

Optimization of many-body wave functions

Cyrus J. Umrigar

QMC summer school, UIUC, 9-19 July 2007

Optimization of many-body wave functions

1. Variance minimization vs. Variational Energy minimization
2. Energy optimization methods that work well:
 - 1 Newton method (CJU, Filippi (2005))
 - 2 Linear method (CJU, Toulouse, Filippi, Sorella, Hennig (2007) (extension of Nightingale's method to nonlinear parms.))
 - 3 Perturbative method (Scemama, Filippi(2006); Toulouse, CJU (2007)) (simplification of EFP method (Fahy, Filippi, Schautz, Prendergast))
3. Stabilization of methods
4. Results for:
 - 1 NO₂, decapentaene (C₁₀H₁₂) (convergence of Newton, advantage of mixed energy and variance optim.)
 - 2 Si₂, C₂ (widely varying FN error)
 - 3 First-row diatomics (agreement with experiment)
 - 4 Ethene (excited states (Schautz, Filippi))
 - 5 Wigner crystallization in quantum dots (floating gaussians (Guclu))
5. Optimization of DMC energy

Accuracy of Diffusion Monte Carlo

The problem:

Fixed-node error can be **LARGE** for these systems. e.g., the fixed-node error for C_2 total energy is 1.3 eV and for well-depth is 0.7 eV.

Possible solutions:

1. Solve the Fermion sign problem. **Hard!**
2. Develop:
 - 1 Better forms of trial wave functions.
 - 2.1.1 Casula and Sorella on geminals
 - 2.1.2 Schmidt, and, Bajdich and Mitas on Pfaffians
 - 2.1.3 Rios and Needs on backflow for nonfluid systems
 - 2 Powerful optimization methods to systematically improve the nodes of the trial wavefunctions.
 - 2.2.1 CJU, Filippi, PRL, **94**, 150201 (2005)
 - 2.2.2 CJU, Toulouse, Filippi, Sorella, Hennig, PRL **98**, 110201 (2007).
 - 2.2.3 Toulouse, CJU, J. Chem. Phys., **126**, 084102 (2007).

Almost all other errors reduced too

1. Reduce fixed-node error (nodes move during optimization). Fixed node errors can be **LARGE**.
2. Reduce other systematic errors in the energy – pseudopotential locality error, time-step error, population-control error.
3. Reduce systematic error of observables that do not commute with the Hamiltonian (mixed estimators, $\langle \Psi_0 | \hat{A} | \Psi_T \rangle$ not exact even for nodeless Ψ).
4. Reduce statistical error.

Functional form of Trial Wave Function

$$\Psi_T = \left(\sum_n d_n D_n^\uparrow D_n^\downarrow \right) \times \mathcal{J}(r_i, r_j, r_{ij})$$

• **Determinants:** $\sum_n d_n D_n^\uparrow D_n^\downarrow$

D^\uparrow and D^\downarrow are determinants of single-particle orbitals ϕ for up (\uparrow) and down (\downarrow) spin electrons respectively.

The single-particle orbitals ϕ are given by:

$$\phi(\mathbf{r}_i) = \sum_{\alpha k} c_{k\alpha} N_{k\alpha} r_{i\alpha}^{n_{k\alpha}-1} e^{-\zeta_{k\alpha} r_{i\alpha}} Y_{l_{k\alpha} m_{k\alpha}}(\hat{\mathbf{r}}_{i\alpha})$$

• **Jastrow:** $\mathcal{J}(r_i, r_j, r_{ij}) = \prod_{\alpha i} \exp(A_{\alpha i}) \prod_{ij} \exp(B_{ij}) \prod_{\alpha ij} \exp(C_{\alpha ij})$

$A_{\alpha i} \Rightarrow$ electron-ion correlation $\sim N_{\text{atomtype}}$ of \mathcal{J} parms.

$B_{ij} \Rightarrow$ electron-electron correlation $\sim N_{\text{atomtype}}$ of $\zeta_{k\alpha}$ parms.

$C_{\alpha ij} \Rightarrow$ electron-electron-ion correlation $\sim N_{\text{atom}}^2$ of $c_{k\alpha}$ parms.

$\sim e^{N_{\text{atom}}}$ of d_n parms.

$d_n, c_{k\alpha}, \zeta_{k\alpha}$ and parms in \mathcal{J} are optimized.

Power of QMC:

\mathcal{J} parms. do work of d_n parms.

Progress in optimization of Many-Body Wavefunctions

Naive energy optim. → Variance optim. → Efficient energy optim.

- 1988 naive energy optimization, few (~ 3) parameters
- 1988 – 2001 variance optimization, ~ 100 parameters
- 2001 – efficient energy optimization, > 1000 parameters

Optimization of Many-Body Wavefunctions

A major advantage of quantum Monte Carlo methods is that there is no restriction on the form of $\Psi_T(\mathbf{R})$. Hence any insight one may have, as regards the nature of the many-body correlations, can be built into $\Psi_T(\mathbf{R})$ and tested. To exploit this freedom it is necessary to have a method for optimizing arbitrary wavefunctions.

First thought: Minimize the energy on MC sample.

$$\bar{E} = \sum_{i=1}^{N_{\text{conf}}} \frac{\mathcal{H}\Psi_T(\mathbf{R}_i; \{p\})}{\Psi_T(\mathbf{R}_i; \{p\})} w_i, \quad w_i = \frac{|\Psi_T(\mathbf{R}_i)|^2}{|\Psi_T^0(\mathbf{R}_i)|^2} \bigg/ \sum_{i=1}^{N_{\text{conf}}} \frac{|\Psi_T(\mathbf{R}_i)|^2}{|\Psi_T^0(\mathbf{R}_i)|^2}$$

Second thought: Minimize the variance of the local energy.

$$\sigma^2 = \sum_{i=1}^{N_{\text{conf}}} \left(\frac{\mathcal{H}\Psi_T(\mathbf{R}_i; \{p\})}{\Psi_T(\mathbf{R}_i; \{p\})} - \bar{E} \right)^2 w_i$$

Third thought: Minimize the energy using MC but not on MC sample.

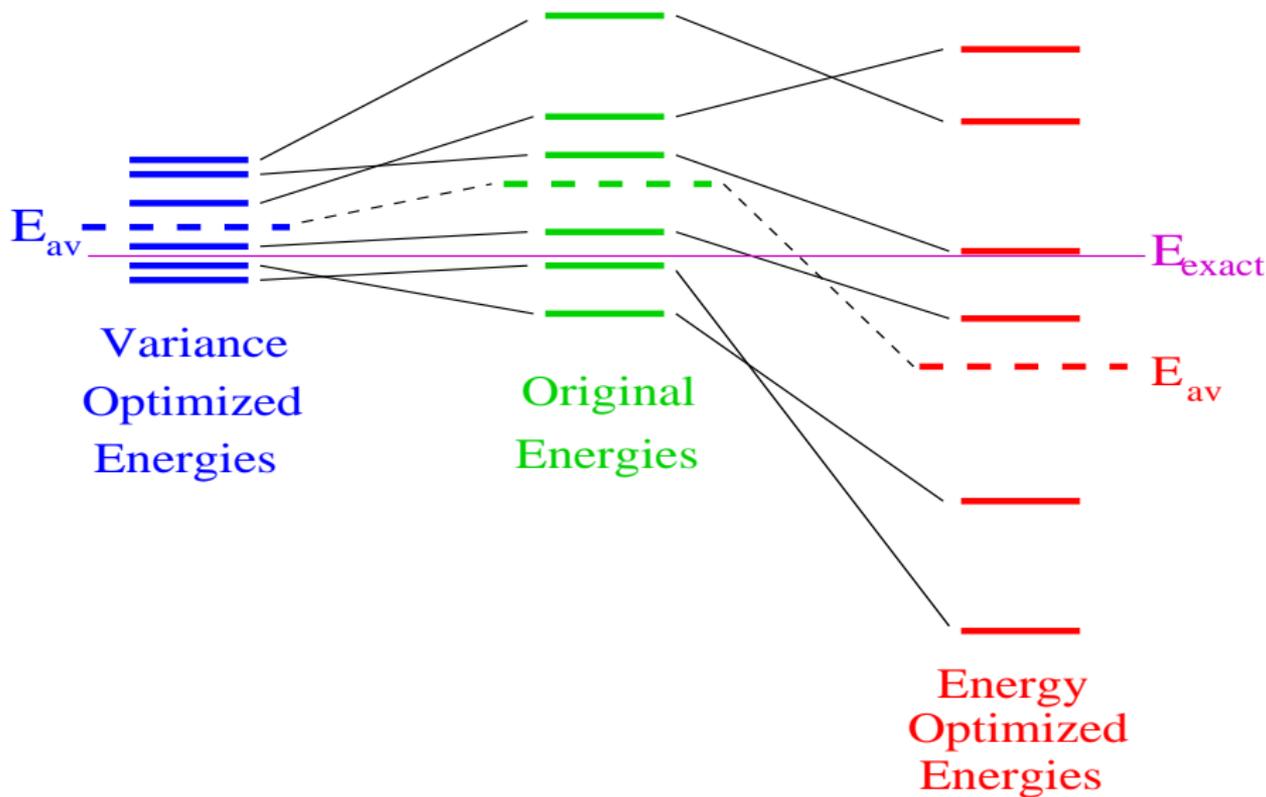
Variance

$$\sigma^2 = \sum_{i=1}^{N_{\text{conf}}} \left(\frac{\mathcal{H}\Psi_{\text{T}}(\mathbf{R}_i)}{\Psi_{\text{T}}(\mathbf{R}_i)} - \bar{E} \right)^2$$

vs.

Energy

$$\bar{E} = \sum_{i=1}^{N_{\text{conf}}} \frac{\mathcal{H}\Psi_{\text{T}}(\mathbf{R}_i)}{\Psi_{\text{T}}(\mathbf{R}_i)}$$



Advantages of Energy (or Mixed) Optim. vs. Variance Optim.

1. Want lowest energy; fluctuations are of secondary importance. Energy and variance do not always go hand-in-hand enough.
2. Some parameters couple more strongly to energy than variance.
3. Energy-optimized Ψ can be better for other expectation values (Rothstein).
4. Hellman-Feynman theorem can be used for forces (when combined with variance reduction methods).

Early energy optimization methods that worked well

Linear method for linear parameters:

M. P. Nightingale and Melik-Alaverdian, Phys. Rev. Lett., **87**, 043401 (2001).

Effective fluctuation potential method:

- ▶ S. Fahy, in *Quantum Monte Carlo Methods in Physics and Chemistry*, edited by M. P. Nightingale and C. J. Umrigar, [NATO ASI Ser. C. **525** 101, 1999];
- ▶ Filippi and S. Fahy, J. Chem. Phys., **112**, 3523 (2000);
- ▶ F. Schautz and S. Fahy, J. Chem. Phys., **116**, 3533 (2002);
- ▶ D. Prendergast, D. Bevan and S. Fahy, Phys. Rev. B, **66**, 155104 (2002);
- ▶ Friedemann Schautz and Claudia Filippi, J. Chem. Phys., **120**, 10931 (2004).

Stochastic reconfiguration method:

- ▶ Sandro Sorella, Phys. Rev. B, **64**, 024512 (2001);
- ▶ Casula and Sorella, J. Chem. Phys., **119**, 6500 (2003).
- ▶ Sorella, PRB 71, 241103 (2005).

Current best energy-optimization methods

1. Newton method [CJU, Filippi, PRL 94, 150201 \(2005\)](#):

Add terms to the Hessian that contribute nothing in the limit of an infinite MC sample, but cancel much of the fluctuations for a finite MC sample.

Gain in efficiency: 3 orders of magnitude for NO_2 , more for $\text{C}_{10}\text{H}_{12}$ compared to Newton of Lin-Zhang-Rappe.

2. Linear method (generalized eigenvalue problem):

1 Linear parameters: [Nightingale, et al., PRL, 87, 043401 \(2001\)](#)

Use asymmetric H to have zero variance property in the limit that the basis functions span an invariant subspace.

2 Nonlinear parameters:

[CJU, Toulouse, Filippi, Sorella, Hennig, PRL 98, 110201 \(2007\)](#).

Choose freedom of normalization $\Psi(\mathbf{p}, \mathbf{R}) = N(\mathbf{p}) \Phi(\mathbf{p}, \mathbf{R})$ to make a near optimal change in the parameters.

3. Perturbation theory in an arbitrary nonorthog. basis:

[Toulouse, CJU, J. Chem. Phys., 126, 084102 \(2007\)](#).

(Small modification of [Scemama-Filippi \(2006\)](#) perturbative EFP, modification of the [Fahy-Filippi-Prendergast-Schautz](#) EFP method.)

Newton energy minimization method

Find parameter changes by solving linear equations:

$$\mathbf{h}\delta\mathbf{p} = -\mathbf{g},$$

\mathbf{h} is the Hessian and \mathbf{g} is the gradient of the energy.

Two modifications of the straightforward Newton method:

1. Add terms to the Hessian that have zero expectation value on an infinite MC sample but that cancel much of the fluctuations on a finite sample.
2. Replace certain terms in the Hessian by other terms that are approximately proportional to them but that fluctuate less.

These 2 changes improve the efficiency of the simple Newton method of Lin, Zhang and Rappe by about 3 orders of magnitude.

Newton energy minimization method

$$\bar{E} = \frac{\int d^3\mathbf{R} \psi H\psi}{\int d^3\mathbf{R} |\psi|^2} = \langle E_L \rangle; \quad E_L(\mathbf{R}) = \frac{H\psi(\mathbf{R})}{\psi(\mathbf{R})}$$

$$\bar{E}_i = \left\langle \frac{\psi_i}{\psi} E_L + \frac{H\psi_i}{\psi} - 2\bar{E} \frac{\psi_i}{\psi} \right\rangle = 2 \left\langle \frac{\psi_i}{\psi} (E_L - \bar{E}) \right\rangle \quad (\text{by Hermiticity}).$$

$$\bar{E}_{ij} = 2 \left[\left\langle \left(\frac{\psi_{ij}}{\psi} + \frac{\psi_i \psi_j}{|\psi|^2} \right) (E_L - \bar{E}) \right\rangle - \left\langle \frac{\psi_i}{\psi} \right\rangle \bar{E}_j - \left\langle \frac{\psi_j}{\psi} \right\rangle \bar{E}_i + \left\langle \frac{\psi_i}{\psi} E_{L,j} \right\rangle \right].$$

Identical to the Hessian in Lin, Zhang, Rappe.

Two changes:

- 1) Symmetrize – eigensystem is real
- 2) Noting that $\langle E_{L,j} \rangle = 0$, rewrite in the form of a covariance

$$\langle \langle ab \rangle - \langle a \rangle \langle b \rangle \rangle$$

The fluctuations of $\langle ab \rangle - \langle a \rangle \langle b \rangle$ are in most cases smaller than those of $\langle ab \rangle$, (e.g. if a and b are weakly correlated), and, they are much smaller if $\sqrt{\langle a^2 \rangle - \langle a \rangle^2} \ll |\langle a \rangle|$ and a is not strongly correlated with $1/b$.

