with

$$g_{L,i}/2 = \left\langle \frac{\psi_i(\mathbf{R})}{\psi_0(\mathbf{R})} \frac{H(\mathbf{R})\psi_0(\mathbf{R})}{\psi_0(\mathbf{R})} \right\rangle_{\psi_0^2} \quad \text{and} \quad g_{R,j}/2 = \left\langle \frac{\psi_0(\mathbf{R})}{\psi_0(\mathbf{R})} \frac{H(\mathbf{R})\psi_j(\mathbf{R})}{\psi_0(\mathbf{R})} \right\rangle_{\psi_0^2}$$
$$H_{ij} = \left\langle \frac{\psi_i(\mathbf{R})}{\psi_0(\mathbf{R})} \frac{H(\mathbf{R})\psi_j(\mathbf{R})}{\psi_0(\mathbf{R})} \right\rangle_{\psi_0^2} \quad \text{and} \quad S_{ij} = \left\langle \frac{\psi_i(\mathbf{R})}{\psi_0(\mathbf{R})} \frac{\psi_j(\mathbf{R})}{\psi_0(\mathbf{R})} \right\rangle_{\psi_0^2}$$

non-symmetric!

Linear optimization method: on a finite VMC sample

The generalized eigenvalue equation is estimated as

$$\begin{pmatrix} E_0 & \mathbf{g}_R^T/2 \\ \mathbf{g}_L/2 & \mathbf{H} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix} = E_{\text{lin}} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}$$

with

$$g_{L,i}/2 = \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{H(\mathbf{R})\Psi_0(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle_{\Psi_0^2} \quad \text{and} \quad g_{R,j}/2 = \left\langle \frac{\Psi_0(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{H(\mathbf{R})\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle_{\Psi_0^2}$$
$$H_{ij} = \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{H(\mathbf{R})\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle_{\Psi_0^2} \quad \text{and} \quad S_{ij} = \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi_0(\mathbf{R})} \frac{\Psi_j(\mathbf{R})}{\Psi_0(\mathbf{R})} \right\rangle_{\Psi_0^2}$$

non-symmetric!

⇒ **Zero-variance principle** of Nightingale *et al.* (PRL 2001):

If there are some $\Delta \mathbf{p}$ so that $\Psi_0(\mathbf{R}) + \sum_j \Delta p_j \Psi_j(\mathbf{R}) = \Psi_{\text{exact}}(\mathbf{R})$
then $\Delta \mathbf{p}$ are found with **zero variance**.

In practice, these non-symmetric estimators reduce the fluctuations on $\Delta \mathbf{p}$ by 1 or 2 orders of magnitude.

How to minimize the energy variance with the linear method?

$$V = \min_{\Delta \mathbf{p}} \left\{ V_0 + \mathbf{g}_V^T \cdot \Delta \mathbf{p} + \frac{1}{2} \Delta \mathbf{p}^T \cdot \mathbf{h}_V \cdot \Delta \mathbf{p} \right\}$$

How to minimize the energy variance with the linear method?

$$V = \min_{\Delta \mathbf{p}} \left\{ V_0 + \mathbf{g}_V^T \cdot \Delta \mathbf{p} + \frac{1}{2} \Delta \mathbf{p}^T \cdot \mathbf{h}_V \cdot \Delta \mathbf{p} \right\}$$

$$\Leftrightarrow V = \min_{\Delta \mathbf{p}} \frac{\begin{pmatrix} 1 & \Delta \mathbf{p}^T \end{pmatrix} \begin{pmatrix} V_0 & \mathbf{g}_V^T/2 \\ \mathbf{g}_V/2 & \mathbf{h}_V/2 + V_0 \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}}{\begin{pmatrix} 1 & \Delta \mathbf{p}^T \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}}$$

How to minimize the energy variance with the linear method?

