Hartree-Fock Stability and its Relation to Strong Correlation

Evan Curtin

November 2, 2017

University of Illinois at Urbana-Champaign

Table of contents

Background Information

Correlation Energy

- The Molecular Electronic S.E. in the Born-Oppenheimer approximation

$$\left(\hat{T}_{elec} + \hat{V}_{elec-nuc} + \hat{V}_{\text{elec-elec}}\right)\Psi = E_{exact}\Psi$$

Correlation Energy

- The Molecular Electronic S.E. in the Born-Oppenheimer approximation

$$\left(\hat{T}_{elec} + \hat{V}_{elec-nuc} + \hat{V}_{\text{elec-elec}}\right)\Psi = E_{exact}\Psi$$

- in Hartree-Fock theory is replaced by

$$\left(\hat{T}_{elec} + \hat{V}_{elec-nuc} + \hat{V}_{\text{mean-field}}\right)\Psi = E_{HF}\Psi$$

Correlation Energy

- The Molecular Electronic S.E. in the Born-Oppenheimer approximation

$$\left(\hat{T}_{elec} + \hat{V}_{elec-nuc} + \hat{V}_{\text{elec-elec}}\right)\Psi = E_{exact}\Psi$$

- in Hartree-Fock theory is replaced by

$$\left(\hat{T}_{elec} + \hat{V}_{elec-nuc} + \hat{V}_{\text{mean-field}}\right)\Psi = E_{HF}\Psi$$

- the difference between the two energies is the correlation energy

$$E_{corr} = E_{exact} - E_{HF}$$

Correlation Energy

- The Molecular Electronic S.E. in the Born-Oppenheimer approximation

$$\left(\hat{T}_{elec} + \hat{V}_{elec-nuc} + \hat{V}_{\text{elec-elec}}\right)\Psi = E_{exact}\Psi$$

- in Hartree-Fock theory is replaced by

$$\left(\hat{T}_{elec} + \hat{V}_{elec-nuc} + \hat{V}_{\text{mean-field}}\right)\Psi = E_{HF}\Psi$$

- the difference between the two energies is the correlation energy

$$E_{corr} = E_{exact} - E_{HF}$$

- The correlation energy is recovered by post Hartree-Fock methods (CC, MBPT) which perform better the closer the HF solution is to the exact.

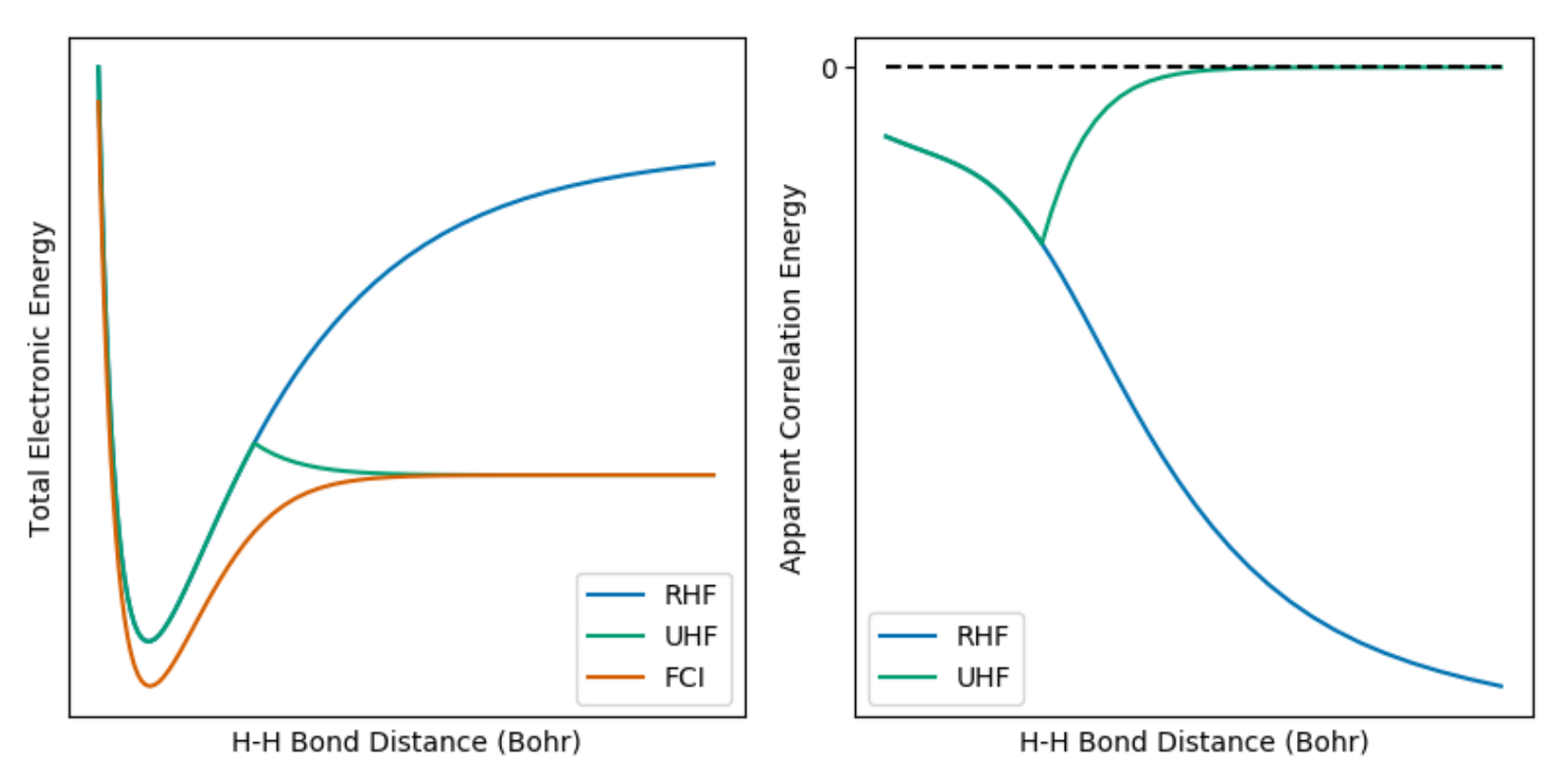
Hartree-Fock Theory is Almost Always Restricted