Newton energy minimization method (cont.)

$$\begin{aligned}
 \bar{E}_{ij} &= 2 \left[\left\langle \left(\frac{\psi_{ij}}{\psi} + \frac{\psi_i \psi_j}{|\psi|^2} \right) (E_L - \bar{E}) \right\rangle - \left\langle \frac{\psi_i}{\psi} \right\rangle \bar{E}_j - \left\langle \frac{\psi_j}{\psi} \right\rangle \bar{E}_i \right] \\
 &\quad + \left\langle \frac{\psi_i}{\psi} E_{L,j} \right\rangle - \left\langle \frac{\psi_i}{\psi} \right\rangle \langle E_{L,j} \rangle + \left\langle \frac{\psi_j}{\psi} E_{L,i} \right\rangle - \left\langle \frac{\psi_j}{\psi} \right\rangle \langle E_{L,i} \rangle \\
 &= 2 \left[\left\langle \left(\frac{\psi_{ij}}{\psi} - \frac{\psi_i \psi_j}{|\psi|^2} \right) (E_L - \bar{E}) \right\rangle \quad (0 \text{ for lin. } p_i \text{ in expon}) \right. \\
 &\quad \left. + 2 \left\langle \left(\frac{\psi_i}{\psi} - \left\langle \frac{\psi_i}{\psi} \right\rangle \right) \left(\frac{\psi_j}{\psi} - \left\langle \frac{\psi_j}{\psi} \right\rangle \right) (E_L - \bar{E}) \right\rangle \right] \\
 &\quad + \left\langle \frac{\psi_i}{\psi} E_{L,j} \right\rangle - \left\langle \frac{\psi_i}{\psi} \right\rangle \langle E_{L,j} \rangle + \left\langle \frac{\psi_j}{\psi} E_{L,i} \right\rangle - \left\langle \frac{\psi_j}{\psi} \right\rangle \langle E_{L,i} \rangle.
 \end{aligned}$$

1) Additional terms =0 for infinite sample but cancel most of the fluctuations for a finite sample.

2) Red terms fluctuate less than blue terms away from minimum. Ratio of red terms to blue terms depends on ψ but is roughly independent of i and j . Why?? Exploit that!

Optimization of determinantal parameters

Different issues arise in optimizing Jastrow parameters and determinantal parameters:

Jastrow: eigenvalues of Hessian have a range of 11 orders of magnitude!

Determinantal parameters: divergences in elements of Hessian and Hamiltonian matrices.

$$\bar{E}_i = 2 \left\langle \frac{\psi_i}{\psi} (E_L - \bar{E}) \right\rangle \quad (\text{by Hermiticity}).$$

$$\begin{aligned} \bar{E}_{ij} = 2 & \left[\left\langle \left(\frac{\psi_{ij}}{\psi} + \frac{\psi_i \psi_j}{|\psi|^2} \right) (E_L - \bar{E}) \right\rangle - \left\langle \frac{\psi_i}{\psi} \right\rangle \bar{E}_j - \left\langle \frac{\psi_j}{\psi} \right\rangle \bar{E}_i \right] \\ & + \left\langle \frac{\psi_i}{\psi} E_{L,j} \right\rangle - \left\langle \frac{\psi_i}{\psi} \right\rangle \langle E_{L,j} \rangle + \left\langle \frac{\psi_j}{\psi} E_{L,i} \right\rangle - \left\langle \frac{\psi_j}{\psi} \right\rangle \langle E_{L,i} \rangle \end{aligned}$$

Leading divergences cancel! They cancel in the linear method too.

Remaining divergences do not seem to be problematic. Do 2nd order ones cancel?

Linear method for linear parameters

M. P. Nightingale and Melik-Alaverdian, PRL, **87**, 043401 (2001).

Given N basis functions, $\psi_i(\mathbf{R}_\sigma)$ (dets. \times Jastrow) find the lowest eigenstate of the (projected) Hamiltonian.

$$\mathcal{H}\psi_i(\mathbf{R}_\sigma) = \sum_j \psi_j(\mathbf{R}_\sigma) E_{ji} \quad \text{or} \quad \mathbf{B}' = \mathbf{B}\mathbf{E}$$

where $B_{\sigma i} = \psi_i(\mathbf{R}_\sigma)$ and $B'_{\sigma i} = \mathcal{H}\psi_i(\mathbf{R}_\sigma)$ are $N_{\text{MC}} \times N$ and \mathbf{E} is $N \times N$. The least-squares solution of these overdetermined Eqs. is

$$\mathbf{E} = (\mathbf{B}^T \mathbf{B})^{-1} (\mathbf{B}^T \mathbf{B}') = \mathbf{S}^{-1} \mathbf{H}$$

$$S_{ij} = \left\langle \frac{\psi_i}{\psi_0} \frac{\psi_j}{\psi_0} \right\rangle_{\psi_0^2}, \quad H_{ij} = \left\langle \frac{\psi_i}{\psi_0} \frac{\mathcal{H}\psi_j}{\psi_0} \right\rangle_{\psi_0^2}$$

and $\sum_{i=1}^N c_i \psi_i$ is an eigenstate of \hat{H} , with eigenvalue ϵ , if \mathbf{c} satisfies

$$\mathbf{E}\mathbf{c} = \epsilon\mathbf{c} \quad \text{i.e.} \quad \mathbf{H}\mathbf{c} = \epsilon\mathbf{S}\mathbf{c}$$

Note: H_{ij} evaluated on a finite MC sample is **not** symmetric, though the true H_{ij} , and H_{ij} that minimizes sample energy, are symmetric. Nonsymmetric H_{ij} needed for **zero-variance** property in limit that functions span an invariant subspace.

Linear method for nonlinear parameters

CJU, Toulouse, Filippi, Sorella, Hennig PRL (2007);
Toulouse, CJU, JCP (2007).

Make linear order Taylor expansion of wavefunction:

$$\psi = \psi_0 + \sum_{i=1}^{N_{\text{parm}}} \delta p_i \psi_i$$

ψ_0 = current wave function

ψ = next wave function

ψ_i = derivative of ψ_0 wrt i^{th} parameter.

No unique way to obtain new parameters.

The simplest procedure: is $p_i^{\text{new}} = p_i + \delta p_i$.

More complicated procedure: fit wave function form to the optimal linear combination.

Simpler, yet efficient approach, freedom of norm to make linear approximation better

$$\begin{aligned}\psi(\mathbf{p}, \mathbf{R}) &= N(\mathbf{p}) \phi(\mathbf{p}, \mathbf{R}), \quad N(\mathbf{p}_0) = 1 \\ \psi_i(\mathbf{p}_0, \mathbf{R}) &= \phi_i(\mathbf{p}_0, \mathbf{R}) + N_i(\mathbf{p}_0, \mathbf{R}) \phi(\mathbf{p}_0, \mathbf{R})\end{aligned}$$

Perturbation Theory in an arbitrary nonorthogonal basis (Toulouse, CJU, JCP (2007))

Given a Hamiltonian \hat{H} and an arbitrary nonorthogonal basis, $\{|\Psi_i\rangle\}$, use perturbation theory to get approximate eigenstates of \hat{H} .

Define dual basis: $\langle\bar{\Psi}_i|\Psi_j\rangle = \delta_{ij}$ and zeroth order Hamiltonian, $\hat{H}^{(0)}$:

$$\langle\bar{\Psi}_i| = \sum_{j=0}^{N_{\text{opt}}} (\mathbf{S}^{-1})_{ij} \langle\Psi_j|, \quad \hat{H}^{(0)} = \sum_{i=0}^{N_{\text{opt}}} E_i |\Psi_i\rangle \langle\bar{\Psi}_i|$$

First order perturbation correction is

$$|\Psi^{(1)}\rangle = - \sum_{i=1}^{N_{\text{opt}}} |\Psi_i\rangle \sum_{j=1}^{N_{\text{opt}}} (\mathbf{S}^{-1})_{ij} \frac{\langle\Psi_j|\hat{H}|\Psi_0\rangle}{E_i - E_0}$$

Want \hat{H} and $\hat{H}^{(0)}$ close – choose E_i so \hat{H} and $\hat{H}^{(0)}$ have same diagonals

$$E_i = \frac{\langle\Psi_i|\hat{H}|\Psi_i\rangle}{\langle\Psi_i|\Psi_i\rangle}$$

If the E_i are evaluated without the Jastrow factor then this is the same as the perturbative eff. fluct. pot. (EFP) method of Scemama and Filippi.

Stabilization

If far from the minimum, or, N_{MC} , is small, then the Hessian, \bar{E}_{ij} , need not be positive definite (whereas variance-minimization Levenberg-Marquardt \bar{E}_{ij} is positive definite).

Even for positive definite \bar{E}_{ij} , the new parameter values may make the wave function worse if quadratic approximation is not good.

Determine eigenvalues, ϵ , of \bar{E}_{ij} and add to it $(\max(0, -\epsilon_{\min}) + a_{\text{diag}})\mathcal{I}$. This shifts the eigenvalues by the added constant. As a_{diag} is increased, the proposed parameter changes become smaller and rotate from the Newtonian direction to the steepest descent direction, but in practice a_{diag} is tiny.

The linear method and the perturbative method can be approximately recast into the Newton method. Consequently we can use the same idea for the linear and perturbative methods too.

Stabilization with Correlated Sampling

Each method has a parameter a_{diag} that automatically adjusts to make the method totally stable:

1. Do a MC run to compute the gradient and the Hessian (or overlap and Hamiltonian).
2. Using the above gradient and Hessian (or overlap and Hamiltonian), use 3 different values of a_{diag} to predict 3 different sets of updated parameters.
3. Do a short correlated sampling run for the 3 different wave functions to compute the energy differences for the 3 wave functions more accurately than the energies themselves.
4. Fit a parabola through the 3 energies to find the optimal a_{diag} .
5. Use this optimal a_{diag} to predict a new wave function, using the gradient and Hessian computed in step 1.
6. Loop back

Comparison of Newton, linear and perturbative methods

Programming effort and cost per iteration:

1. Newton method requires ψ , ψ_i , ψ_{ij} , $\hat{H}\psi$, $\hat{H}\psi_i$.
2. Linear method requires ψ , ψ_i , $\hat{H}\psi$, $\hat{H}\psi_i$.
3. Perturbative method requires ψ , ψ_i , $\hat{H}\psi$, $\hat{H}\psi_i$.
Perturbative method with approximate denominators requires ψ , ψ_i .

Correct to which order:

1. Newton has correct 2nd order terms in E .
2. Linear does not have correct 2nd order terms in E but has all order terms coming from linear terms in Ψ .

Convergence with number of iterations:

1. Newton and linear methods converge in 1-5 iterations for all parameters (CSF, orbital and Jastrow).
2. Perturbative method converges in 1-5 iterations for CSF and orbital parameters but is very slow for Jastrow because eigenvalues of Hessian for Jastrow span 9-12 orders of magnitude.

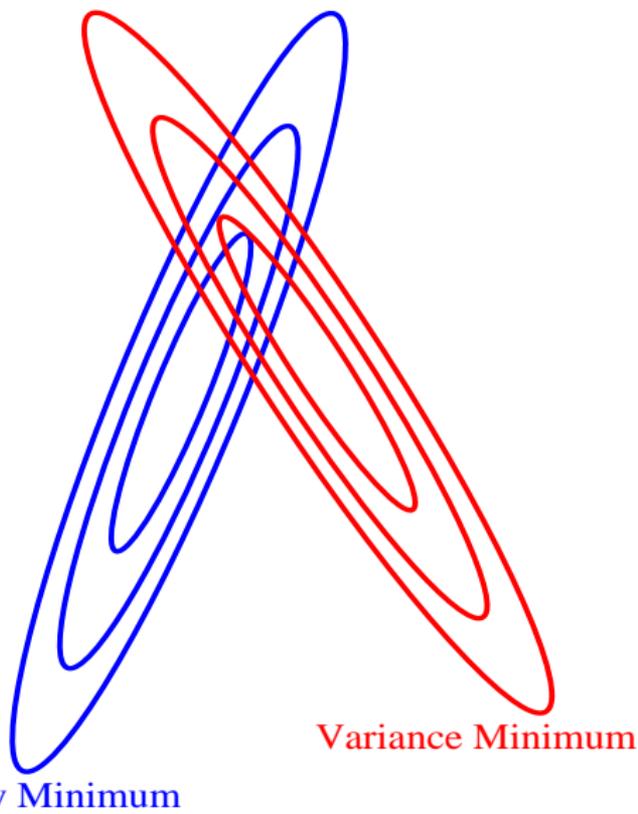
Things to note

Eigenvalues of \bar{E}_{ij} for Jastrow parameters span 11 orders of magnitude. So steepest descent would be horribly slow to converge!

Linear and Newton methods can be used for all parameters, even basis-set exponents.

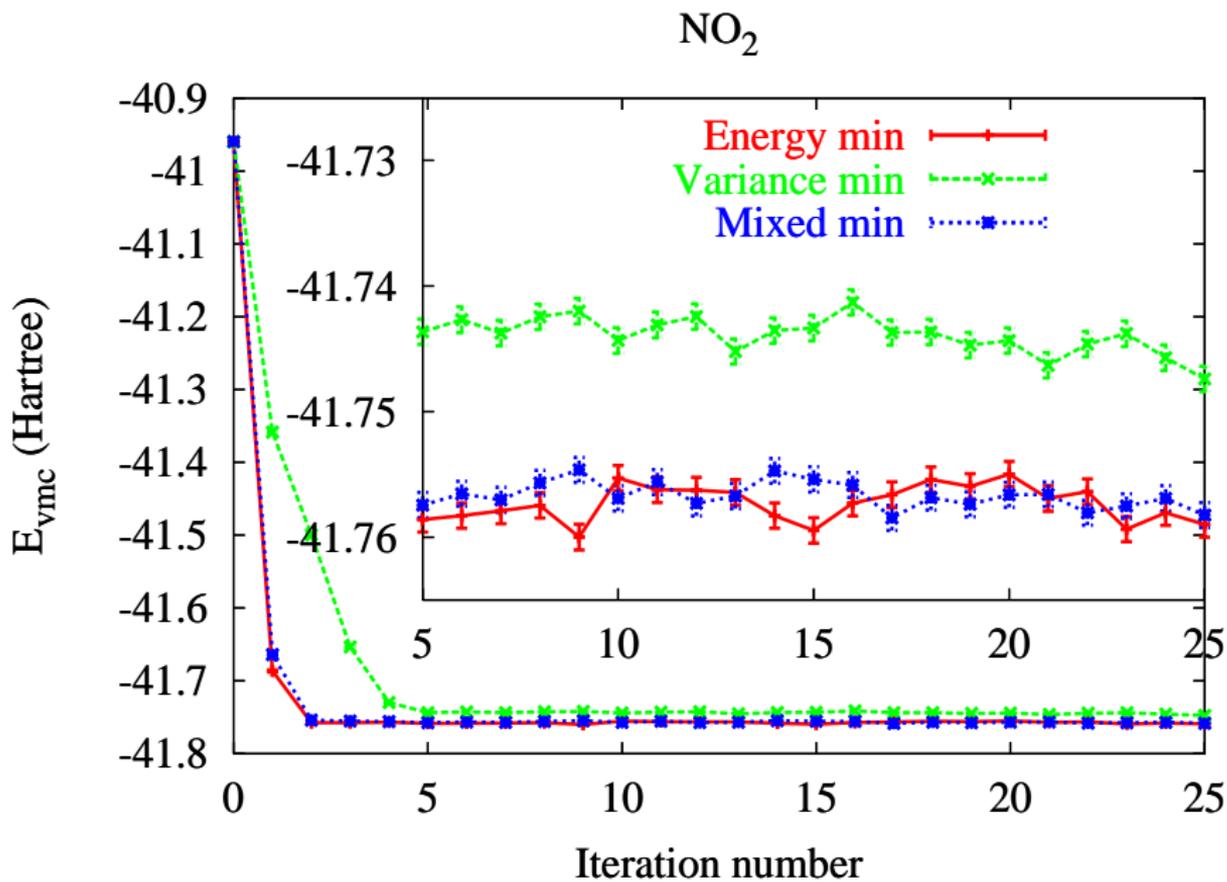
Take Home Message: Any method that attempts to minimize the energy, by minimizing the energy evaluated on a set of MC points, will require a very large sample and be highly inefficient. Each of the 3 methods presented above avoids doing this.

