Fragment Molecular Orbital Molecular Dynamics

Kurt Brorsen

Iowa State University

CSGF Program Review

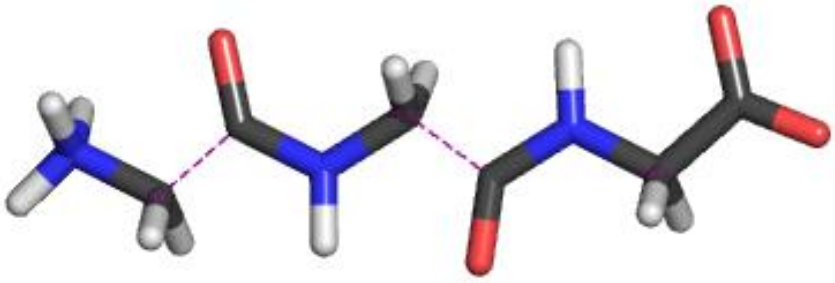
Computational Chemistry

- *Ab Initio* quantum chemistry
- Time Independent Schrodinger Equation

$$\hat{H}\psi = E\psi$$

- Hartree-Fock, Configuration Interaction, Perturbation Theory, Density Functional Theory, Coupled Cluster…

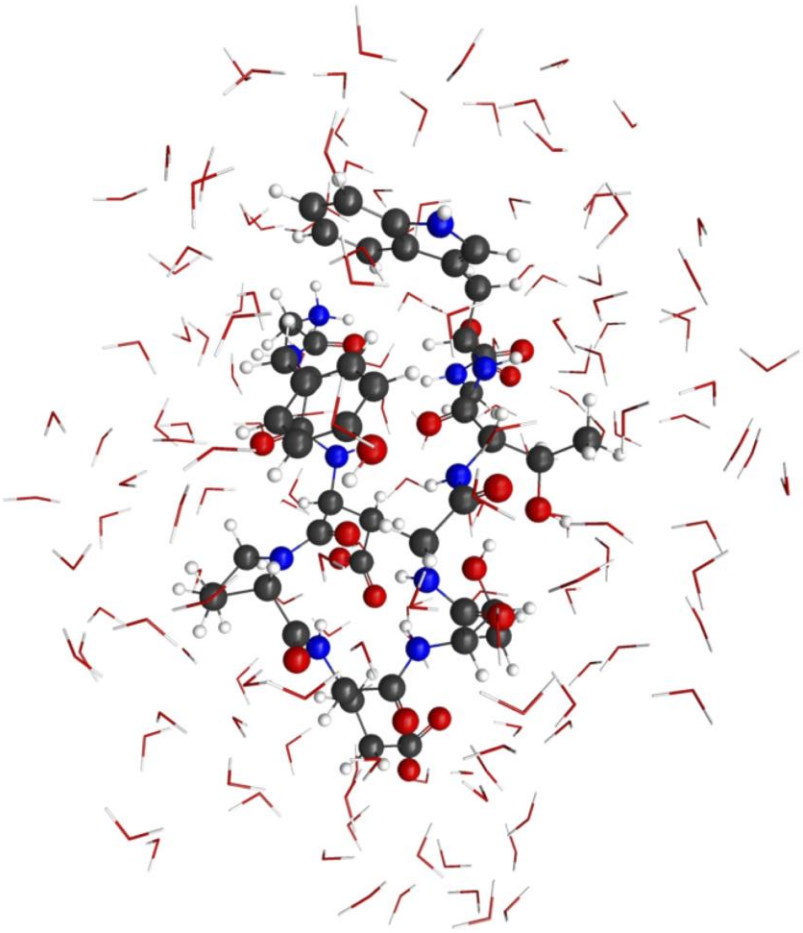
The Scaling Problem



- Basic implementations of *Ab Initio* methods exhibit poor scaling
- Hartree-Fock - N^4
- MP2 - N^5
- CCSD(T) - N^7