$$V = \min_{\Delta \mathbf{p}} \left\{ V_0 + \mathbf{g}_V^T \cdot \Delta \mathbf{p} + \frac{1}{2} \Delta \mathbf{p}^T \cdot \mathbf{h}_V \cdot \Delta \mathbf{p} \right\}$$

$$\Leftrightarrow V = \min_{\Delta \mathbf{p}} \frac{\begin{pmatrix} 1 & \Delta \mathbf{p}^T \end{pmatrix} \begin{pmatrix} V_0 & \mathbf{g}_V^T/2 \\ \mathbf{g}_V/2 & \mathbf{h}_V/2 + V_0 \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}}{\begin{pmatrix} 1 & \Delta \mathbf{p}^T \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}}$$

$$\Leftrightarrow \begin{pmatrix} V_0 & \mathbf{g}_V^T/2 \\ \mathbf{g}_V/2 & \mathbf{h}_V/2 + V_0 \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix} = V \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}$$

How to minimize the energy variance with the linear method?

$$V = \min_{\Delta \mathbf{p}} \left\{ V_0 + \mathbf{g}_V^T \cdot \Delta \mathbf{p} + \frac{1}{2} \Delta \mathbf{p}^T \cdot \mathbf{h}_V \cdot \Delta \mathbf{p} \right\}$$

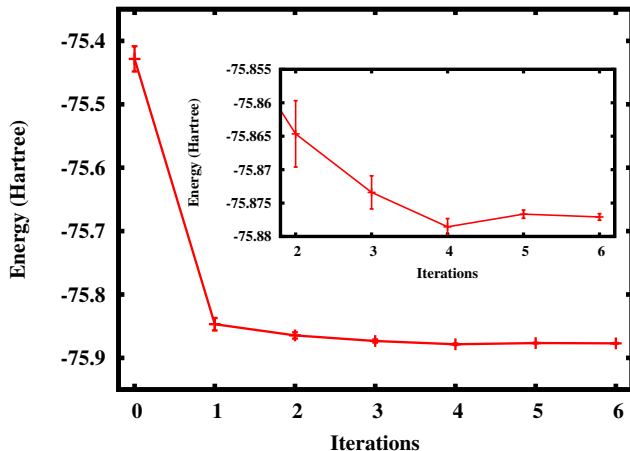
$$\Leftrightarrow V = \min_{\Delta \mathbf{p}} \frac{\begin{pmatrix} 1 & \Delta \mathbf{p}^T \end{pmatrix} \begin{pmatrix} V_0 & \mathbf{g}_V^T/2 \\ \mathbf{g}_V/2 & \mathbf{h}_V/2 + V_0 \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}}{\begin{pmatrix} 1 & \Delta \mathbf{p}^T \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}}$$

$$\Leftrightarrow \begin{pmatrix} V_0 & \mathbf{g}_V^T/2 \\ \mathbf{g}_V/2 & \mathbf{h}_V/2 + V_0 \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix} = V \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \mathbf{p} \end{pmatrix}$$

matrix to add to the energy matrix

Simultaneous optimization of all parameters in VMC

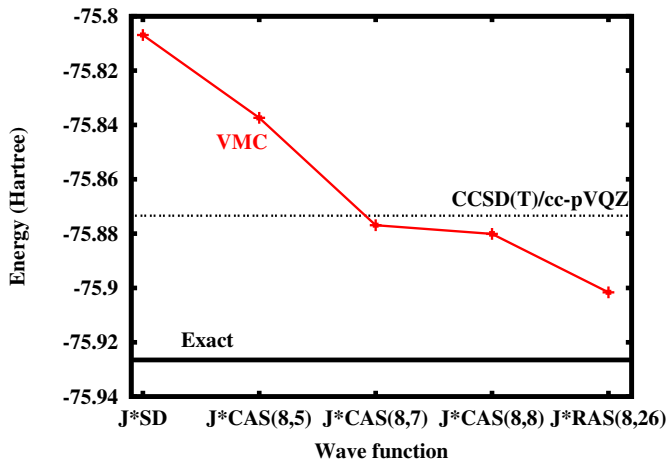
Optimization of 24 Jastrow, 49 CSF, 64 orbital and 12 exponent parameters for the C_2 molecule:



⇒ The energy converges with an accuracy of 1 mhartree in about 4 or 5 iterations

Systematic improvement in QMC

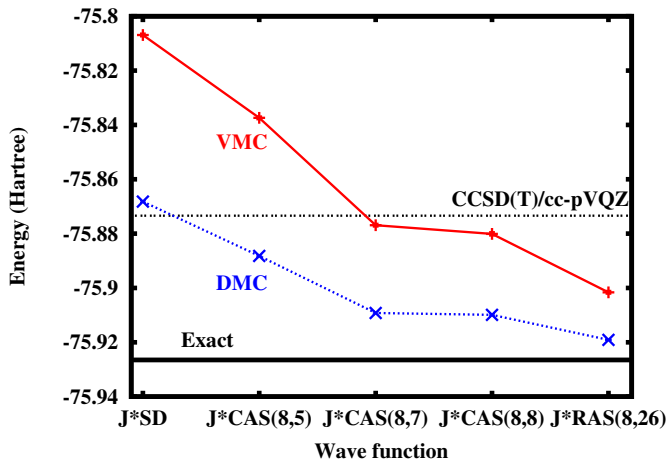
For C_2 molecule: total energies for a **series of fully optimized Jastrow-Slater wave functions**:



⇒ **Systematic improvement in VMC**

Systematic improvement in QMC

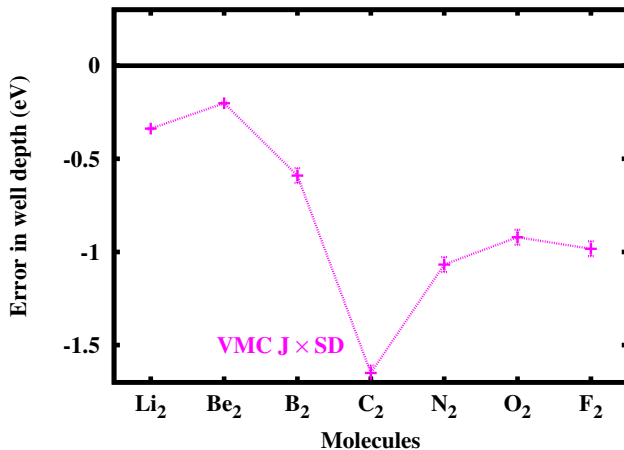
For C_2 molecule: total energies for a **series of fully optimized Jastrow-Slater wave functions**:



⇒ **Systematic improvement in VMC and DMC!**

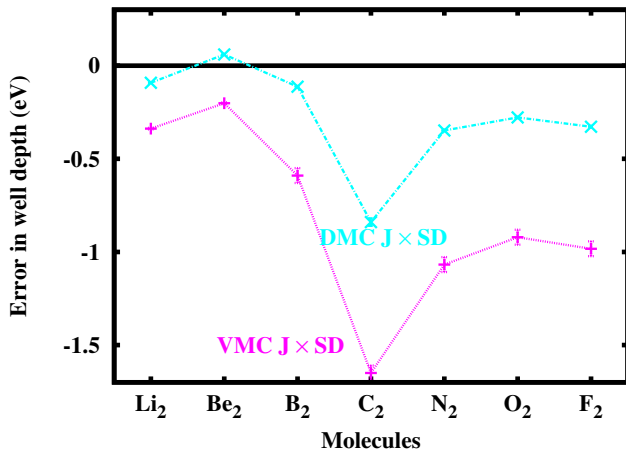
Well depths of second-row homonuclear diatomic molecules

VMC and **DMC** errors in well depths for some **fully optimized Jastrow-Slater wave functions**:



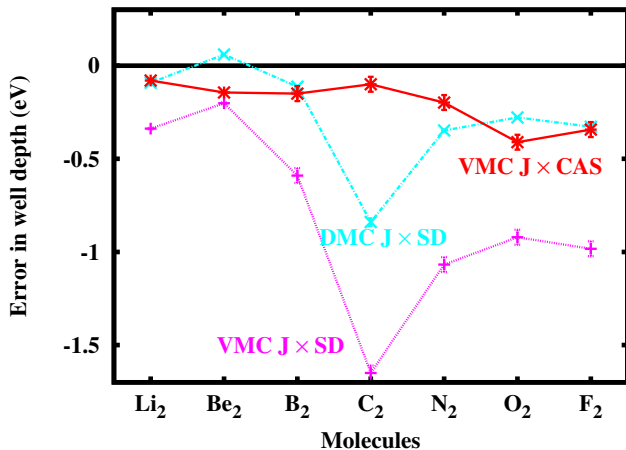
Well depths of second-row homonuclear diatomic molecules

VMC and **DMC** errors in well depths for some **fully optimized Jastrow-Slater wave functions**:



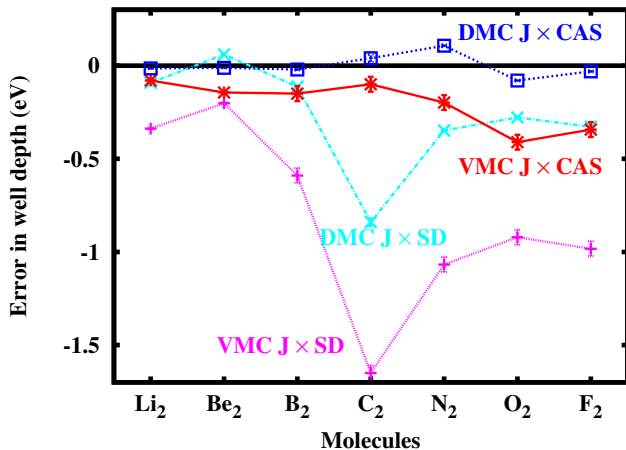
Well depths of second-row homonuclear diatomic molecules

VMC and **DMC** errors in well depths for some **fully optimized Jastrow-Slater wave functions**:



Well depths of second-row homonuclear diatomic molecules

VMC and **DMC** errors in well depths for some **fully optimized Jastrow-Slater wave functions**:



⇒ Near chemical accuracy in **DMC** with Jastrow × **CAS**

- 1 Wave function optimization
- 2 Calculation of pair densities
- 3 Conclusions

Calculation of pair densities

Spherically and system-averaged pair density = intracule density

$$I(u) = \sum_{i < j} \int \frac{d\Omega_{\mathbf{u}}}{4\pi} \int d\mathbf{R} \Psi(\mathbf{R})^2 \delta(\mathbf{r}_{ij} - \mathbf{u})$$

e.g., gives the Coulombic electron-electron interaction energy:

$$W_{ee} = \int_0^{\infty} du 4\pi u^2 I(u) \frac{1}{u}$$

Usefulness

- **qualitative analysis** of electronic structure (Cioslowski, Gill, Ugalde, etc...)
- **quantitative predictions** beyond usual DFT (Gori-Giorgi, Perdew, Savin, etc...)

Usual “histogram” method

$$I(u) \approx \sum_{i < j} \int \frac{d\Omega_{\mathbf{u}}}{4\pi} \int d\mathbf{R} \Psi(\mathbf{R})^2 \frac{1_{[u-\Delta u/2, u+\Delta u/2]}(\mathbf{r}_{ij})}{\Delta u^3}$$

Problems

- large statistical uncertainty due to large variance, especially at small u
- systematic error due to approximate $\Psi(\mathbf{R})$ but also due to discretization over u

Local energy

- Estimator: $E_L(\mathbf{R}) = \frac{H(\mathbf{R})\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$
 - Systematic error: $\delta E = \mathcal{O}(\delta\Psi^2)$
 - Variance: $\sigma^2(E_L) = \mathcal{O}(\delta\Psi^2)$
- Quadratic Zero-Variance
Zero-Bias property**

Calculation of an observable in VMC

Local energy

- Estimator: $E_L(\mathbf{R}) = \frac{H(\mathbf{R})\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$
 - Systematic error: $\delta E = \mathcal{O}(\delta\Psi^2)$
 - Variance: $\sigma^2(E_L) = \mathcal{O}(\delta\Psi^2)$
- Quadratic Zero-Variance
Zero-Bias property**