Optimization of linear combination of energy and variance

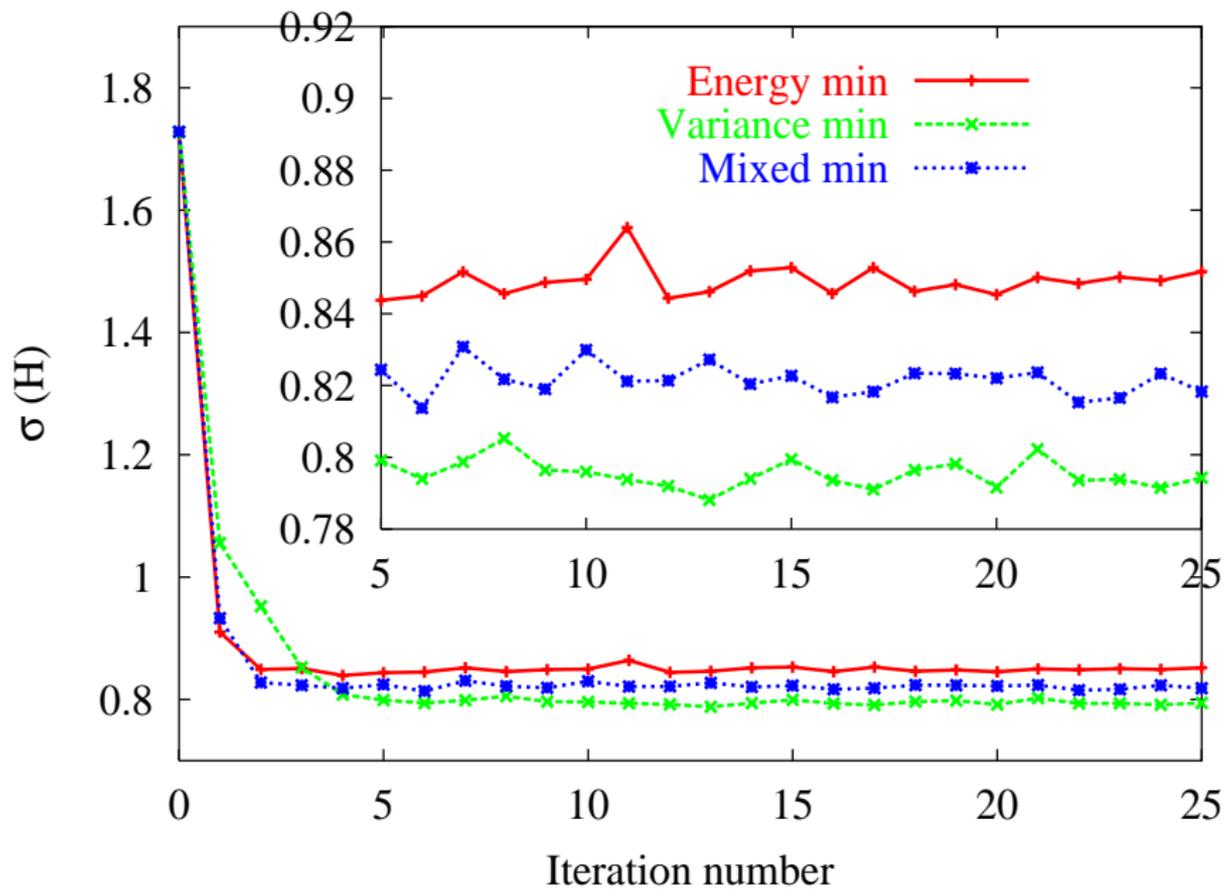


1. Can reduce the variance, without sacrificing appreciably the energy, by minimizing a linear combination, particularly since the ratio of hard to soft directions is 11 orders of magnitude.
2. Easy to do – gradient and Hessian of linear combination are linear combinations of the gradient and Hessian.
3. Measure of efficiency of the wave function is $\sigma^2 T_{\text{corr}}$.

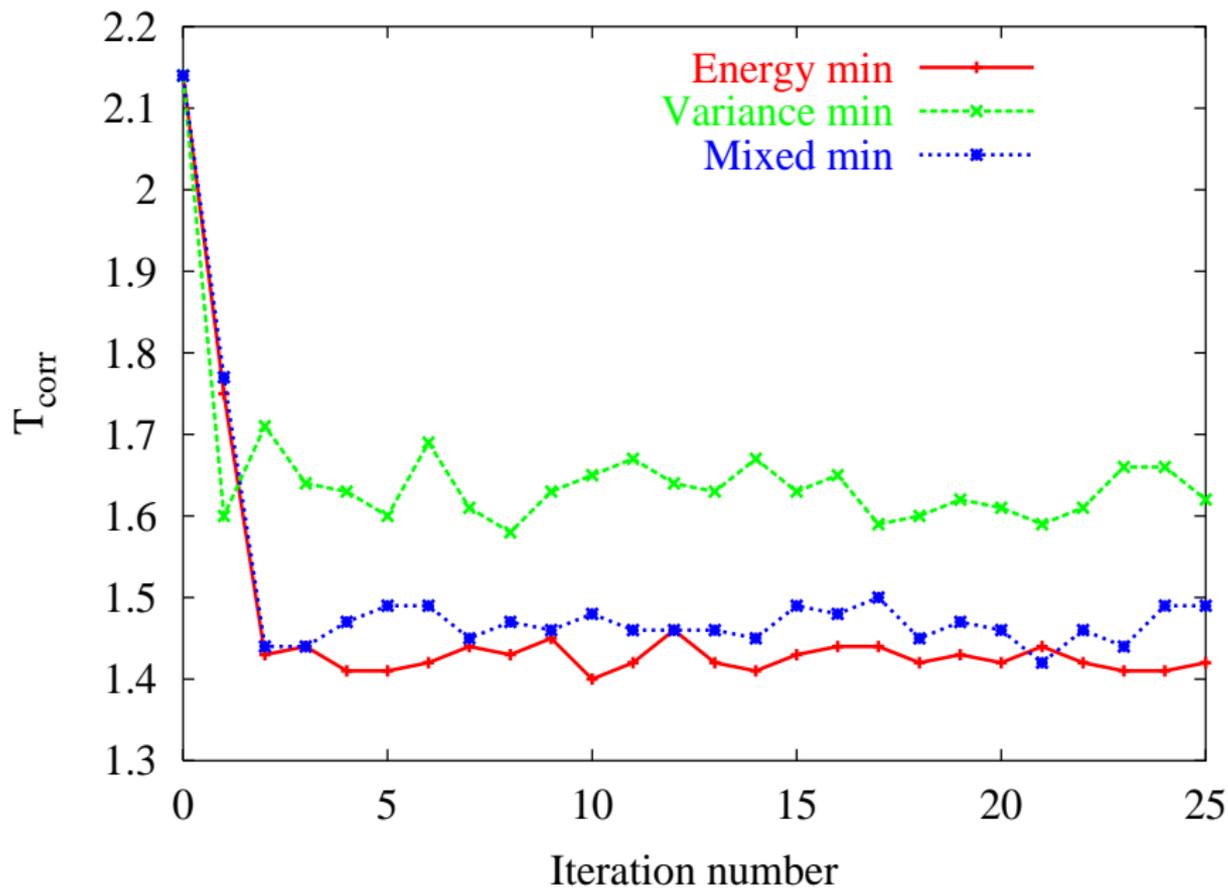
Convergence of energy of NO₂



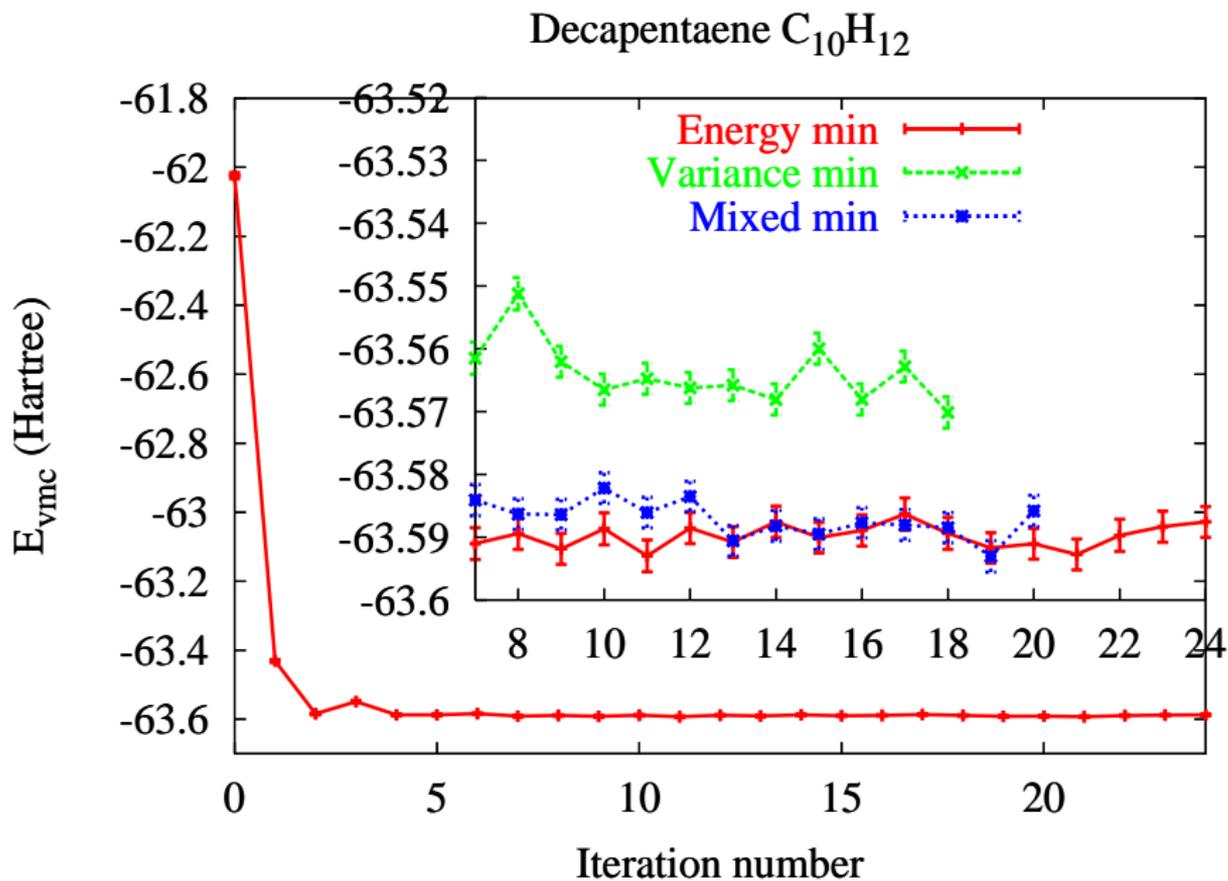
Convergence of energy fluctuations, σ , of NO₂



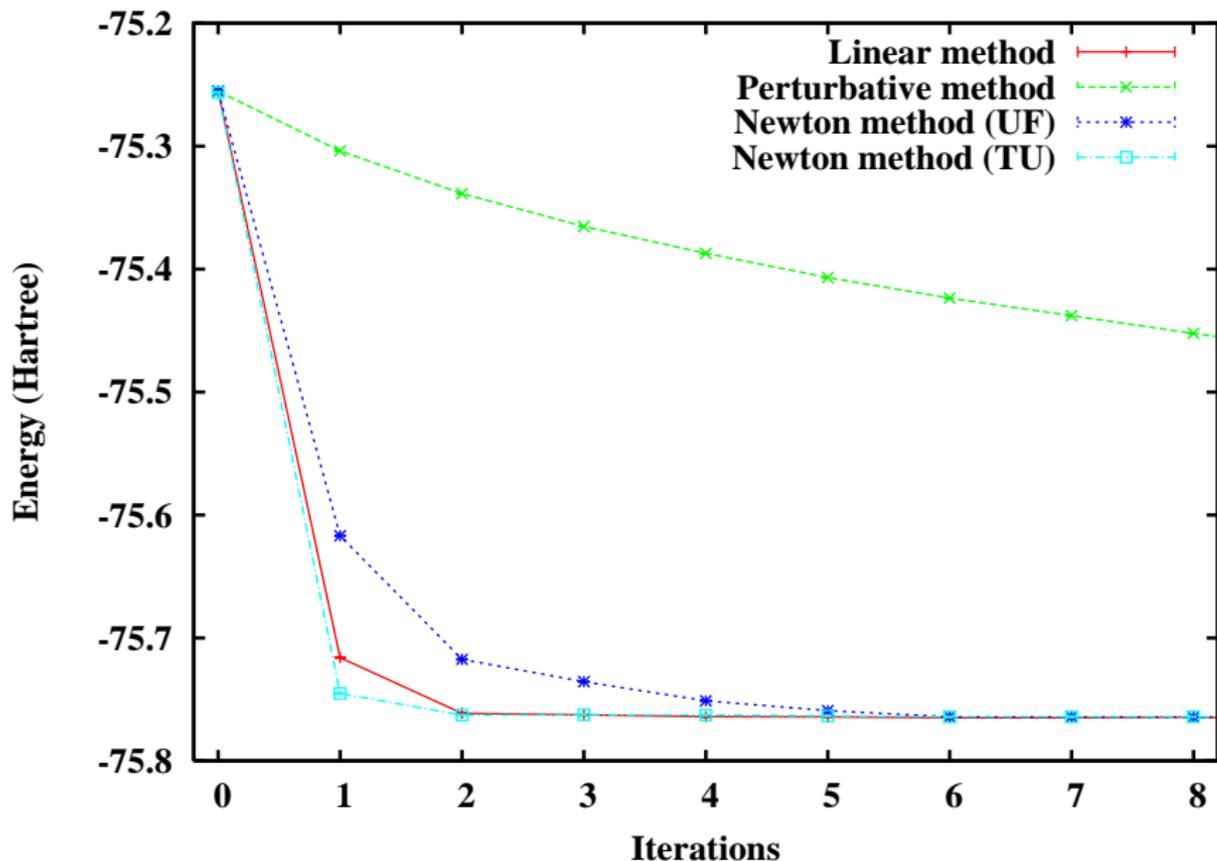
Convergence of autocorrelation time, T_{corr} , of NO_2