Fixing the Scaling Problem

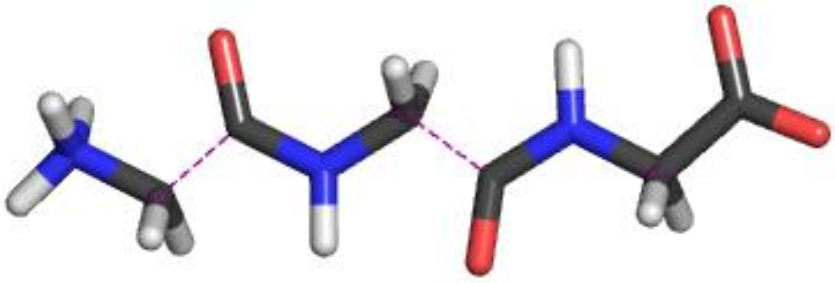
- Chemistry is largely local
- Local methods
- Fragment Based Methods
  - Divide and Conquer, Systematic Fragmentation Method, X-Pol, Fragment Molecular Orbital, etc…



Fragment Molecular Orbital (FMO)

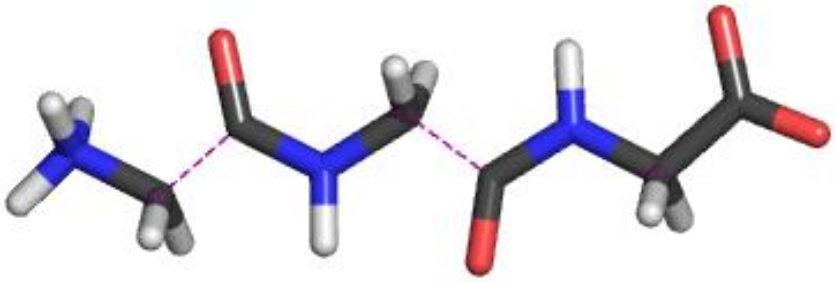
- Introduced in 1999 by Kitaura et al
- System is divided into fragments chosen by the user.
  - Trivial for clusters.
  - Chemical intuition must be used when dividing bonded systems.
- Each fragment (monomer) is iterated to self consistency in a Coulombic field due to the other fragments.
- A dimer calculation is carried out on pairs of fragments to obtain two body properties.
  - This is not iterated to self consistency

FMO Fragment Series



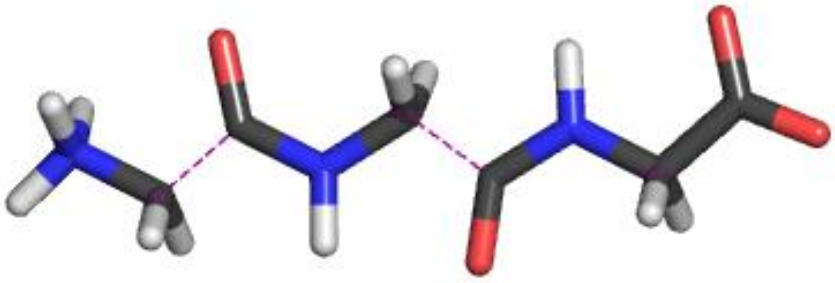
$$E = \sum_{I}^{N} E_I + \sum_{I>J}^{N} (E_{IJ} - E_I - E_J)$$
$$+ \sum_{I>J>K}^{N} \{(E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J)$$
$$- (E_{JK} - E_J - E_K) - (E_{KI} - E_K - E_I)\} + \cdots,$$

FMO Fragment Series



$$E = \boxed{\text{FMO1}\ \sum_{I}^{N} E_I} + \sum_{I>J}^{N} (E_{IJ} - E_I - E_J)$$
$$+ \sum_{I>J>K}^{N} \{(E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J)$$
$$- (E_{JK} - E_J - E_K) - (E_{KI} - E_K - E_I)\} + \cdots,$$