Method	Spinorbital	DoF	Eigenfunction of
Restricted	$\chi_j^{\alpha}(\vec{r},\sigma) = \sum_{i=1}^{N} c_{ij}\phi_i(\vec{r})\,\alpha(\sigma)$ $\chi_j^{\beta}(\vec{r},\sigma) = \sum_{i=1}^{N} c_{ij}\phi_i(\vec{r})\,\beta(\sigma)$	N/2	$\hat{S}^2$, $\hat{S}_z$
Unrestricted	$\chi_j^{\alpha}(\vec{r},\sigma) = \sum_{i=1}^{N} c_{ij}^{\alpha}\phi_i(\vec{r})\alpha(\sigma)$ $\chi_j^{\beta}(\vec{r},\sigma) = \sum_{i=1}^{N} c_{ij}^{\beta}\phi_i(\vec{r})\beta(\sigma)$	N	$\hat{S}_z$
General	$\chi_j(\vec{r},\sigma) = \sum_{i=1}^{N}[c_{ij}^{\alpha}\phi_i(\vec{r})\,\alpha(\sigma) + c_{ij}^{\beta}\phi_i(\vec{r})\beta(\sigma)]$	2N	Neither

Seeger, R.; Pople, J. A. J. Chem. Phys. 1977, 66 (7), 3045.

Case Study: H_2 Dissociation

Is this strong correlation or a bad reference?



Project Questions

- How can we improve our HF solutions?

Project Questions

- How can we improve our HF solutions?
- Does this actually help in finding the exact energy?

Project Questions

- How can we improve our HF solutions?
- Does this actually help in finding the exact energy?
- Apply to systems (known molecules, periodic systems)

Hartree-Fock Stability

Hartree-Fock Stability Conditions

- Solving the HF equations guarantees only that the energy is stationary. The first order variation is therefore 0.

Hartree-Fock Stability Conditions

- Solving the HF equations guarantees only that the energy is stationary. The first order variation is therefore 0.
- We need to know if this is indeed a minimum.

Hartree-Fock Stability Conditions

- Solving the HF equations guarantees only that the energy is stationary. The first order variation is therefore 0.
- We need to know if this is indeed a minimum.
- This not the case if any Orbital Hessian (aka stability matrix, electronic Hessian) eigenvalues are negative

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \omega \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \tag{1}$$

Hartree-Fock Stability Conditions

- Solving the HF equations guarantees only that the energy is stationary. The first order variation is therefore 0.
- We need to know if this is indeed a minimum.
- This not the case if any Orbital Hessian (aka stability matrix, electronic Hessian) eigenvalues are negative

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \omega \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \tag{1}$$

- Where

$$\begin{aligned} A_{ia,jb} &= \langle {}^{a}_{i} | H - E_0 | {}^{b}_{j} \rangle = (\epsilon_a - \epsilon_i)\, \delta_{ij} \delta_{ab} + \langle aj || ib \rangle \\ B_{ia,jb} &= \langle {}^{ab}_{ij} | H - E_0 | 0 \rangle = \langle ab || ij \rangle \,. \end{aligned} \tag{2}$$

Hartree-Fock Stability Conditions

The Matrix Equation Factorizes

Solution Type	Space Type					
	Real RHF	Complex RHF	Real UHF	Complex UHF	Real GHF	Complex GHF
Real RHF	${}^1\mathbf{A}' + {}^1\mathbf{B}'$	${}^1\mathbf{A}' - {}^1\mathbf{B}'$	${}^3\mathbf{A}' + {}^3\mathbf{B}'$	${}^3\mathbf{A}' - {}^3\mathbf{B}'$	${}^3\mathbf{A}' + {}^3\mathbf{B}'$	${}^3\mathbf{A}' - {}^3\mathbf{B}'$
Complex RHF	-	${}^1\mathbf{H}'$	-	${}^3\mathbf{H}'$	-	${}^3\mathbf{H}'$
Real UHF	-	-	$\mathbf{A}' + \mathbf{B}'$	$\mathbf{A}' - \mathbf{B}'$	$\mathbf{A}'' + \mathbf{B}''$	$\mathbf{A}'' - \mathbf{B}''$
Complex UHF	-	-	-	$\mathbf{H}'$	-	$\mathbf{H}'$
Real GHF	-	-	-	-	$\mathbf{A} - \mathbf{B}$	$\mathbf{A} - \mathbf{B}$
Complex GHF	-	-	-	-	-	$\mathbf{H}$

Table reproduced from Seeger & Pople[1]

1) Seeger, R.; Pople, J. A. J. Chem. Phys. 1977, 66 (7), 3045.