Arbitrary observable \hat{O} (which does not commute with \hat{H})

- Estimator: $O_L(\mathbf{R}) = \frac{O(\mathbf{R})\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$
 - Systematic error: $\delta O = \mathcal{O}(\delta\Psi)$
 - Variance: $\sigma^2(O_L) = \mathcal{O}(1)$
- ~~**Quadratic Zero-Variance
Zero-Bias property**~~

Zero-Variance Zero-Bias estimators (Assaraf & Caffarel)

Consider the λ -dependent Hamiltonian

$$\hat{H}^\lambda = \hat{H} + \lambda \hat{O}$$

with an associated trial wave function

$$\Psi^\lambda = \Psi + \lambda \Psi' + \dots$$

Zero-Variance Zero-Bias estimators (Assaraf & Caffarel)

Consider the λ -dependent Hamiltonian

$$\hat{H}^\lambda = \hat{H} + \lambda \hat{O}$$

with an associated trial wave function

$$\Psi^\lambda = \Psi + \lambda \Psi' + \dots$$

Hellmann-Feynman theorem suggests to define **ZVZB estimator**

$$\langle O_L^{\text{ZVZB}}(\mathbf{R}) \rangle_{\Psi^2} = \left(\frac{dE^\lambda}{d\lambda} \right)_{\lambda=0} = \langle O_L(\mathbf{R}) \rangle_{\Psi^2} + \langle \Delta O_L^{\text{ZV}}(\mathbf{R}) \rangle_{\Psi^2} + \langle \Delta O_L^{\text{ZB}}(\mathbf{R}) \rangle_{\Psi^2},$$

with the **ZV term**

$$\Delta O_L^{\text{ZV}}(\mathbf{R}) = \left[\frac{H(\mathbf{R})\Psi'(\mathbf{R})}{\Psi'(\mathbf{R})} - E_L(\mathbf{R}) \right] \frac{\Psi'(\mathbf{R})}{\Psi(\mathbf{R})}$$

and the **ZB term**

$$\Delta O_L^{\text{ZB}}(\mathbf{R}) = 2 [E_L(\mathbf{R}) - E] \frac{\Psi'(\mathbf{R})}{\Psi(\mathbf{R})}$$

Zero-Variance Zero-Bias estimators (Assaraf & Caffarel)

Hellmann-Feynman theorem suggests to define **ZVZB estimator**

$$\langle O_L^{\text{ZVZB}}(\mathbf{R}) \rangle_{\Psi^2} = \left(\frac{dE^\lambda}{d\lambda} \right)_{\lambda=0} = \langle O_L(\mathbf{R}) \rangle_{\Psi^2} + \langle \Delta O_L^{\text{ZV}}(\mathbf{R}) \rangle_{\Psi^2} + \langle \Delta O_L^{\text{ZB}}(\mathbf{R}) \rangle_{\Psi^2},$$

with the **ZV term**

$$\Delta O_L^{\text{ZV}}(\mathbf{R}) = \left[\frac{H(\mathbf{R})\Psi'(\mathbf{R})}{\Psi'(\mathbf{R})} - E_L(\mathbf{R}) \right] \frac{\Psi'(\mathbf{R})}{\Psi(\mathbf{R})}$$

and the **ZB term**

$$\Delta O_L^{\text{ZB}}(\mathbf{R}) = 2 [E_L(\mathbf{R}) - E] \frac{\Psi'(\mathbf{R})}{\Psi(\mathbf{R})}$$

Quadratic Zero-Variance Zero-Bias property

- Systematic error: $\delta O^{\text{ZVZB}} = \mathcal{O}(\delta\Psi^2 + \delta\Psi \delta\Psi')$
- Variance: $\sigma^2 \left(O_L^{\text{ZVZB}} \right) = \mathcal{O}(\delta\Psi^2 + \delta\Psi'^2 + \delta\Psi \delta\Psi')$