FMO Fragment Series



FMO1

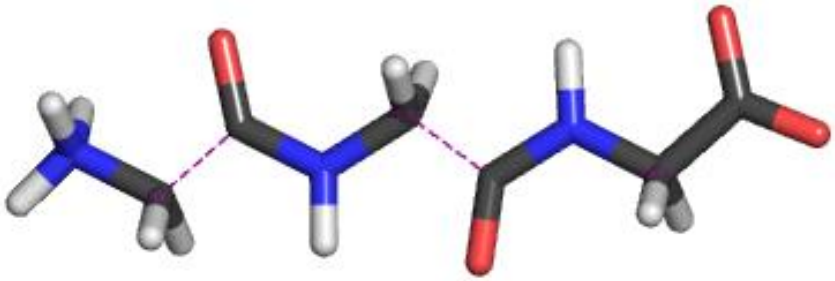
FMO2

$$E = \sum_{I}^{N} E_I + \sum_{I>J}^{N} (E_{IJ} - E_I - E_J)$$

$$+ \sum_{I>J>K}^{N} \{(E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J)$$

$$- (E_{JK} - E_J - E_K) - (E_{KI} - E_K - E_I)\} + \cdots,$$

FMO Fragment Series



FMO1

FMO2

FMO3

$$E = \sum_{I}^{N} E_I + \sum_{I>J}^{N} (E_{IJ} - E_I - E_J)$$

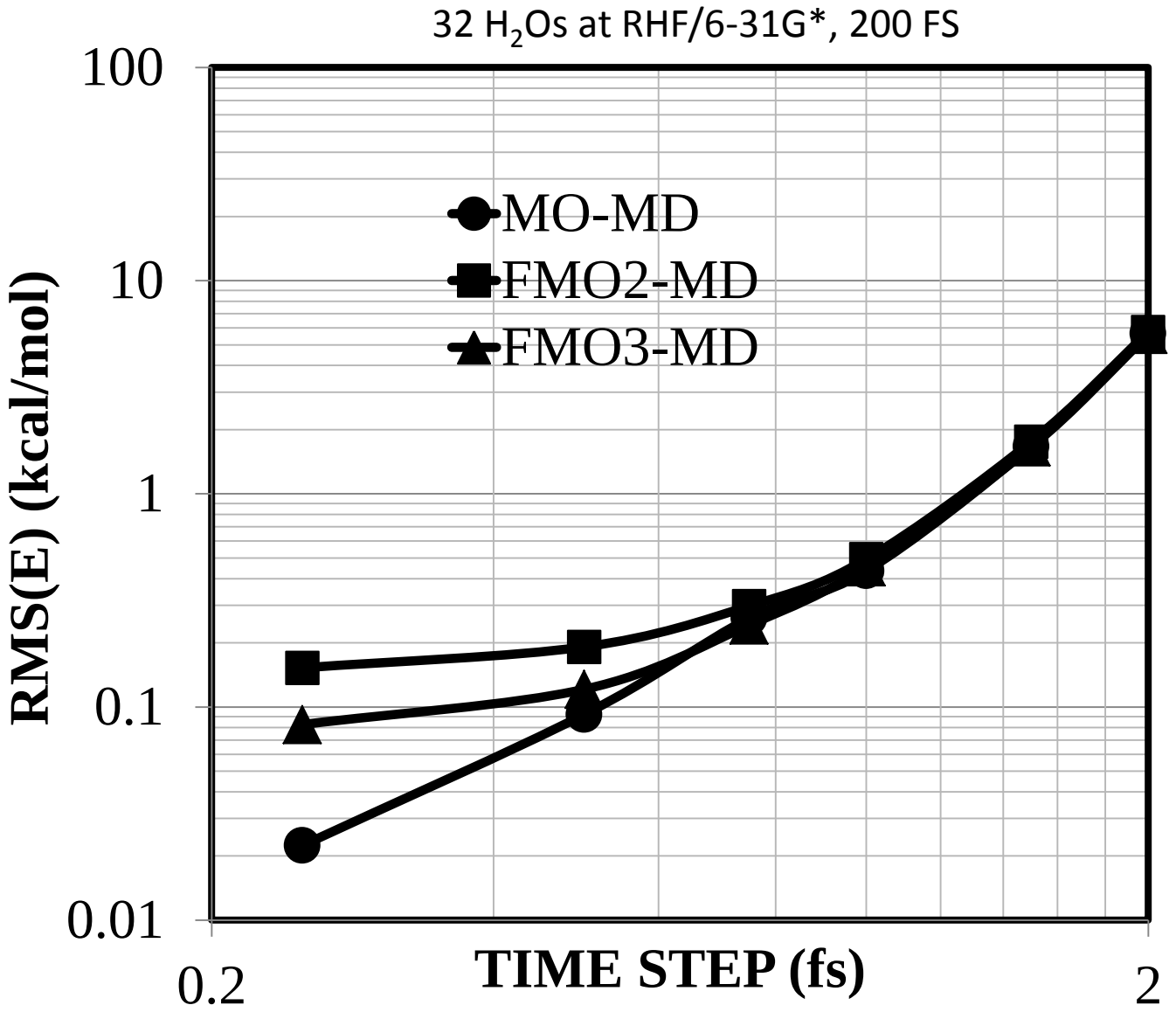
$$+ \sum_{I>J>K}^{N} \{(E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J)$$