Homogeneous Electron Gas

Brief Overview

- Homogeneous Electron Gas (HEG) model, also known as Uniform Electron Gas or Jellium Model.
- Characterized by r_s, inversely proportional to density
- Electrons in a box with "smeared" nuclei $\rightarrow$ uniform positive background charge
- The total charge is constrained to be neutral,

$$V_{bg}(\mathbf{r}) = \sum_i \frac{-Ze^2}{|\mathbf{r} - \mathbf{R_i}|} \rightarrow -e^2 \int \frac{d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}, \tag{3}$$

- and the background and coulomb terms cancel exactly,

$$V_{ee} = e^2 \int \frac{d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}. \tag{4}$$

Brief Overview

- The discretized solutions are given by,

$$\epsilon_{\vec{k}} = \frac{\hbar^2 k^2}{2m} - \sum_{\vec{k}'}^{|\vec{k}'|<k_f} \langle \vec{k}, \vec{k}' | \vec{k}', \vec{k} \rangle \tag{5}$$

- Where the 2D and 3D two electron integral is given by

$$\langle \vec{k}, \vec{k}' | \vec{k}'', \vec{k}''' \rangle = \begin{cases} \frac{\pi}{V} \frac{2^{D-1}}{|\vec{k}-\vec{k}''|^{D-1}} & \vec{k}''' = \vec{k} + \vec{k}' - \vec{k}'' \\ 0 & \text{else} \end{cases}$$

- In 1D, if $V(r_{12}) = V_0 \delta(r_{12})$,

$$\langle k, k' | k'', k''' \rangle = \begin{cases} V_0 \ ; & k''' = k + k' - k'' \\ 0 \ ; & \text{else} \end{cases} \tag{6}$$

Giuliani, G.; Vignale, G. Quantum Theory of the Electron Liquid; 2005.