Convergence of energy of decapentaene $C_{10}H_{12}$

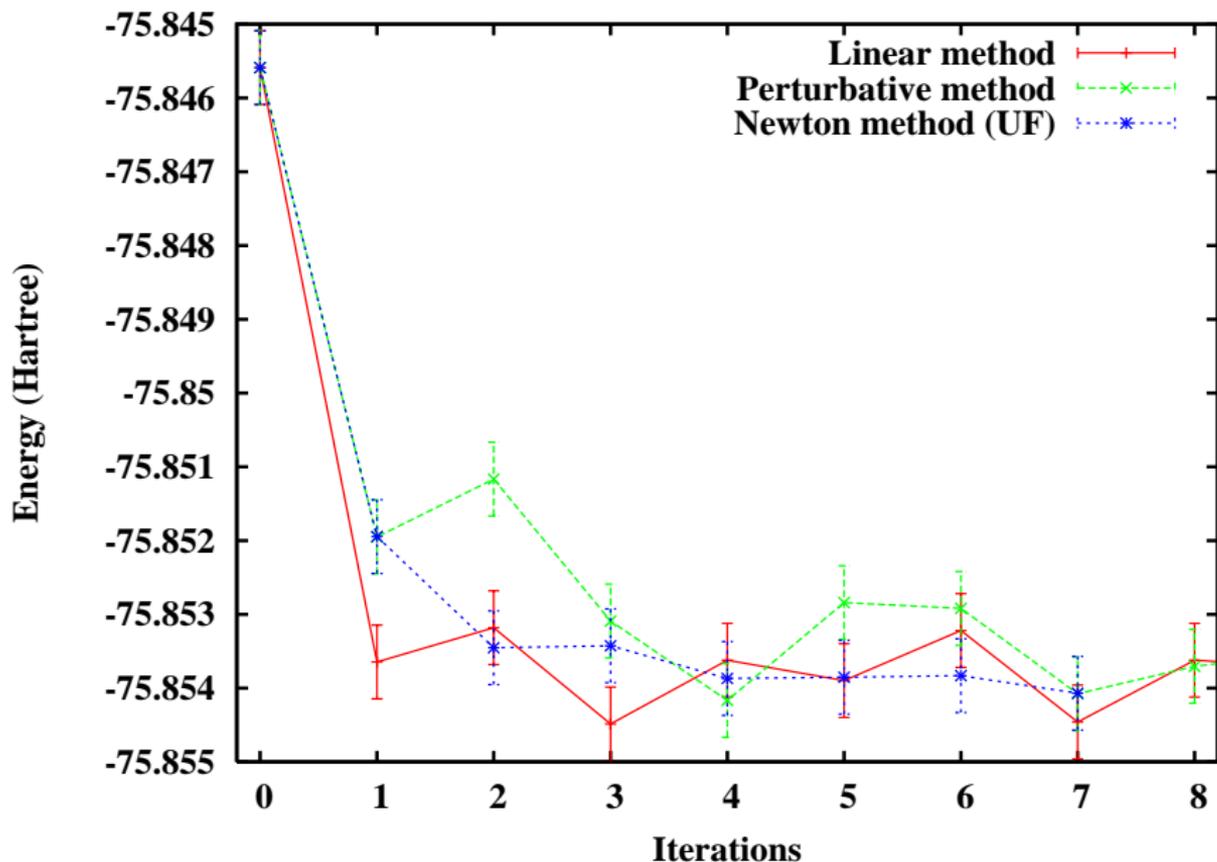


Optimization of Jastrow of all-electron C_2



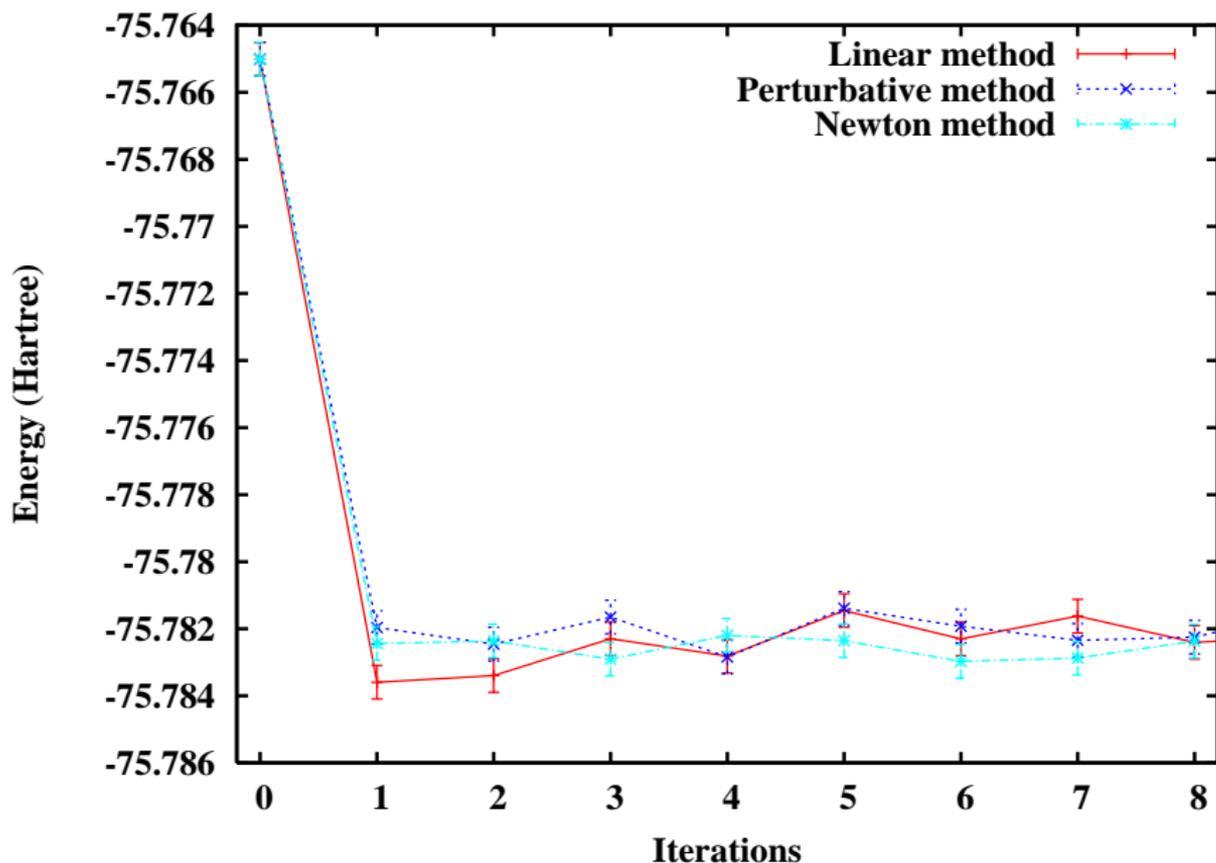
Linear and Newton methods converge in 2-6 iterations, perturbative is slow!

Optimization of CSF coefs. of all-electron C_2



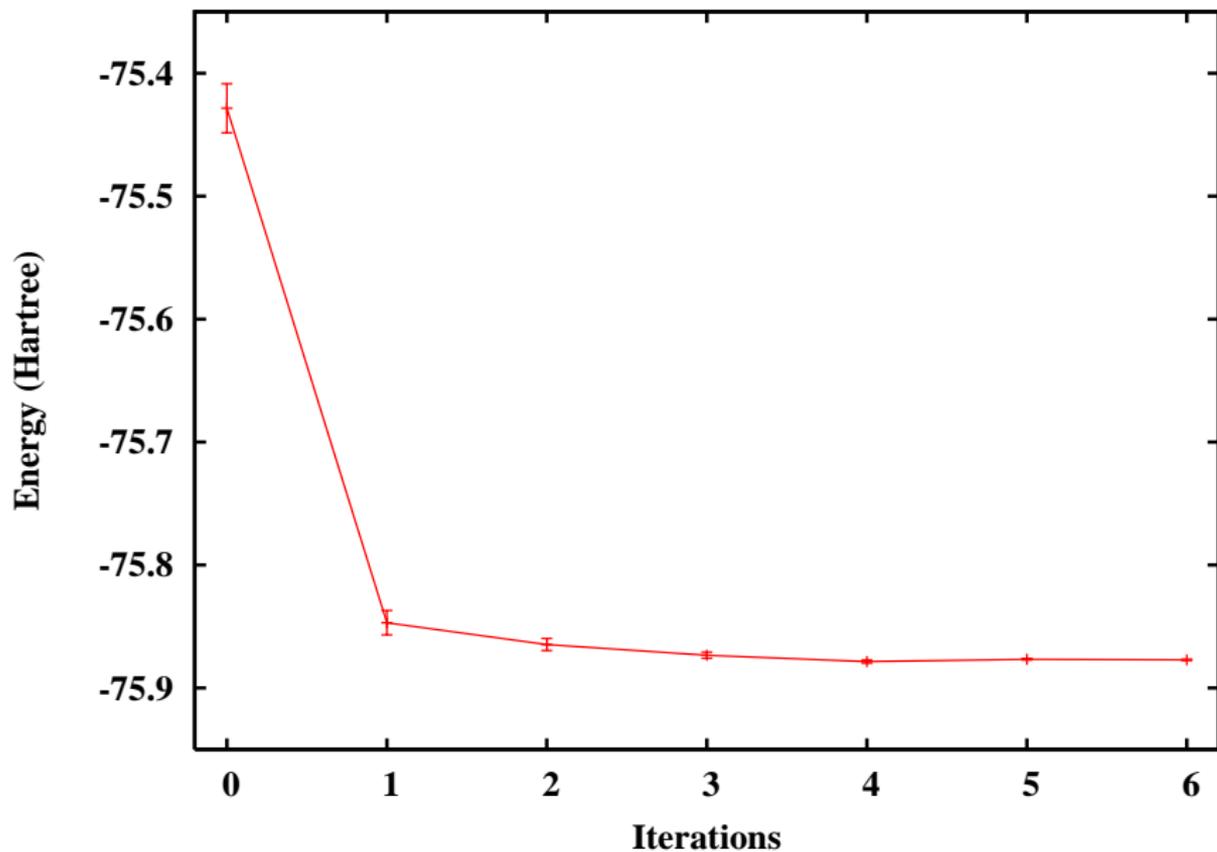
All 3 methods converge in 1-3 iteration!

Optimization of orbitals of all-electron C_2

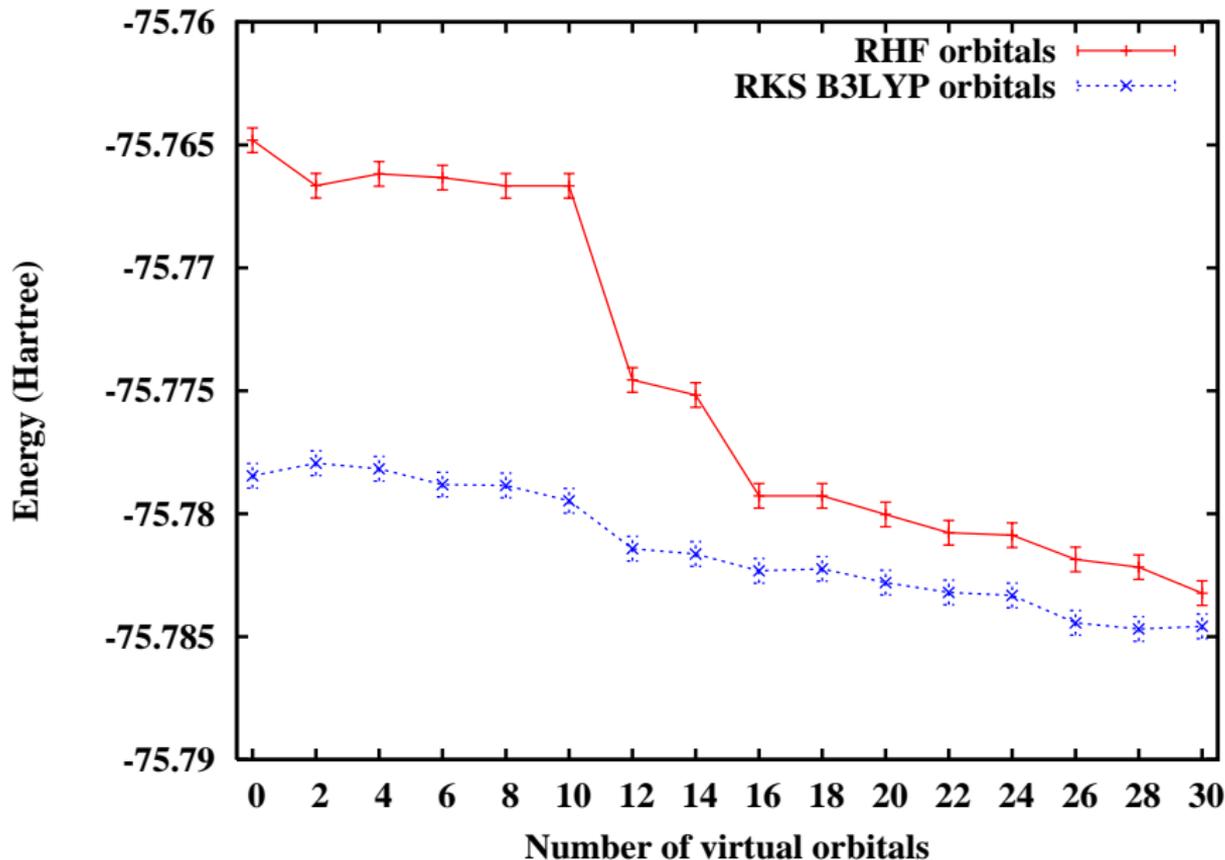


All 3 methods converge in 1 iteration!