$$- (E_{JK} - E_J - E_K) - (E_{KI} - E_K - E_I)\} + \cdots,$$

FMO Gradient

- The gradient for FMO was derived at the same time as the original FMO equations.
- Due to FMO not being fully variational, a fully analytic gradient required the solution of the Coupled Perturbed Hartree-Fock (CPHF) equations.
  - Due to difficulty in solving CPHF equations, their contribution to the gradient was ignored.
  - Gradient was termed “near fully analytic”
- Original gradient checked against numerical results
  - Sufficient for geometry optimizations
- FMO interfaced with molecular dynamics by Komeiji *et al* in 2003

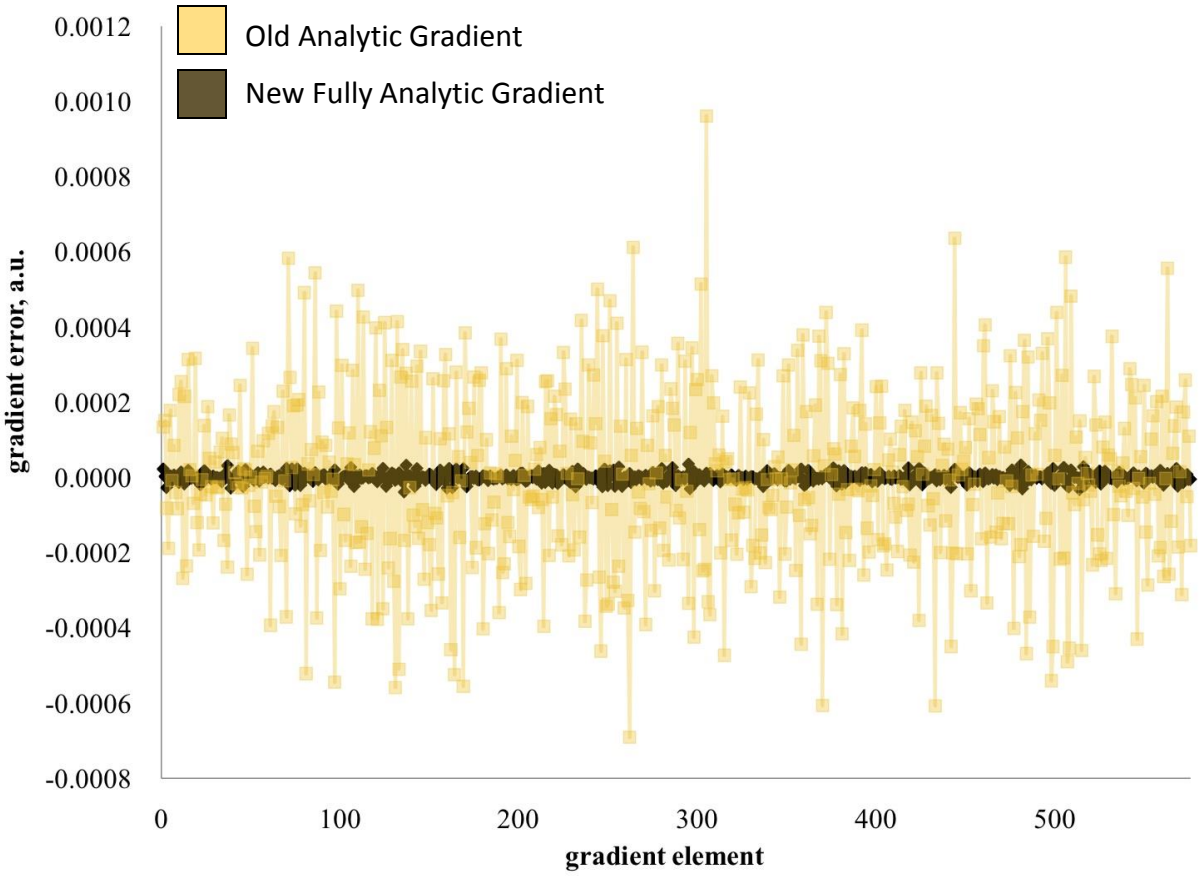
Energy Conservation for FMO-MD

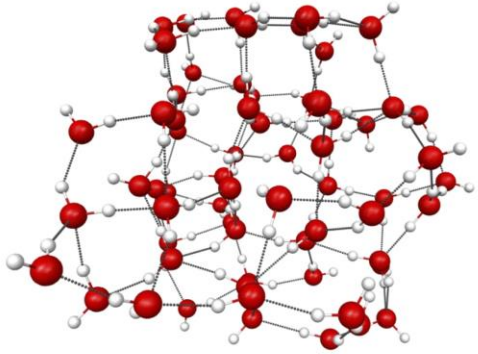


An analytic gradient for FMO

- Response terms were added in 2011
- Used a Z-vector approach based on the self-consistent charge calculation for monomers
- Further extended to MP2 and the use of an approximate electrostatic potential
  - Currently being extended to Density Functional Theory
- Over a factor of ten times more accurate