Exact Results

QMC Energies for various phases

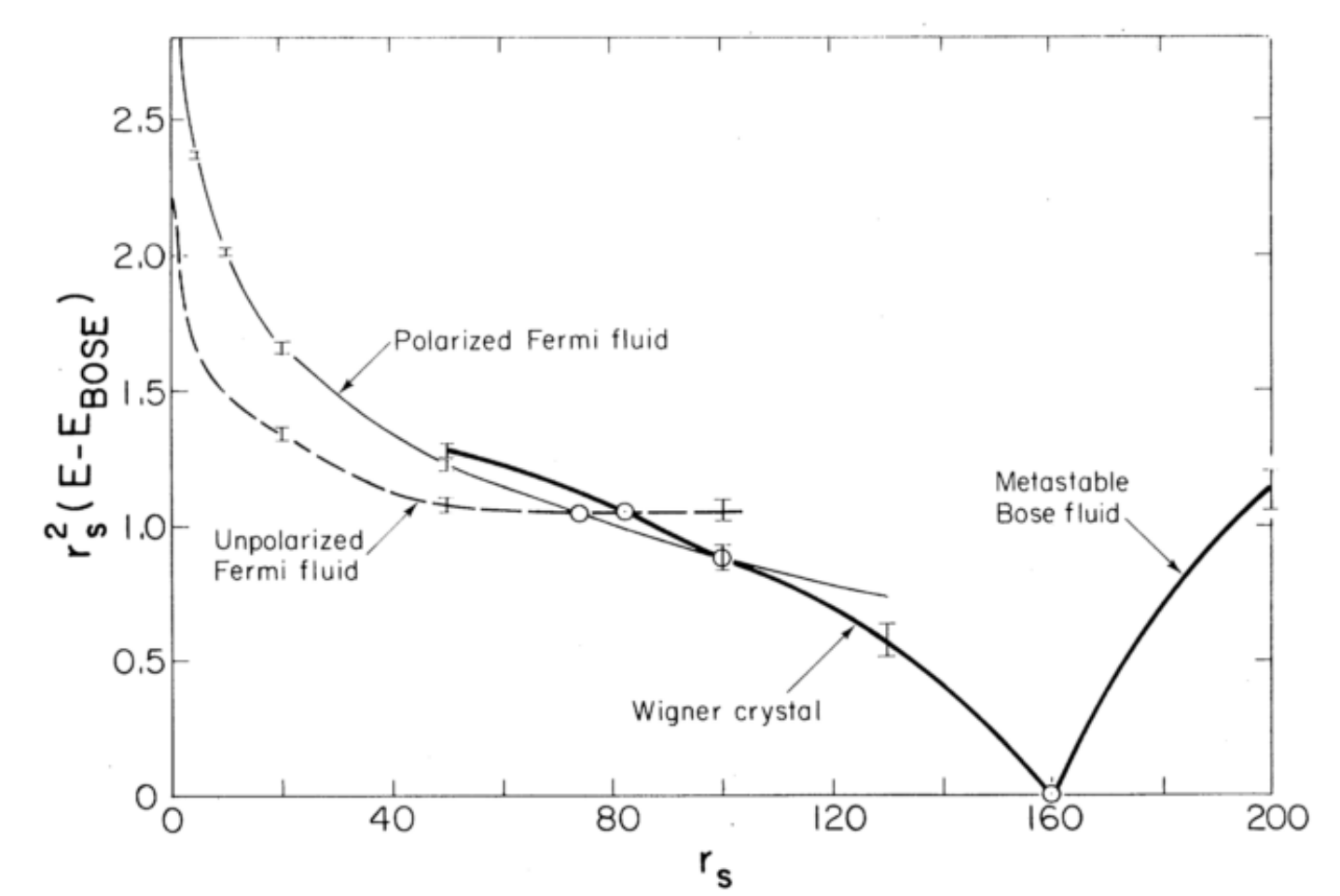


Figure reproduced from: Ceperley, D. M.; Alder, B. J. Phys. Rev. Lett. 1980, 45 (7), 566569.

Project Goal

- Can we use the HF-Stability analysis to determine where the HEG is unstable?

Project Goal

- Can we use the HF-Stability analysis to determine where the HEG is unstable?
- Will this predict the known tendency of crystallization at low density?

Project Goal

- Can we use the HF-Stability analysis to determine where the HEG is unstable?
- Will this predict the known tendency of crystallization at low density?
- It is known that the GHF solution persists at all densities for the HEG[1].

Overhauser, A. W. Phys. Rev. Lett. 1960, 4 (9), 462465.

Project Goal

- Can we use the HF-Stability analysis to determine where the HEG is unstable?
- Will this predict the known tendency of crystallization at low density?
- It is known that the GHF solution persists at all densities for the HEG[1].
- Can we show this numerically?

Overhauser, A. W. Phys. Rev. Lett. 1960, 4 (9), 462465.

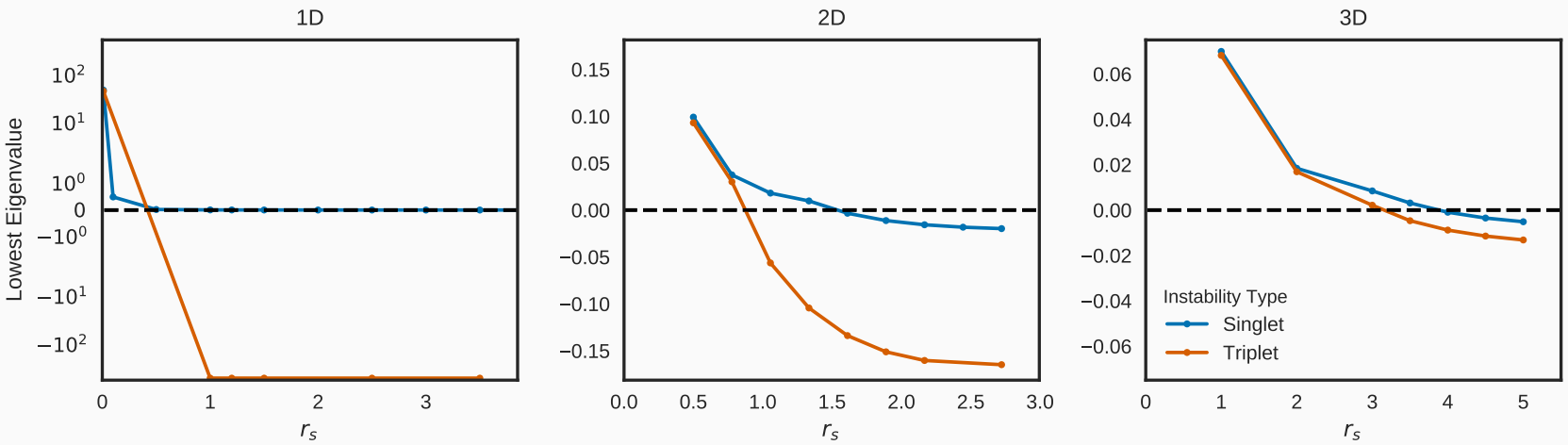
Project Goal

- Can we use the HF-Stability analysis to determine where the HEG is unstable?
- Will this predict the known tendency of crystallization at low density?
- It is known that the GHF solution persists at all densities for the HEG[1].
- Can we show this numerically?
- (Future) Can we find a lower energy GHF solution to reduce "correlation energy"?