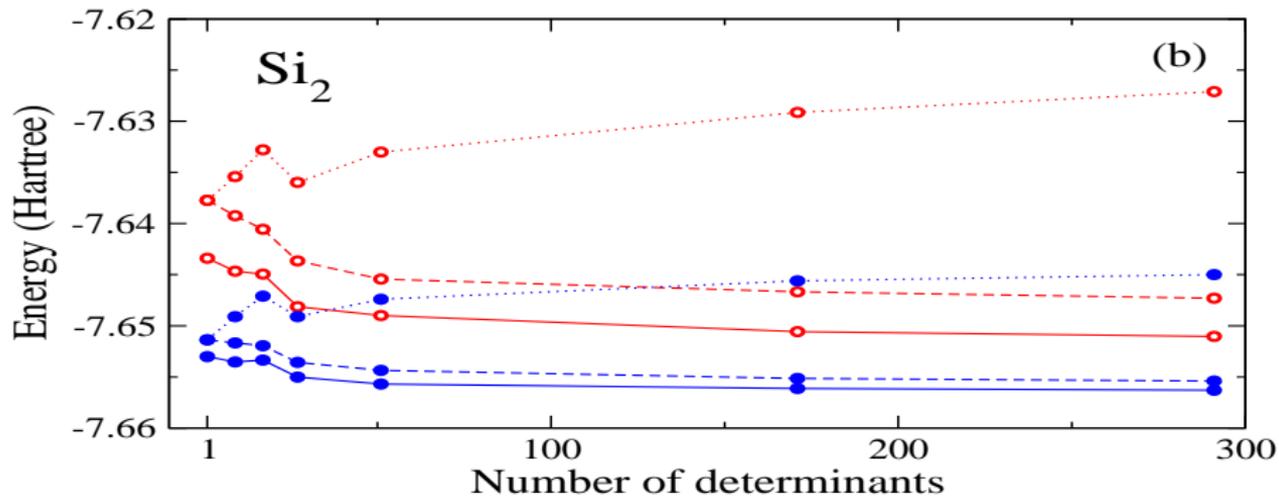
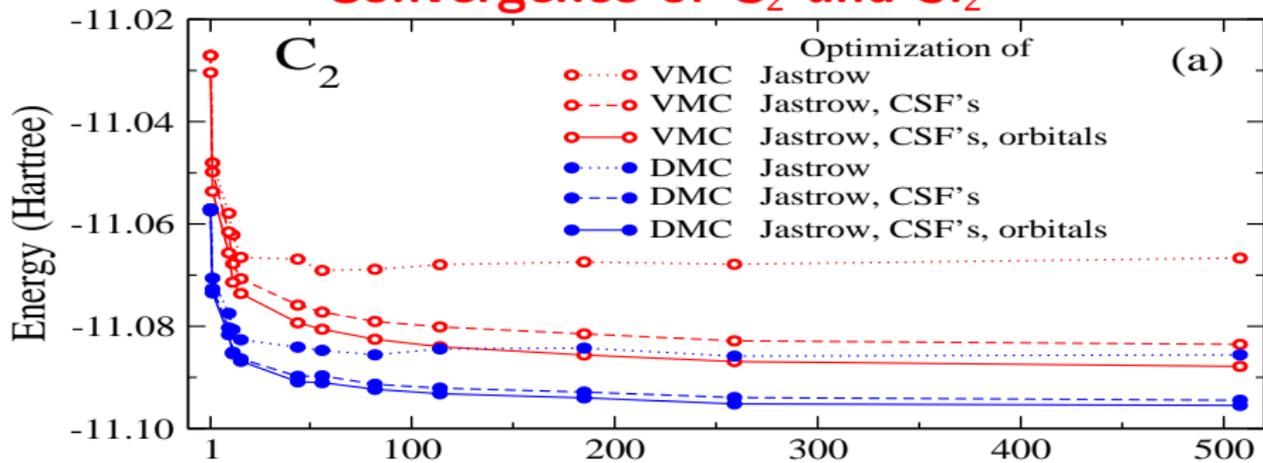
Simultaneous optimization of Jastrow, CSFs and orbitals of all-electron C_2 with linear method

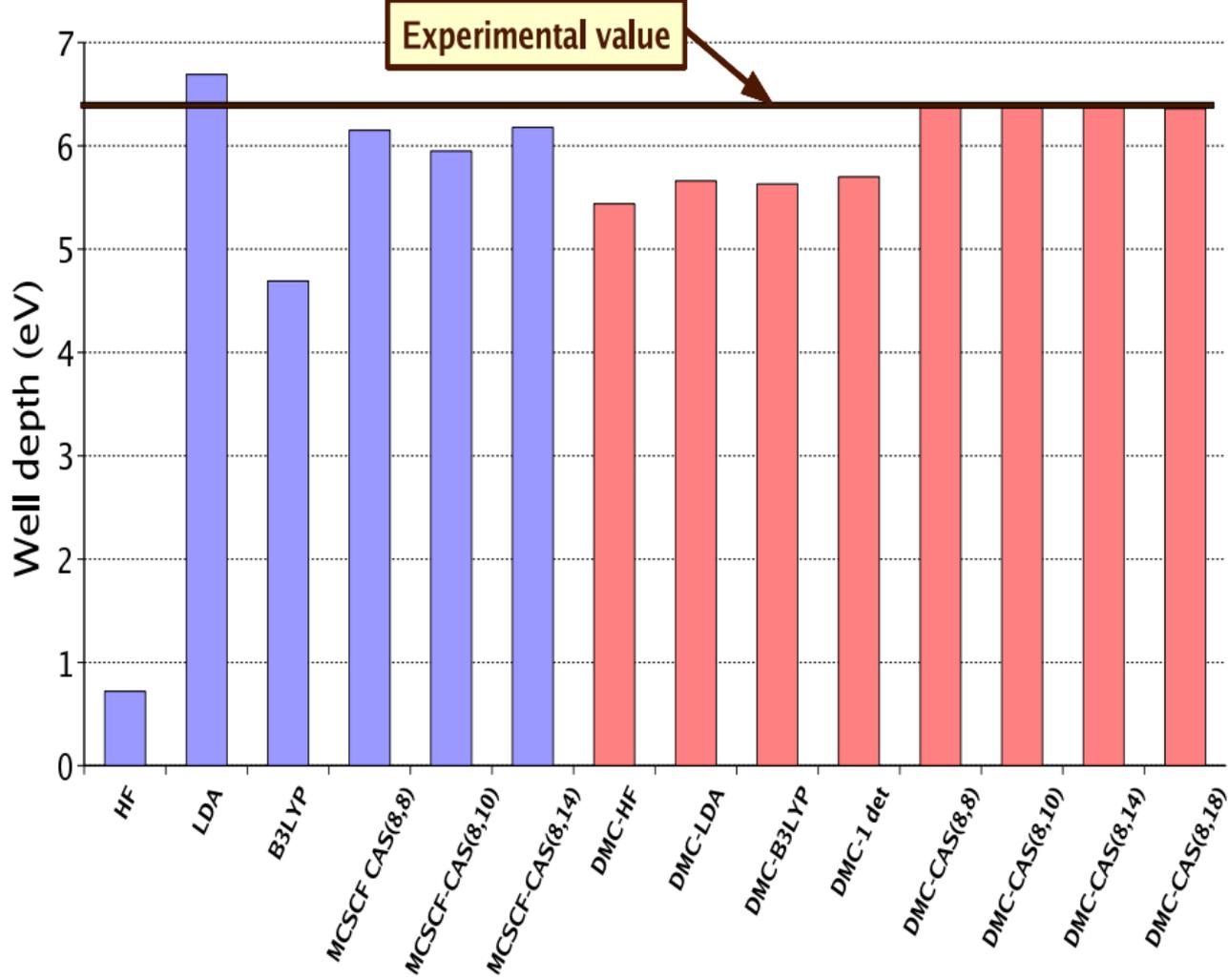


Convergence of energy of all-electron C_2 vs N_{orb}^{virt}

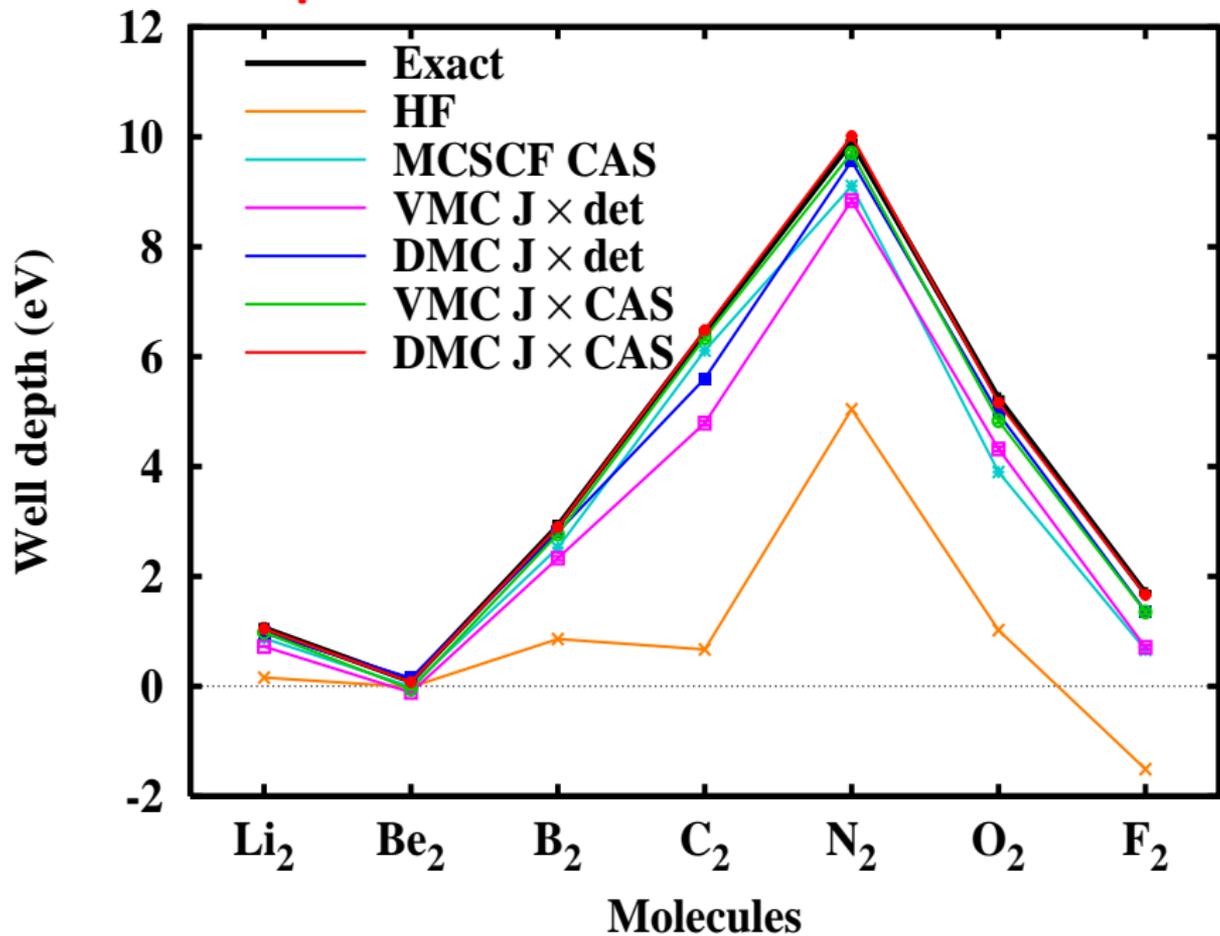


Convergence of C_2 and Si_2

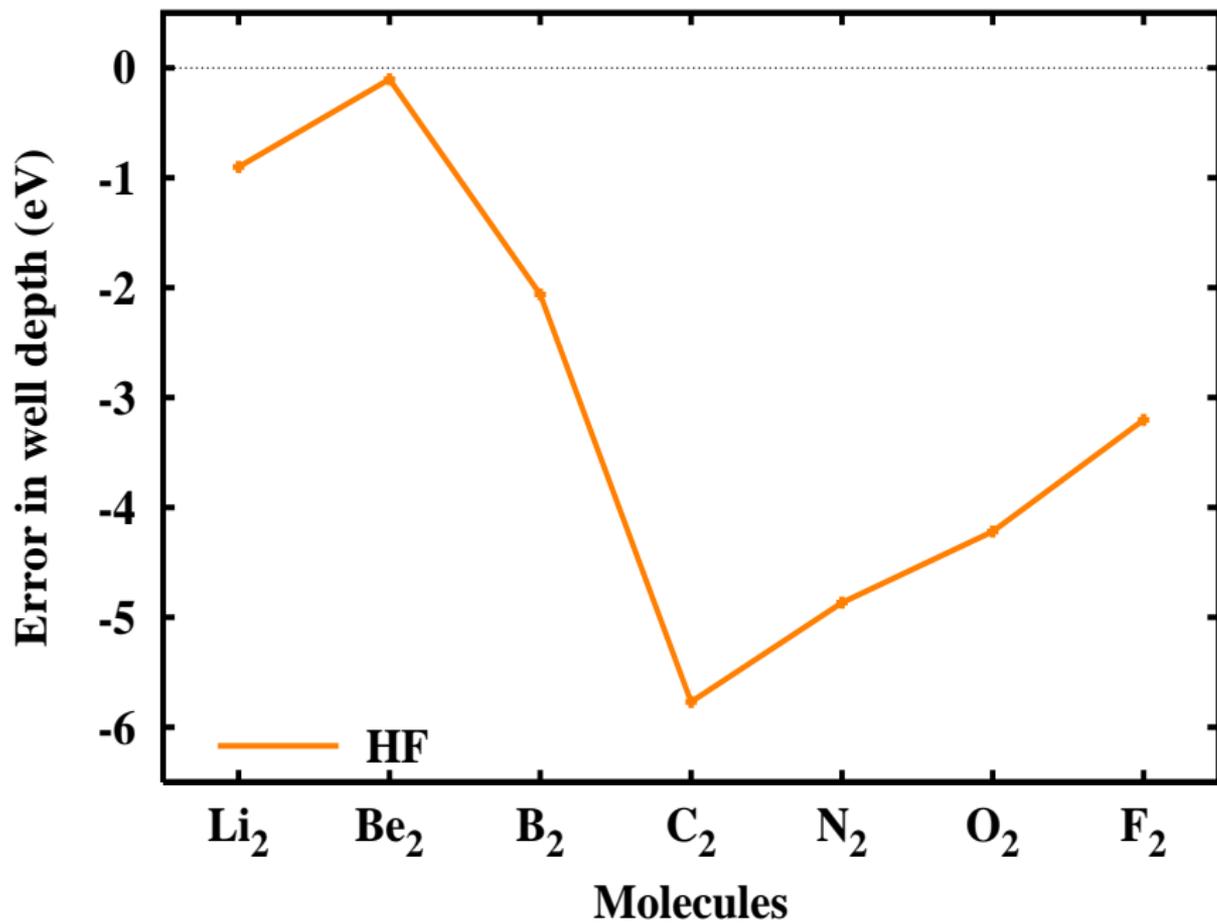




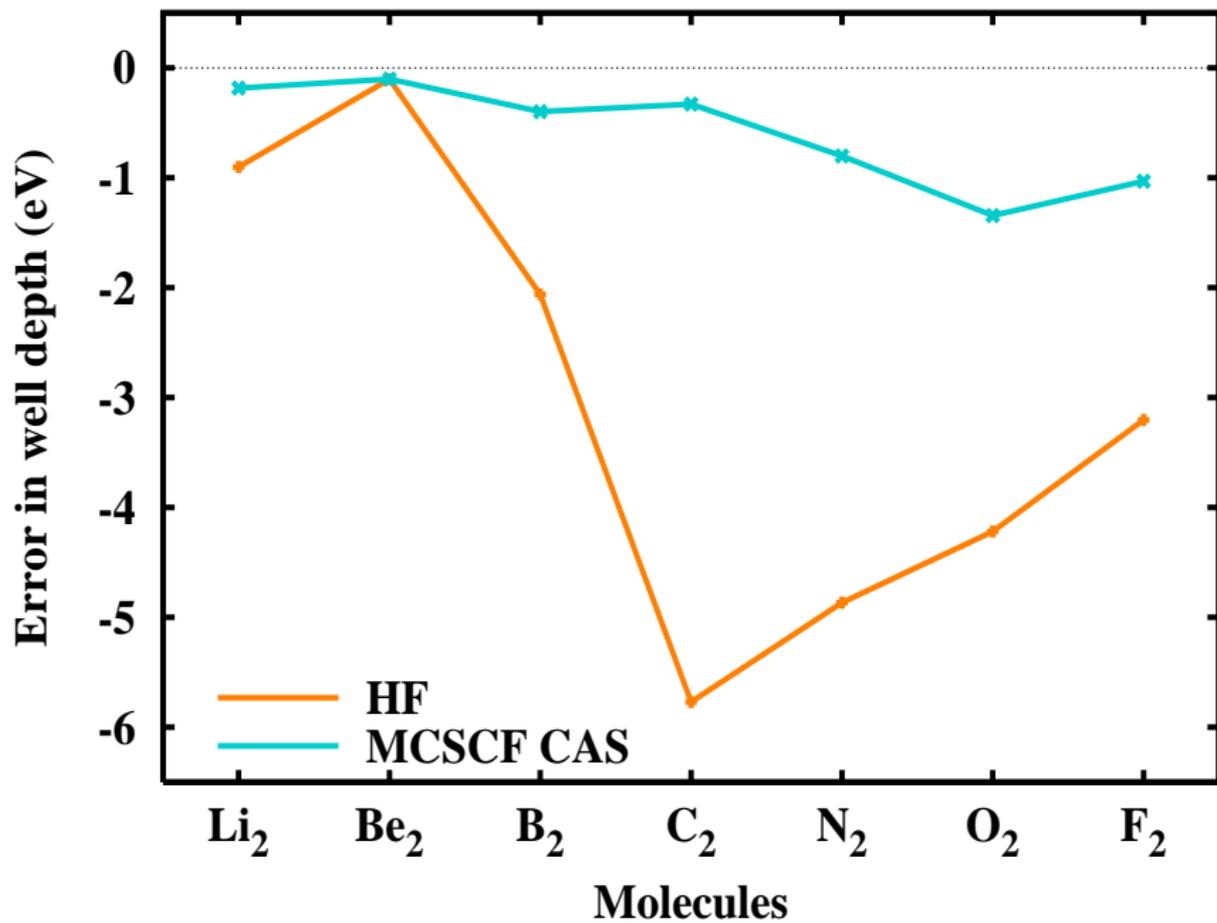
Well-Depth of 1st-Row Diatomic Molecules