Calculation of intracules in QMC

Simplest approximate wave function derivative:

$$\Psi'(\mathbf{R}) = -\frac{1}{4\pi} \sum_{i < j} \int \frac{d\Omega_{\mathbf{u}}}{4\pi} \frac{1}{|\mathbf{r}_{ij} - \mathbf{u}|} \Psi(\mathbf{R})$$

ZVZB improved estimator

$$I(u) = -\frac{1}{2\pi} \sum_{i < j} \int \frac{d\Omega_{\mathbf{u}}}{4\pi} \int d\mathbf{R} \Psi(\mathbf{R})^2 \left[\frac{\nabla_{\mathbf{r}_j} \Psi(\mathbf{R})}{\Psi(\mathbf{R})} \cdot \frac{\mathbf{r}_{ij} - \mathbf{u}}{|\mathbf{r}_{ij} - \mathbf{u}|^3} + (E_L(\mathbf{R}) - E) \frac{1}{|\mathbf{r}_{ij} - \mathbf{u}|} \right]$$

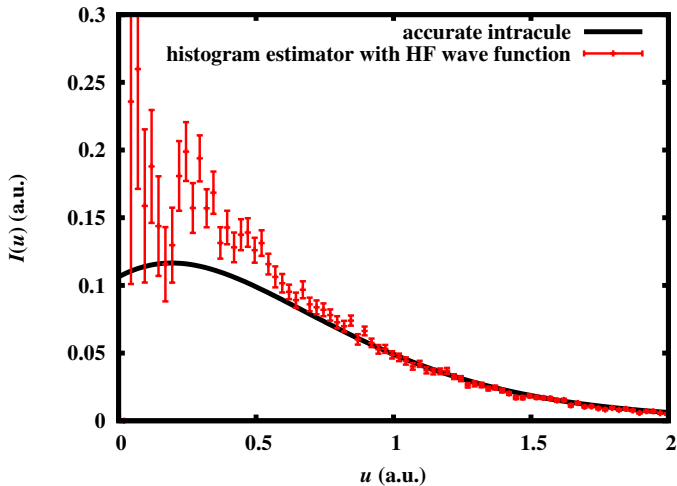
(+ possible refinements)

Advantages

- reduction of variance
- reduction of systematic error

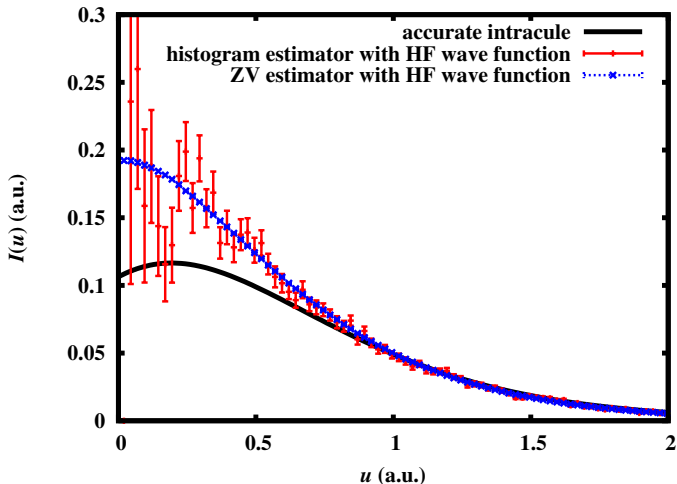
Comparison of the estimators

Intracule $I(u)$ of the He atom in VMC (100000 configurations):



Comparison of the estimators

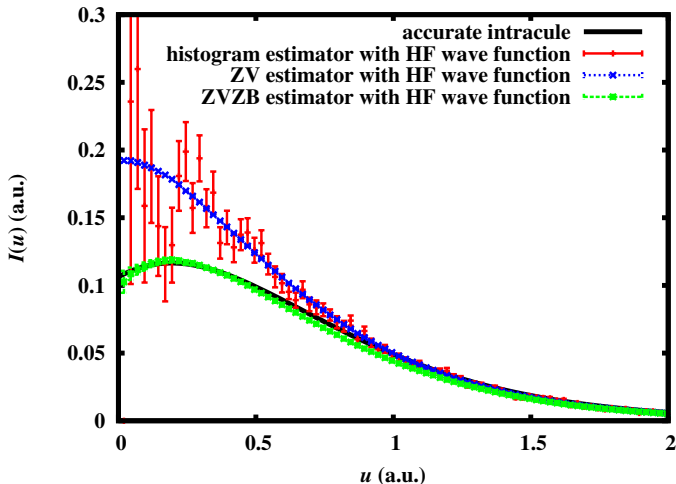
Intracule $I(u)$ of the He atom in VMC (100000 configurations):