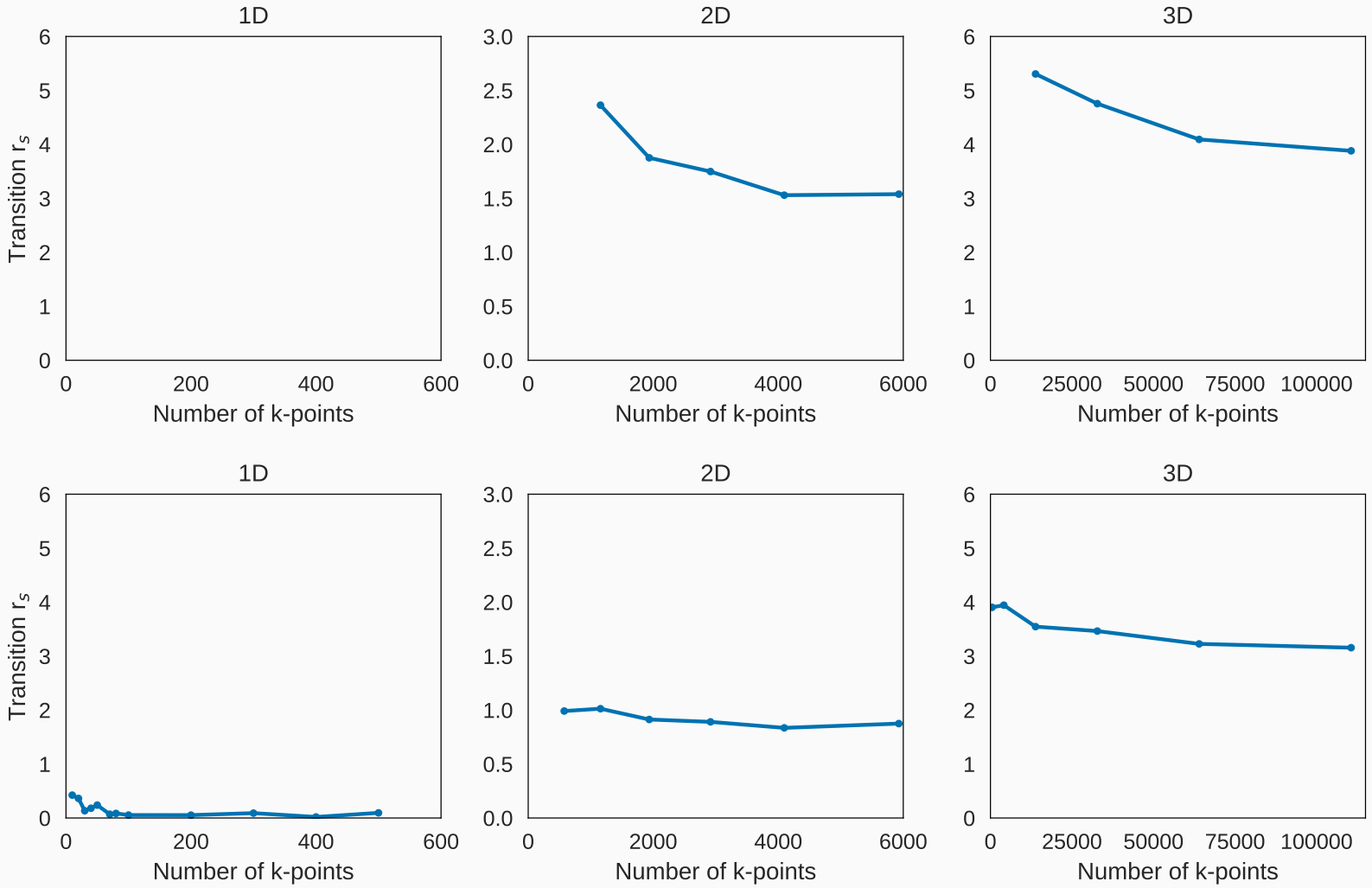
Overhauser, A. W. Phys. Rev. Lett. 1960, 4 (9), 462465.

Results

Stability Curves



Convergence of Stability Curves



Dependence on r_s

# Dimensions	Instability r_s	Literature
1	0.15	0^{1}
2	0.95	0.8^{2}
3	3.0	3.0^{3}

(1) Overhauser, A. W. Phys. Rev. 1962, 128 (3), 14371452.
(2) Bernu, B.; Delyon, F.; Holzmann, M.; Baguet, L. Phys. Rev. B 2011, 84 (11), 115115.
(3) Baguet, L.; Delyon, F.; Bernu, B.; Holzmann, M. Phys. Rev. B 2014, 90 (16), 165131.

Future Directions

Brute Force GHF

- Find the lowest energy GHF solution

(1) Ohno, K. Chem. Rec. 2016, 16 (5), 21982218.
(2) Supady, A.; Blum, V.; Baldauf, C. J. Chem. Inf. Model. 2015, 55 (11), 23382348.

Brute Force GHF

- Find the lowest energy GHF solution
- Use algorithms inspired by similar problems (atomic configuration)

(1) Ohno, K. Chem. Rec. 2016, 16 (5), 21982218.
(2) Supady, A.; Blum, V.; Baldauf, C. J. Chem. Inf. Model. 2015, 55 (11), 23382348.