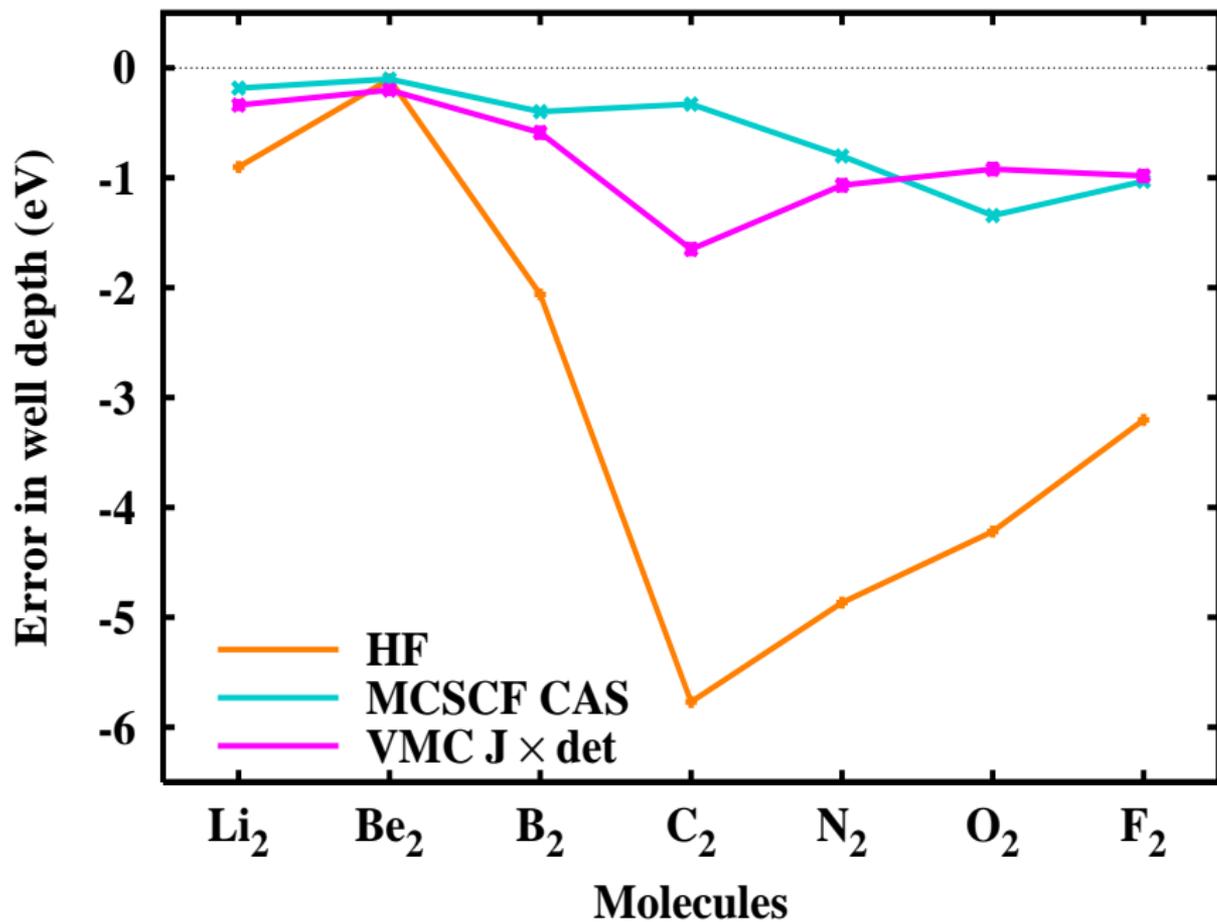
Error in Well-Depth of 1st-Row Diatomic Molecules



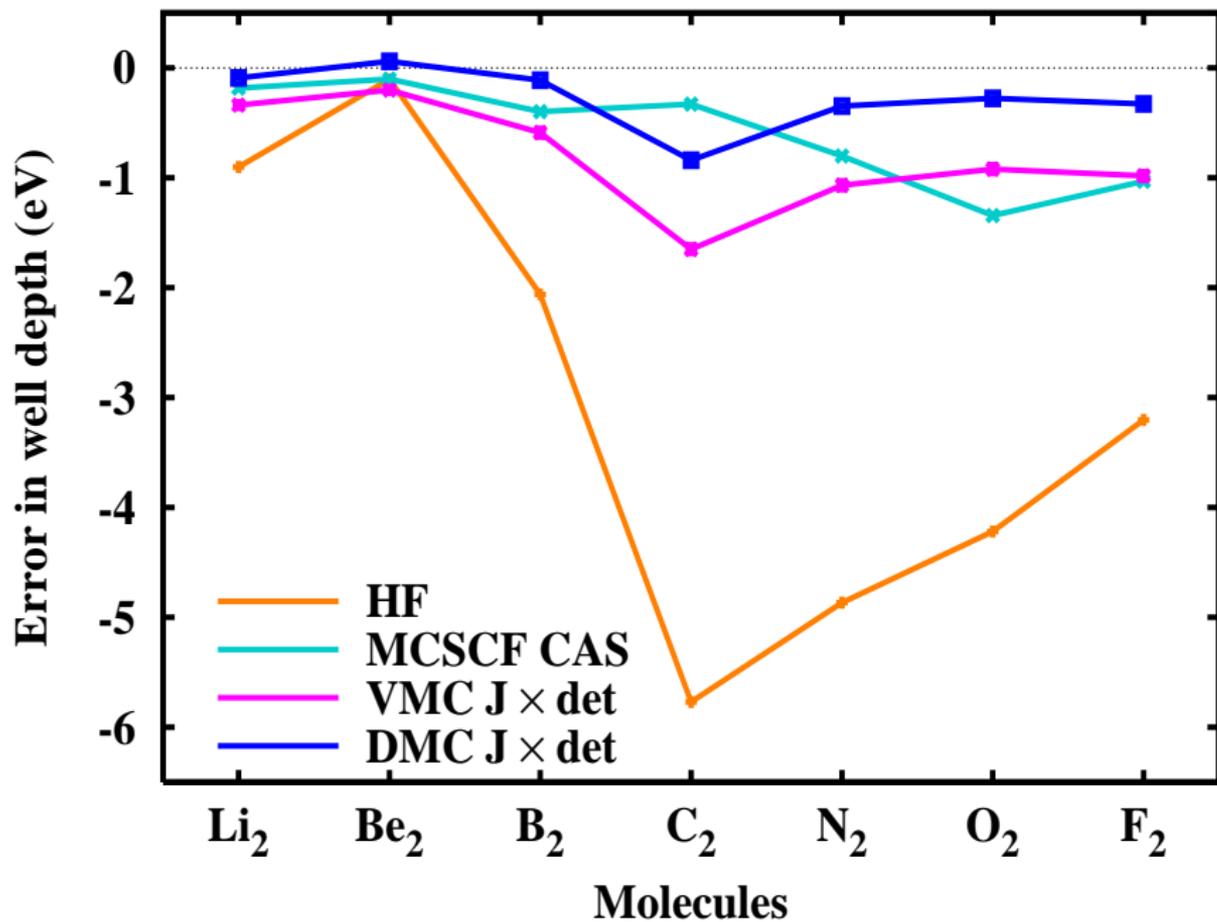
Error in Well-Depth of 1st-Row Diatomic Molecules



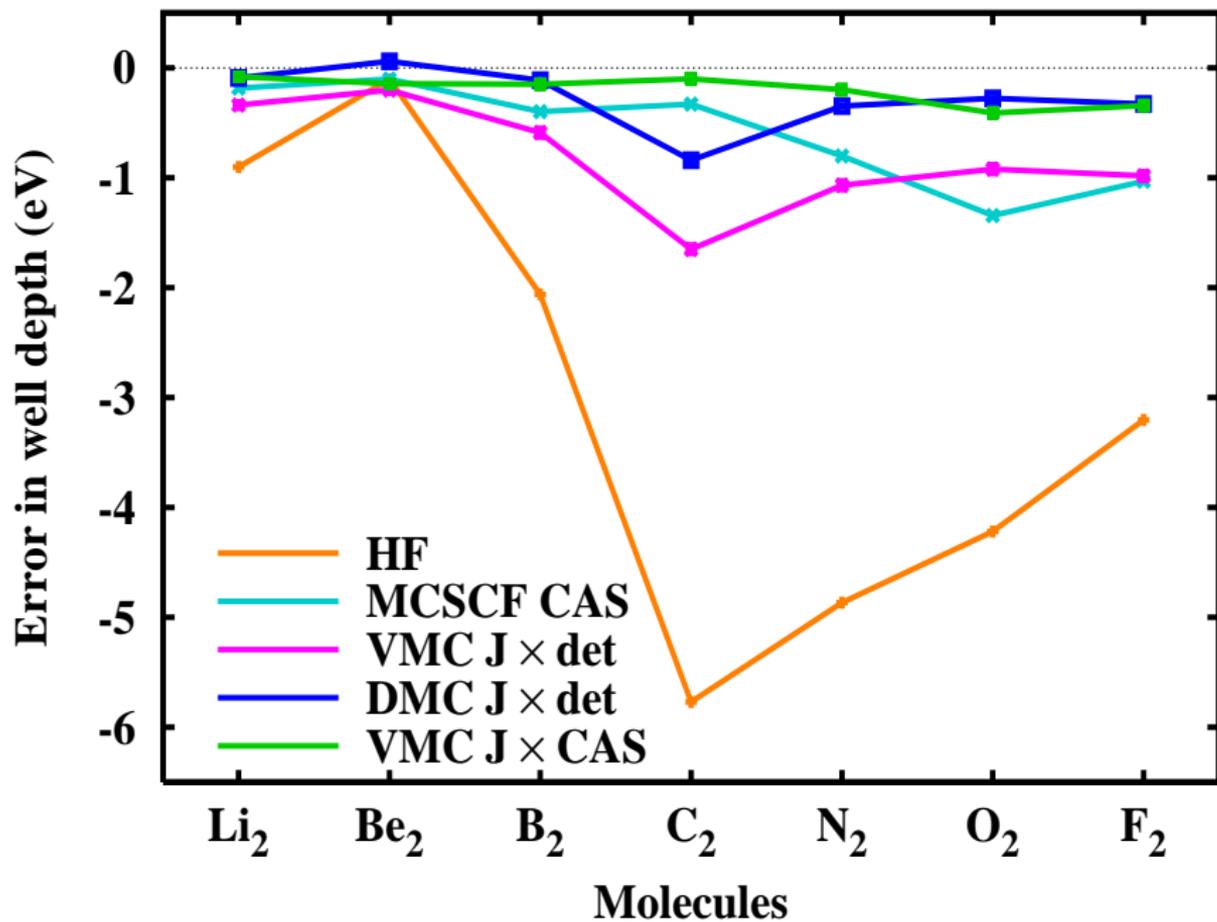
Error in Well-Depth of 1st-Row Diatomic Molecules



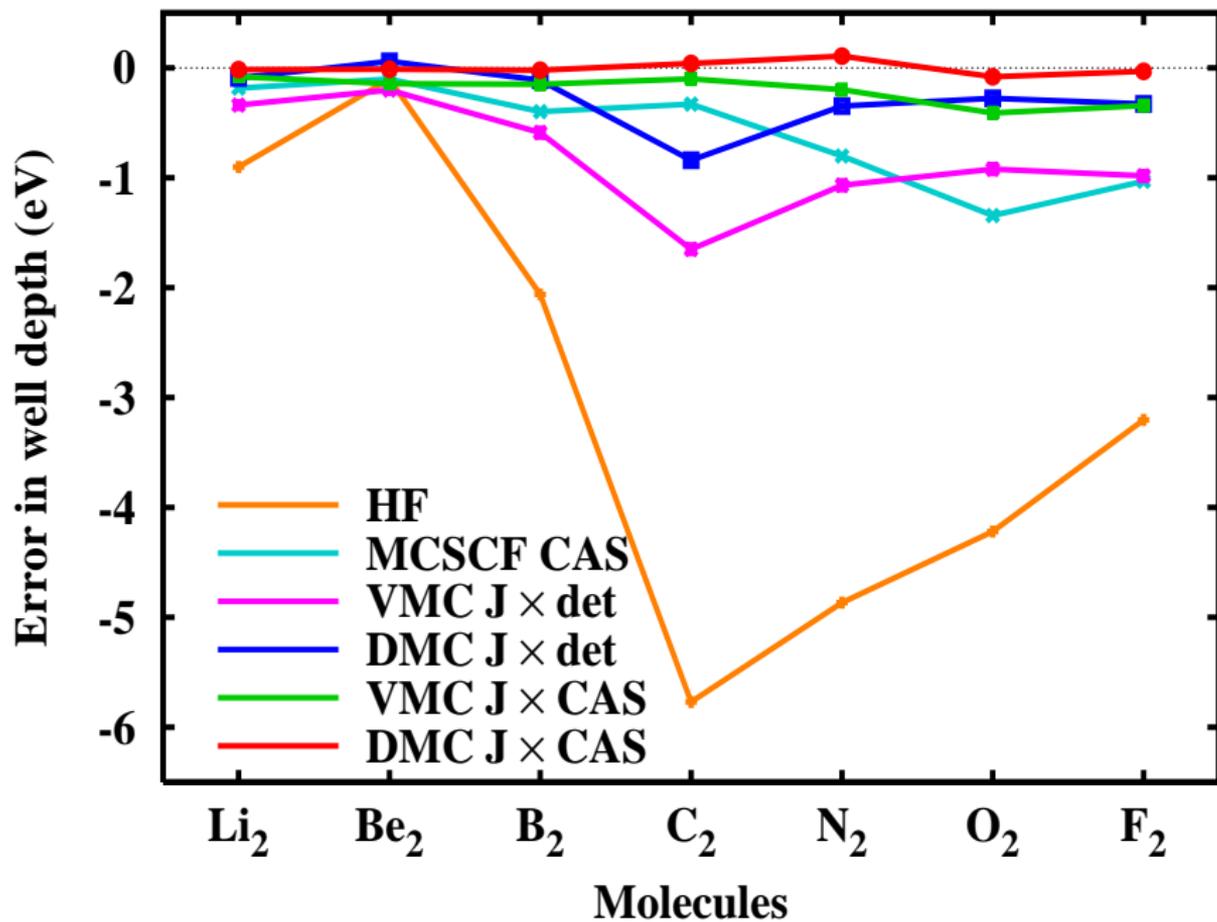
Error in Well-Depth of 1st-Row Diatomic Molecules