⇒ reduction of statistical uncertainty and systematic error by several orders of magnitude

Comparison of the estimators

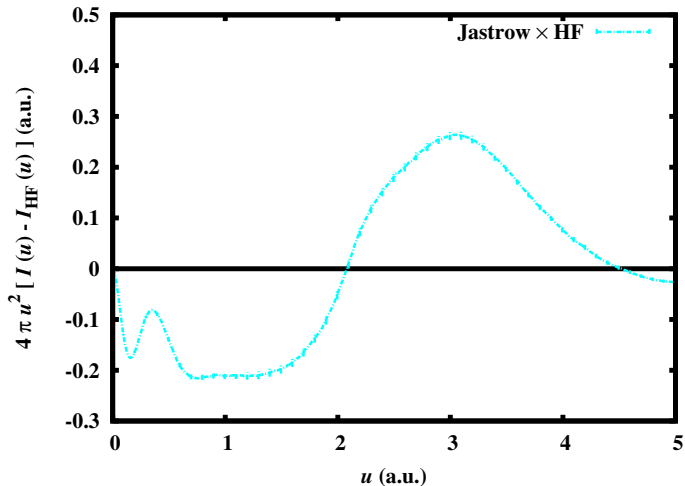
Intracule $I(u)$ of the He atom in VMC (100000 configurations):



⇒ reduction of statistical uncertainty and systematic error
by several orders of magnitude

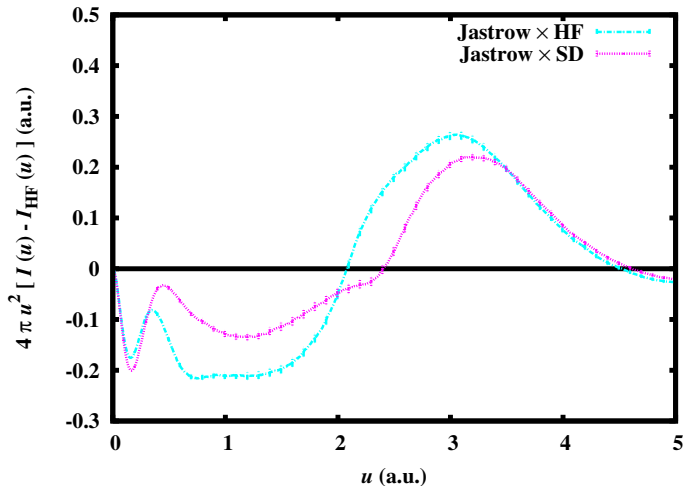
Systematic improvement of the intracule

Correlation hole $4\pi u^2 [I(u) - I_{HF}(u)]$ of the C_2 molecule in VMC for a series of wave functions:



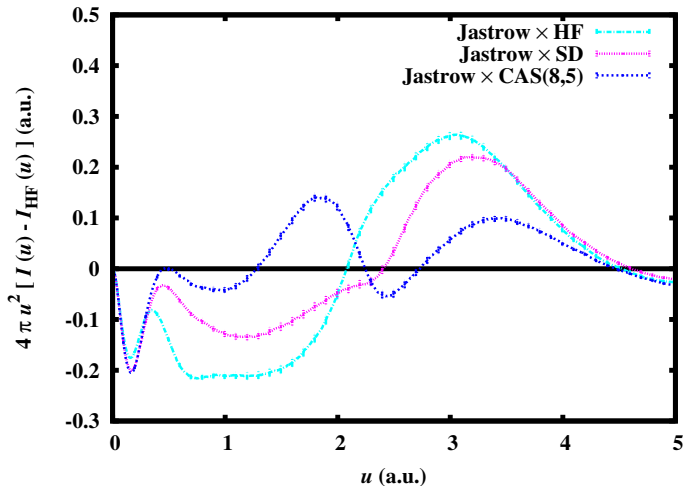
Systematic improvement of the intracule

Correlation hole $4\pi u^2 [I(u) - I_{HF}(u)]$ of the C_2 molecule in VMC for a series of wave functions:



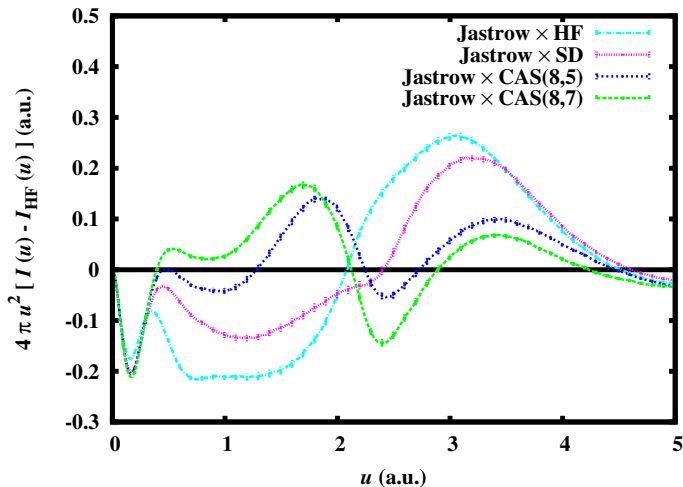
Systematic improvement of the intracule

Correlation hole $4\pi u^2 [I(u) - I_{HF}(u)]$ of the C_2 molecule in VMC for a series of wave functions:



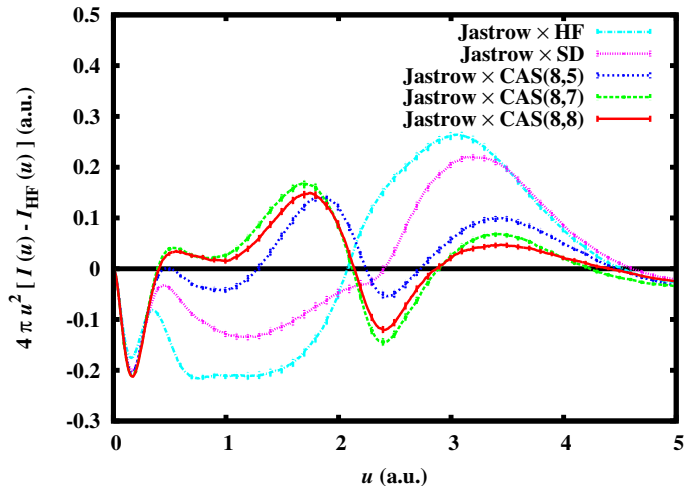
Systematic improvement of the intracule

Correlation hole $4\pi u^2 [I(u) - I_{HF}(u)]$ of the C_2 molecule in VMC for a series of wave functions:



Systematic improvement of the intracule

Correlation hole $4\pi u^2 [I(u) - I_{HF}(u)]$ of the C_2 molecule in VMC for a series of wave functions:



- 1 Wave function optimization
- 2 Calculation of pair densities
- 3 Conclusions**

Summary

- **Efficient wave function optimization method by energy minimization in VMC.**
- Achievement of **systematic improvement** and **near chemical accuracy.**
- **Improved QMC estimators** for calculating **pair densities.**

Toulouse, Umrigar, JCP 126, 084102 (2007)

Umrigar, Toulouse, Filippi, Sorella, Hennig, PRL 98, 110201 (2007)

Toulouse, Assaraf, Umrigar, JCP 126, 244112 (2007)

Web page: www.lct.jussieu.fr/pagesperso/toulouse/

Future work

- Direct minimization of the DMC energy.
- Geometry optimization.