Brute Force GHF

- Find the lowest energy GHF solution
- Use algorithms inspired by similar problems (atomic configuration)
- Global Reaction Route Mapping (GRRM) / Anharmonic Downard Distortion (ADD)[1]

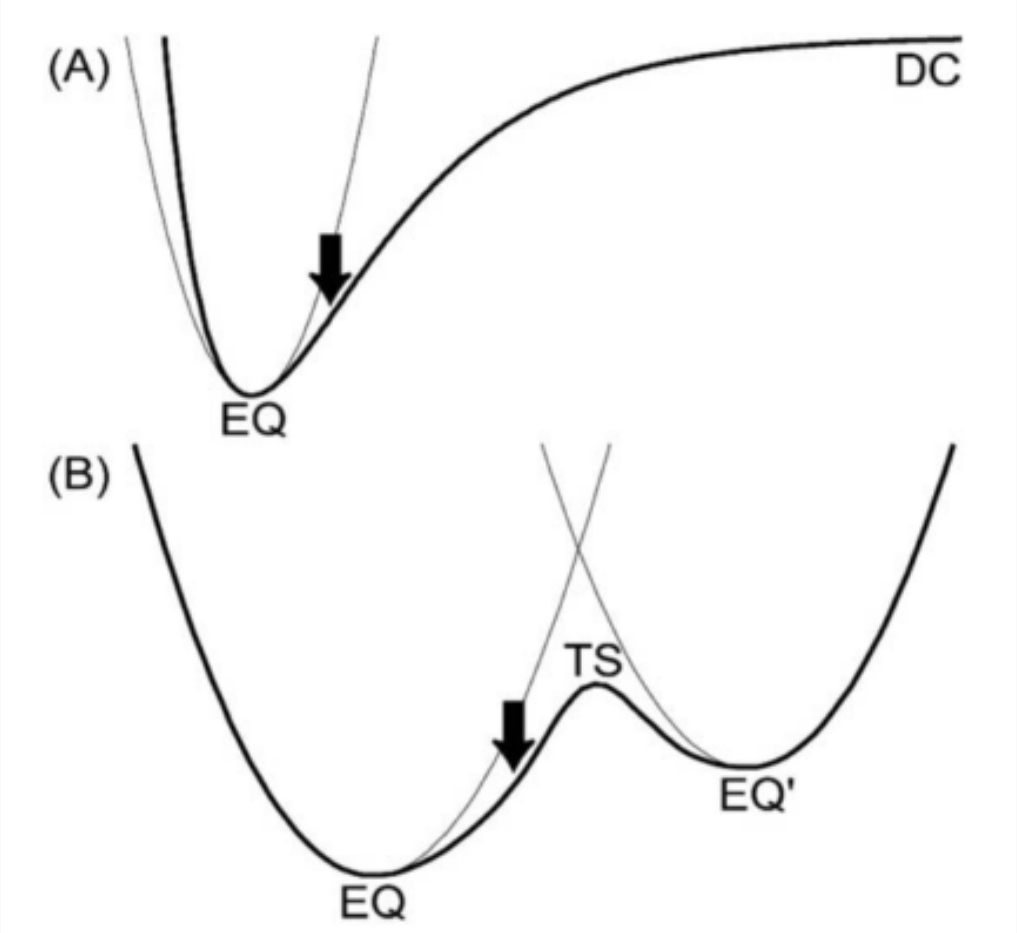
(1) Ohno, K. Chem. Rec. 2016, 16 (5), 21982218.
(2) Supady, A.; Blum, V.; Baldauf, C. J. Chem. Inf. Model. 2015, 55 (11), 23382348.

Brute Force GHF

- Find the lowest energy GHF solution
- Use algorithms inspired by similar problems (atomic configuration)
- Global Reaction Route Mapping (GRRM) / Anharmonic Downard Distortion (ADD)[1]
- Genetic Algorithm[2]

(1) Ohno, K. Chem. Rec. 2016, 16 (5), 21982218.
(2) Supady, A.; Blum, V.; Baldauf, C. J. Chem. Inf. Model. 2015, 55 (11), 23382348.

Anharmonic Downard Distortion



Replace atomic coordinates with MO coefficients and vibrational information with MO eigenvalues/vectors

Figure reproduced from: Ohno, K. Chem. Rec. 2016, 16 (5), 21982218.