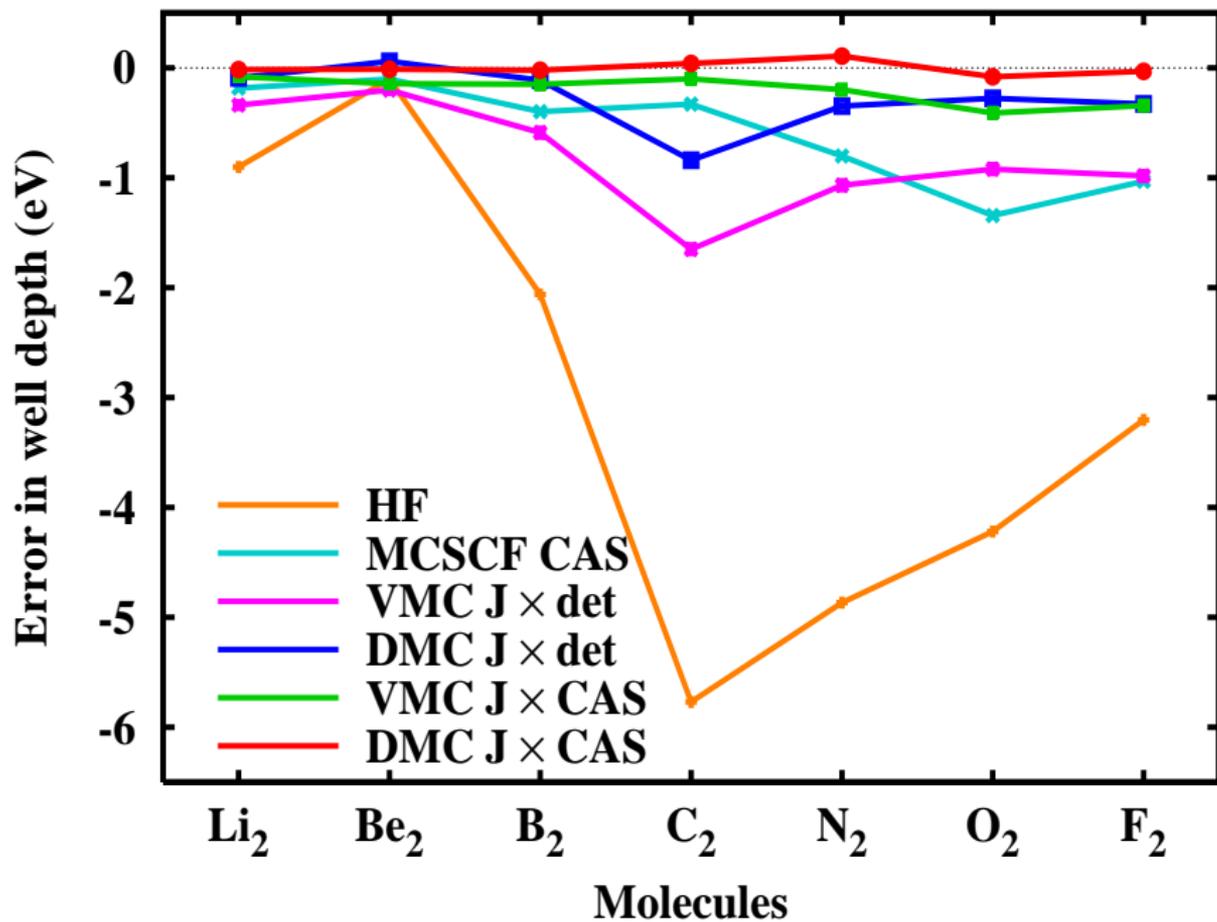
Error in Well-Depth of 1st-Row Diatomic Molecules



Error in Well-Depth of 1st-Row Diatomic Molecules



Error in Well-Depth of 1st-Row Diatomic Molecules

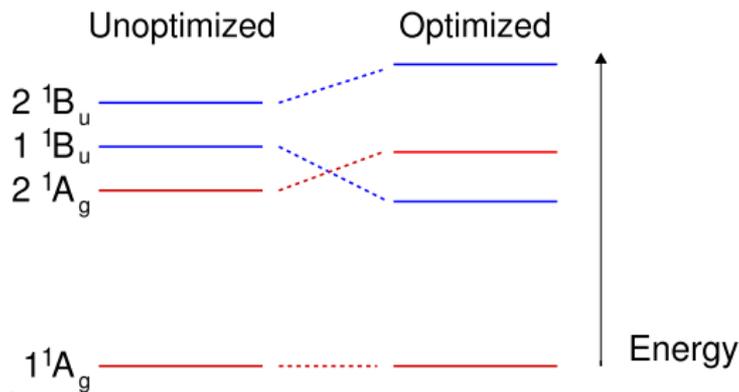


EFP method for ground and excited states

F. Schautz and C. Filippi, JCP **120**, 10931 (2004)

- Excitations of ethene C_2H_4 → Up to 858 optimized parameters

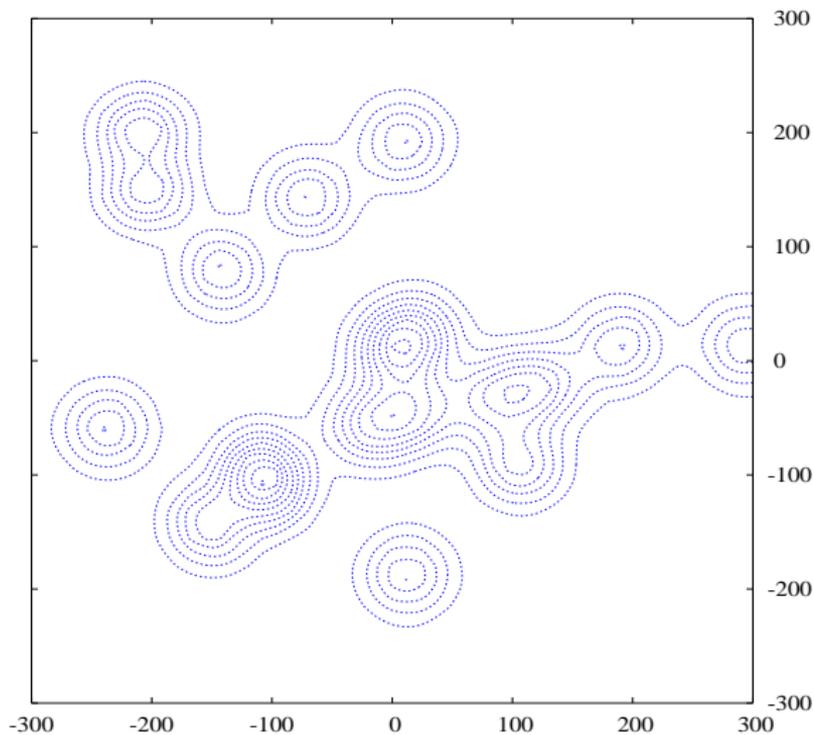
DMC excitation energies



State	Unoptimized	Optimized	expt.	(eV)
$1\ ^1B_{1u}$	8.45(2)	7.93(2)	>7.7	
$2\ ^1A_g$	7.96(2)	8.36(2)	8.29	
$2\ ^1B_{1u}$	9.05(2)	9.37(2)	9.33	

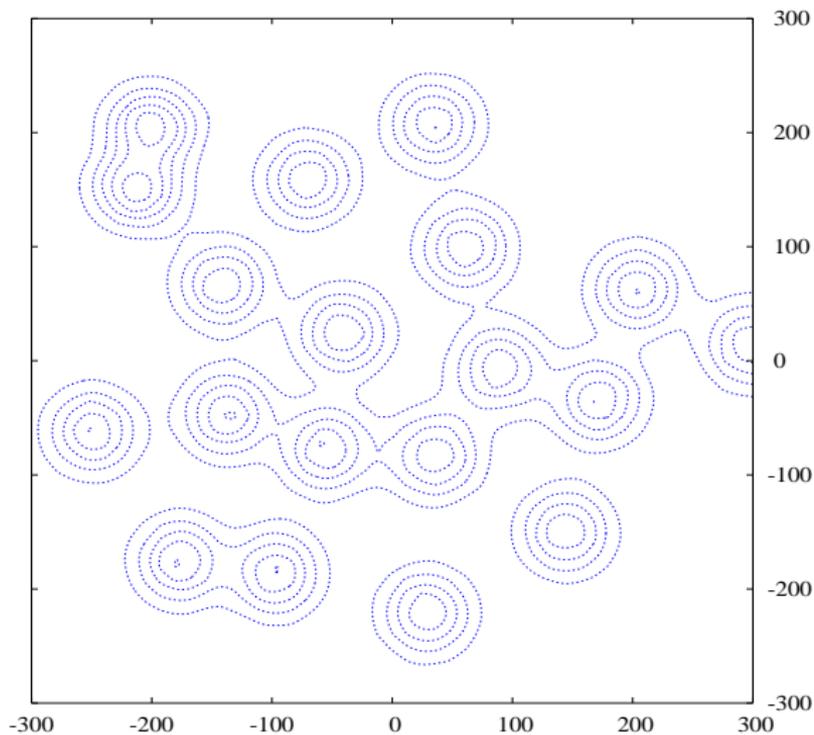
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 0



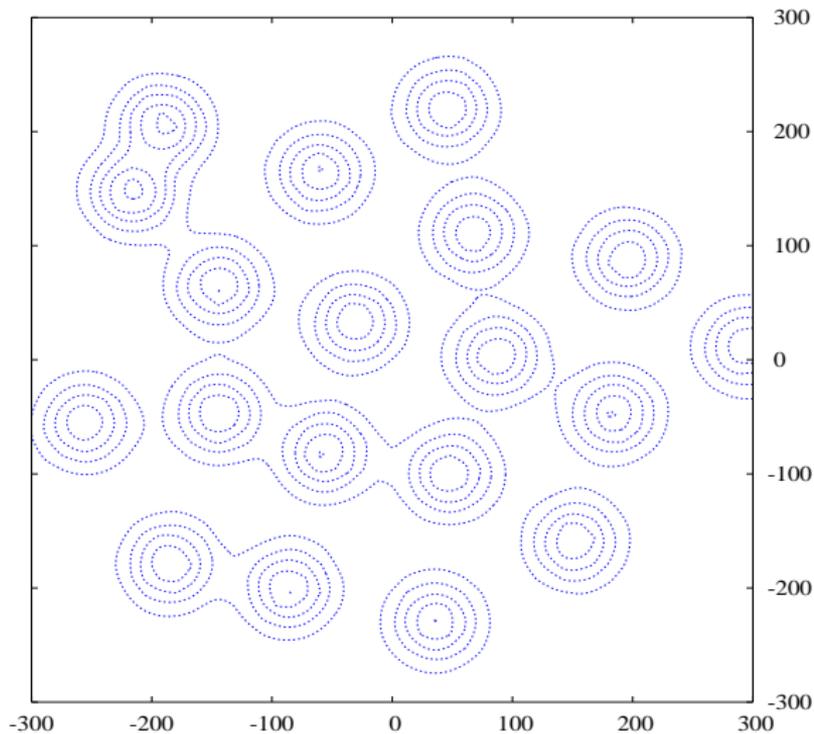
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 2



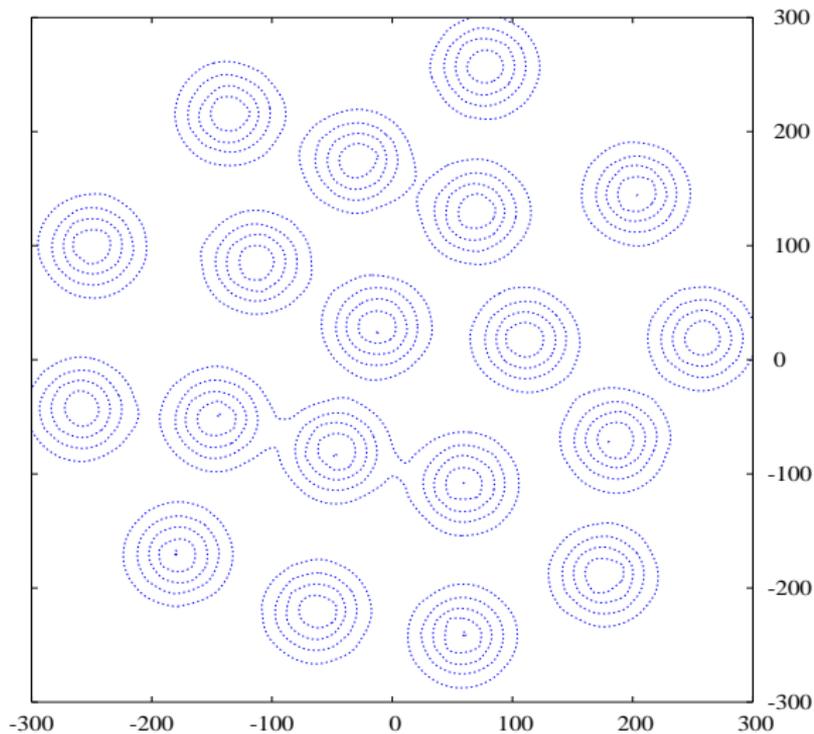
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 5



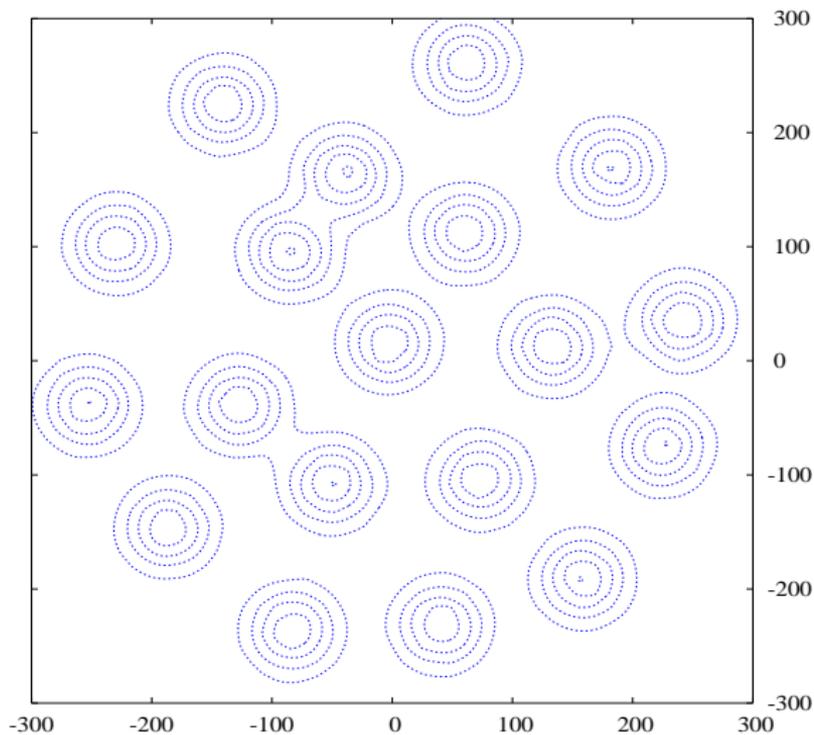
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 7



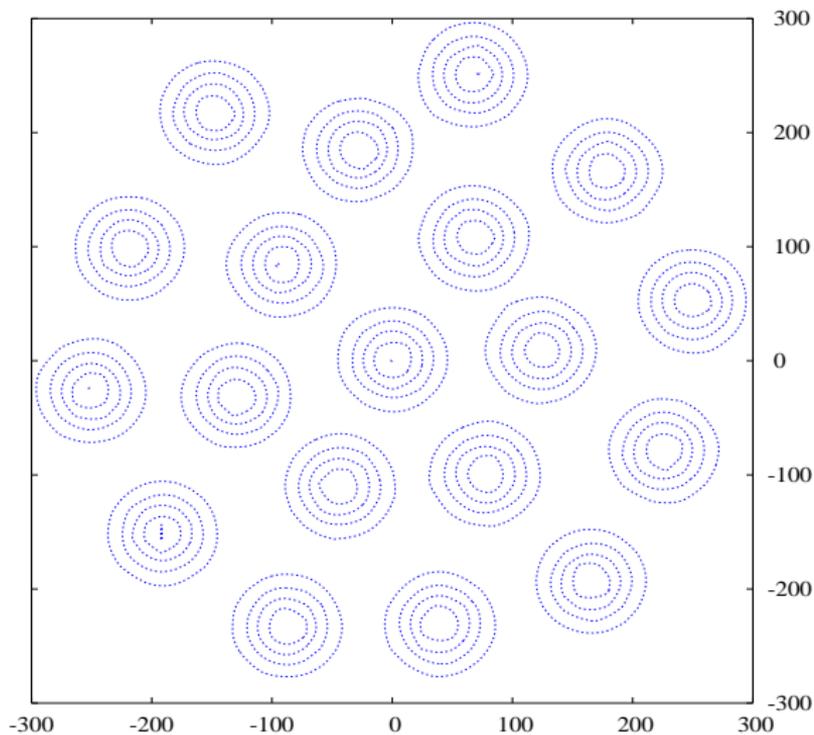
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 12



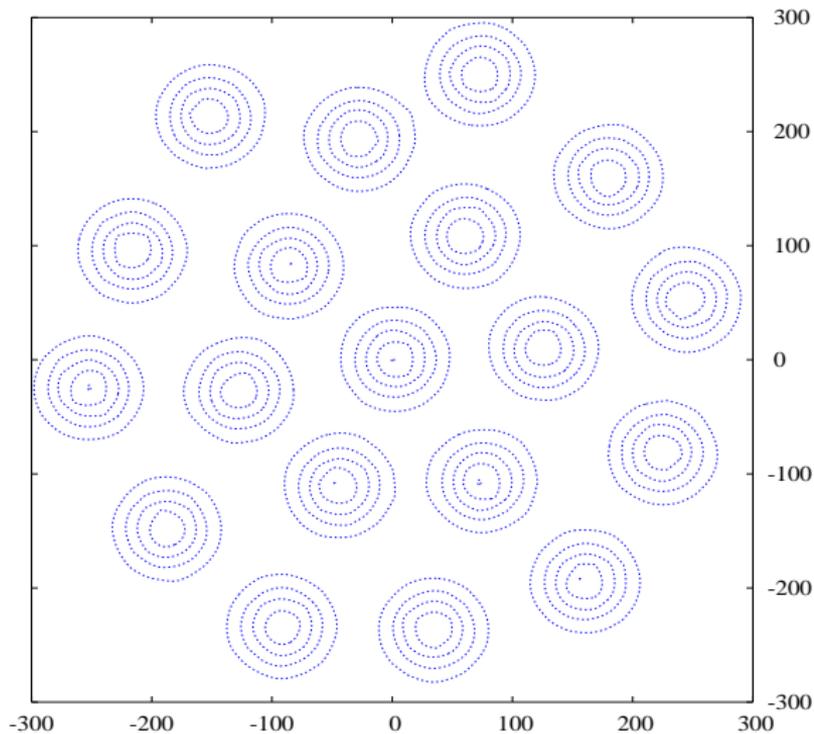
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 17



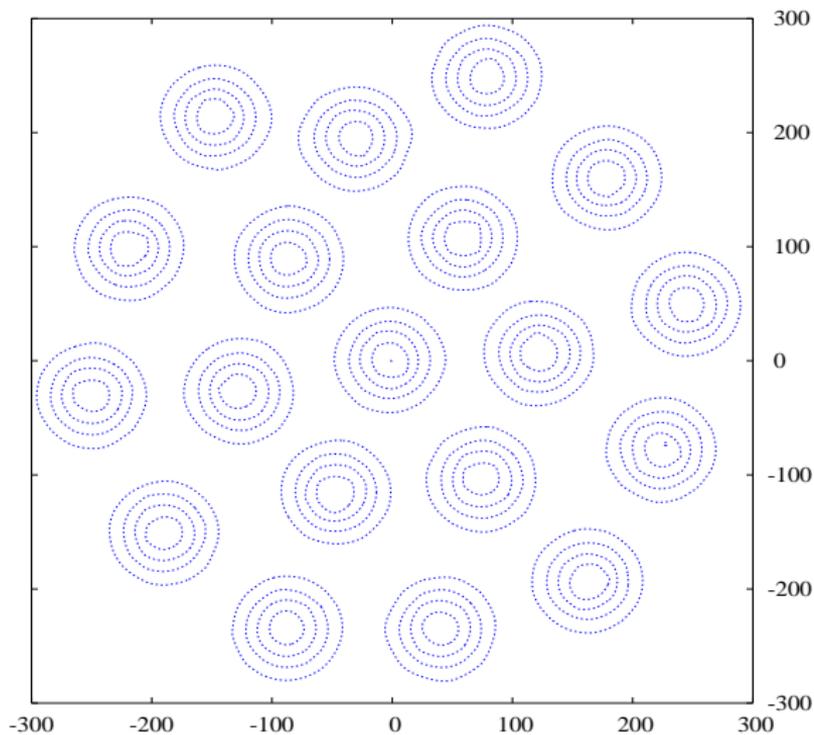
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 20



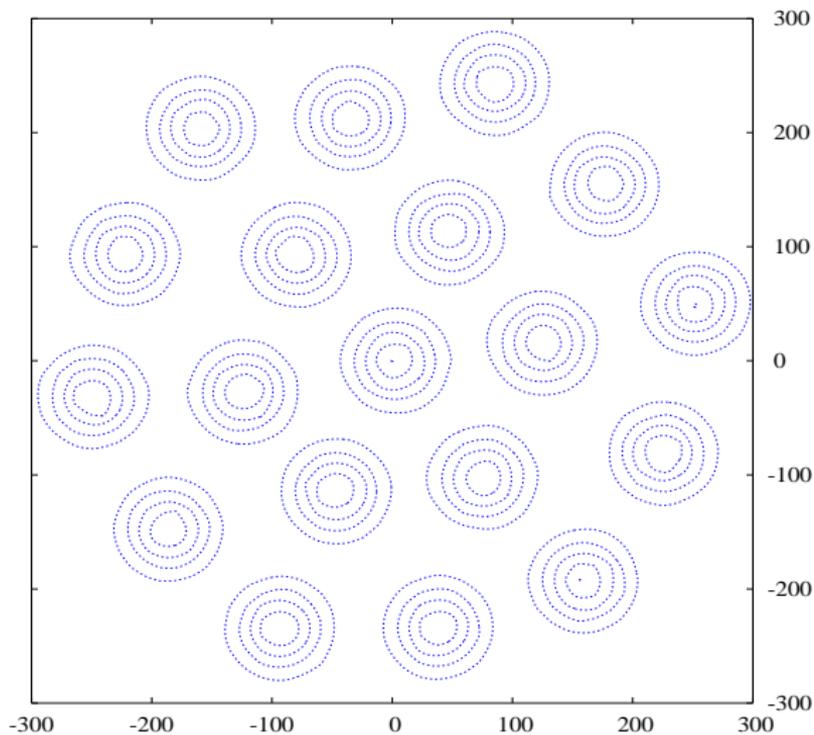
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 22



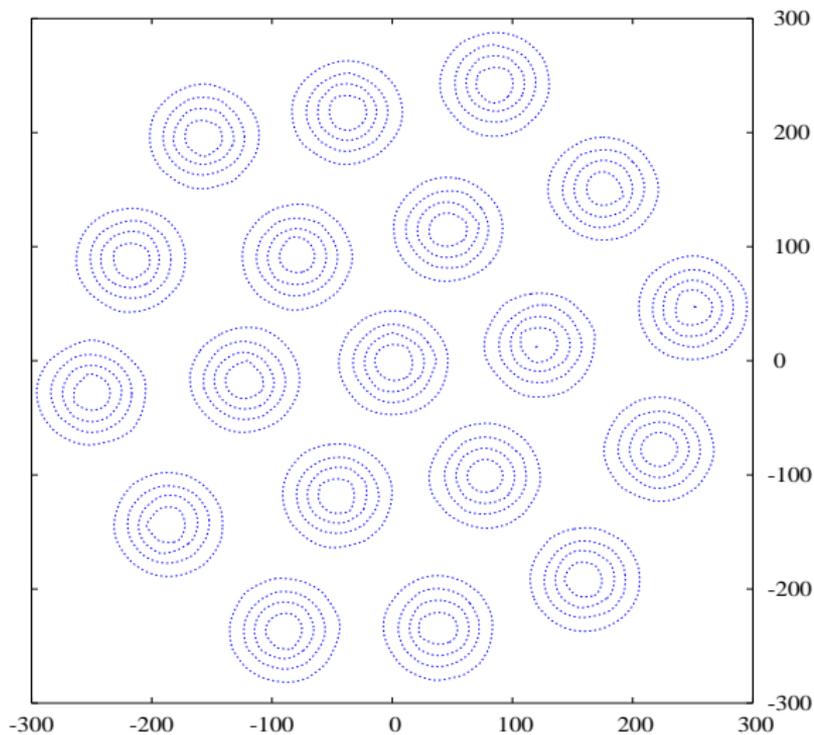
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 25



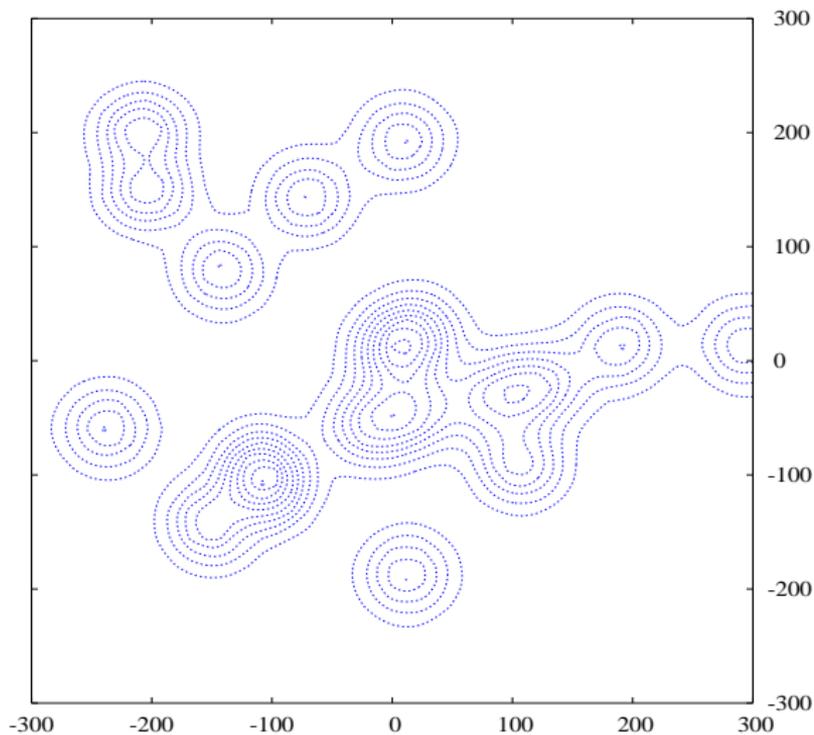
Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 27



Wigner crystallization in quantum dots (Guclu)

$N=19$, $S_z=0.5$, $r_s=45$, Step 0



Optimization of the Diffusion Monte Carlo energy (many-body nodes)

A linear combination of the DMC energy, VMC energy and variance can be optimized using extensions of either the **Newton** or the **linear** methods described above.

For the **linear** method, the ingredients are:

1. The calculation of excited state properties with quantum Monte Carlo
D.M. Ceperley and B. Bernu, *J. Chem. Phys.* **89**, 6316 (1988).

Use basis projected a time $T/2$,

$$H_{ij} = \langle \Psi_i | e^{-\hat{H}T/2} \hat{H} e^{-\hat{H}T/2} | \Psi_j \rangle = \langle \Psi_i | \hat{H} e^{-\hat{H}T} | \Psi_j \rangle$$

$$S_{ij} = \langle \Psi_i | e^{-\hat{H}T} | \Psi_j \rangle$$

2. A nonsymmetric H_{ij} ala **Nightingale**
3. Extension to nonlinear parameters as described above.
4. Integration over all projection times to a) improve statistics, b) mix in VMC energy and c) simplify the method

Conclusions

1. Three simple, robust and highly efficient methods for energy optimization.
Newton and linear methods work for all the parameters.
Perturbation theory is least expensive method for optimizing a large number of orbital parameters but slow for Jastrow.
2. Methods have been used on large systems – 1000 electrons in solid phases of Silicon.
3. All 3 methods have been used to optimize hundreds of parameters and perturbative method for a few thousand orbital parameters (Sorella, Filippi).
4. Optimization of the determinantal coefficients can be very important for reducing fixed-node error of E_{DMC} .
5. Seemingly similar molecules, e.g. C_2 and Si_2 have fixed-node errors for single-determinant wave functions that differ by a factor of 10.