Genetic Algorithm

Genetic Algorithm promising in finding many low energy solutions

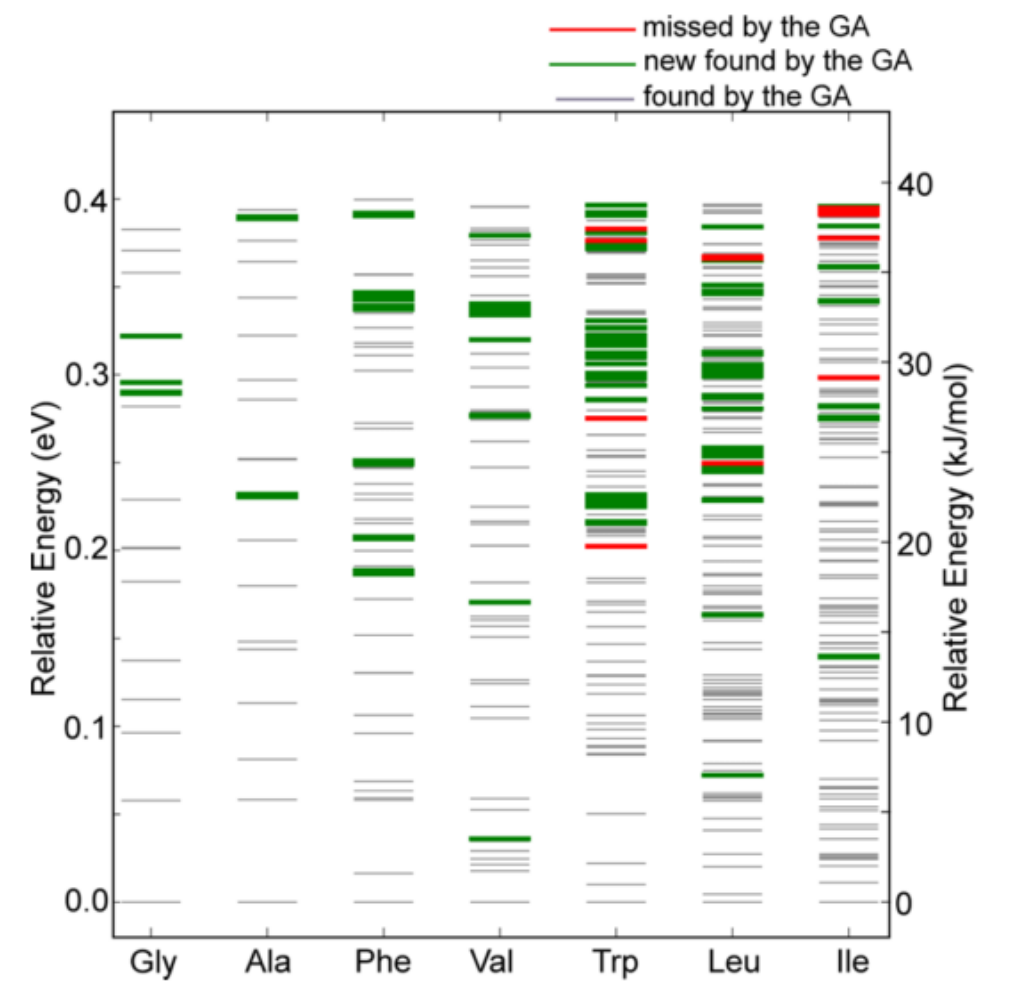


Figure reproduced from: Supady, A.; Blum, V.; Baldauf, C. J. Chem. Inf. Model. 2015, 55 (11), 23382348.

Proposal: Genetic Algorithm

- Use as the gene the density matrix / MO coefficients
- Less human input
- Less dependence on underlying PES structure
- How the HF solution gets found is irrelevant

Figure reproduced from: Supady, A.; Blum, V.; Baldauf, C. J. Chem. Inf. Model. 2015, 55 (11), 23382348.

Concluding Remarks

- Paramagnetic HEG instabilities are successfully reproduced by the MO hessian approach
- HF Stability analysis can show when restrictions should be lifted
- It is feasible that even stable HF are not global minimum
- Other approaches should be used to find even lower energy HF solutions
- Inspiration from similar problems could help; GRRM or Genetic Algorithms

Figure reproduced from: Supady, A.; Blum, V.; Baldauf, C. J. Chem. Inf. Model. 2015, 55 (11), 23382348.

Questions?

Iterative Subspace Eigenvalue Methods

Davidson's Algorithm

$\mathbf{Ax} = \lambda\mathbf{x}$	Eigenvalue Problem
$\mathbf{V} = [\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_M}]$	Guess vectors
$\tilde{\mathbf{A}} = \mathbf{V}^\dagger\mathbf{AV}$	Transform into subspace
$\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\lambda}\tilde{\mathbf{x}}$	Solve the subspace problem
$\mathbf{x_i} \approx \mathbf{x}_i^R = \mathbf{V}\tilde{\mathbf{x}}_i$	Approximate eigenvectors
$\lambda_i \approx \lambda_i^R = \tilde{\lambda}_i$	Approximate eigenvalues
$\mathbf{r}_i = (\mathbf{A} - \lambda_i\mathbf{I})\,\mathbf{x}_i^R$	Calculate the residue
$\delta_i = c_i\mathbf{r}_i$	Correction vectors
$c_i = \frac{1}{\lambda_i\mathbf{I}-\mathbf{D}}$	Diagonal Precondition
$\mathbf{V} = [\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_M}, \delta_1, \delta_2, ..., \delta_l]$	Append to guess and restart
$\mathbf{V} = \mathit{orthonormalized}(\mathbf{V})$	Ensure orthonormal projection

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Davidson, E. R. J. Comput. Phys. 1975, 17 (1), 8794.

Convergence Properties

- The convergence of these subspace algorithms depends on:

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested
  - The block size of the algorithm

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested
  - The block size of the algorithm
  - The density of the eigenvalue spectrum

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested
  - The block size of the algorithm
  - The density of the eigenvalue spectrum
  - The initial guess eigenvectors

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested
  - The block size of the algorithm
  - The density of the eigenvalue spectrum
  - The initial guess eigenvectors
  - The preconditioner (Davidson only)

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested
  - The block size of the algorithm
  - The density of the eigenvalue spectrum
  - The initial guess eigenvectors
  - The preconditioner (Davidson only)
- They can have convergence issues if:

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested
  - The block size of the algorithm
  - The density of the eigenvalue spectrum
  - The initial guess eigenvectors
  - The preconditioner (Davidson only)
- They can have convergence issues if:
  - The block size is too small compared to degeneracy

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested
  - The block size of the algorithm
  - The density of the eigenvalue spectrum
  - The initial guess eigenvectors
  - The preconditioner (Davidson only)
- They can have convergence issues if:
  - The block size is too small compared to degeneracy
  - The approximate eigenvectors become non-orthogonal

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Convergence Properties

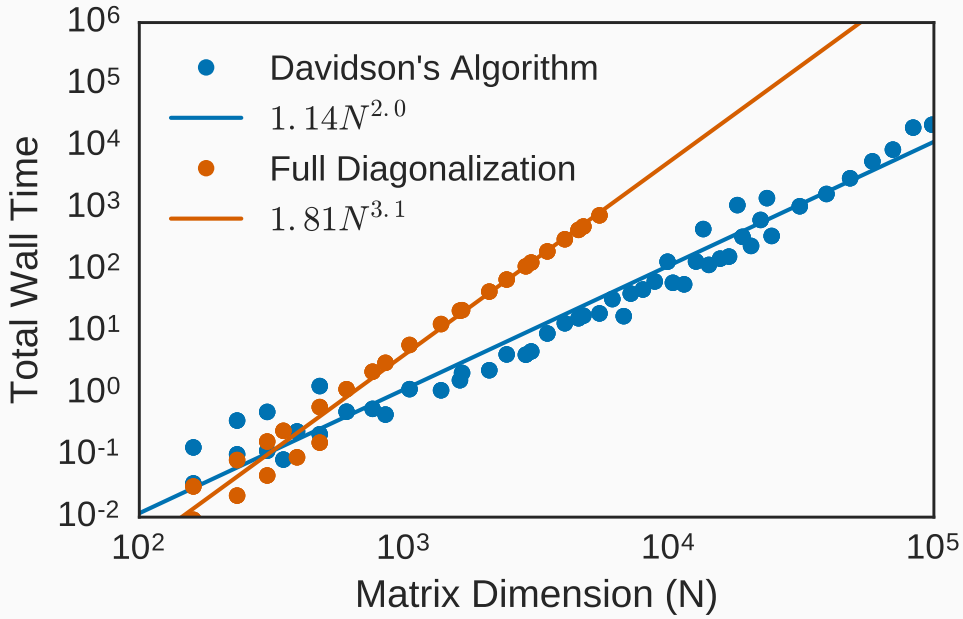
- The convergence of these subspace algorithms depends on:
  - Number of eigenvalues requested
  - The block size of the algorithm
  - The density of the eigenvalue spectrum
  - The initial guess eigenvectors
  - The preconditioner (Davidson only)
- They can have convergence issues if:
  - The block size is too small compared to degeneracy
  - The approximate eigenvectors become non-orthogonal
- My recommendation for guess eigenvectors is

$$v_j^{(i)} = normalize\left(\frac{1}{|A_{ii} - A_{jj}| + 1}\right). \tag{7}$$

1. Saad, Y. Numerical Methods for Large Eigenvalue Problems; SIAM, 2011.
2. Li, R.-C.; Zhang, L.-H. Convergence of Block Lanczos Method for Eigenvalue Clusters; 2013.

Davidson Scaling

- Davidson is Asymptotically quadratic.
- Full diagonalization is almost cubic.
- Matrix multiplication is order[1] $log_2(7) \approx 2.807$



1) http://mathworld.wolfram.com/StrassenFormulas.html

Orthogonalization

- The condition number, κ, is bound from below by

$$\kappa \geq \frac{Max(A_{ii})}{Min(A_{jj})} \tag{8}$$

- The Gram-Schmidt procedure has numerical issues,

$$||\mathbf{I} - \mathbf{Q}^{\mathbf{T}}\mathbf{Q}|| \leq \frac{\alpha\kappa^2}{1 - \beta\kappa^2}. \tag{9}$$

- Modified Gram-Schmidt is better, but not perfect,

$$||\mathbf{I} - \mathbf{Q}^{\mathbf{T}}\mathbf{Q}|| \leq \frac{\gamma\kappa}{1 - \eta\kappa} \tag{10}$$

- May need multiple orthogonalization steps

Hern, V.; Tom, A.; Vidal, V. SLEPc Technical Reports, 2007.