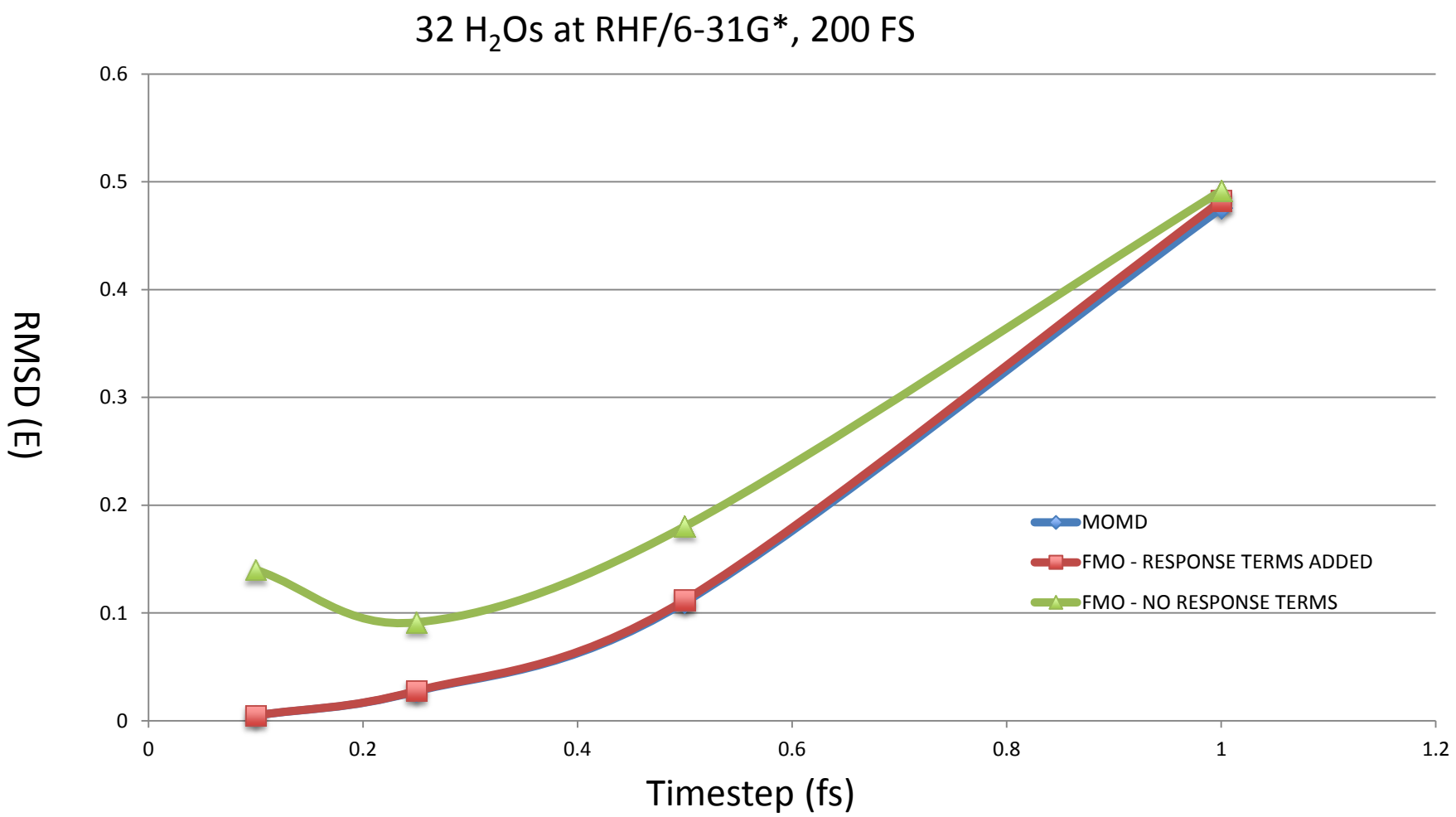
Fully Analytic Gradient for $(H_2O)_{64}$ RHF/6-31G(d)





	Old Analytic Gradient	New Fully Analytic Gradient
Max Gradient Error (a.u.)	0.000961	0.000035
RMS Gradient Error (a.u.)	0.000231	0.000011

FMO-MD and Energy Conservation



Conclusions/Acknowledgments

- FMO-MD requires a gradient that includes response terms
  - Still computationally expensive
- Iowa State University
  - Mark Gordon
- National Institute of Advanced Industrial Science and Technology Japan
  - Takeshi Nagata, Dmitri Federov
- Krell Institute

Alternative FMO Energy Formulation

The energy expression for FMO2 can be rewritten as

$$E = \sum_{I}^{N} E_I^{'} + \sum_{I>J}^{N} (E_{IJ}^{'} - E_I^{'} - E_J^{'}) + \sum_{I>J}^{N} Tr(\Delta \mathbf{D}^{IJ} \mathbf{V}^{IJ})$$

Where E_x is

$$E_X^{'} = \sum_{\mu\nu \in X} D_{\mu\nu}^X h_{\mu\nu}^X + \frac{1}{2} \sum_{\mu\nu\lambda\sigma \in X} \left[D_{\mu\nu}^X D_{\lambda\sigma}^X - \frac{1}{2} D_{\mu\lambda}^X D_{\nu\sigma}^X \right] (\mu\nu \mid \lambda\sigma) + \mathrm{E}_X^{NR},$$

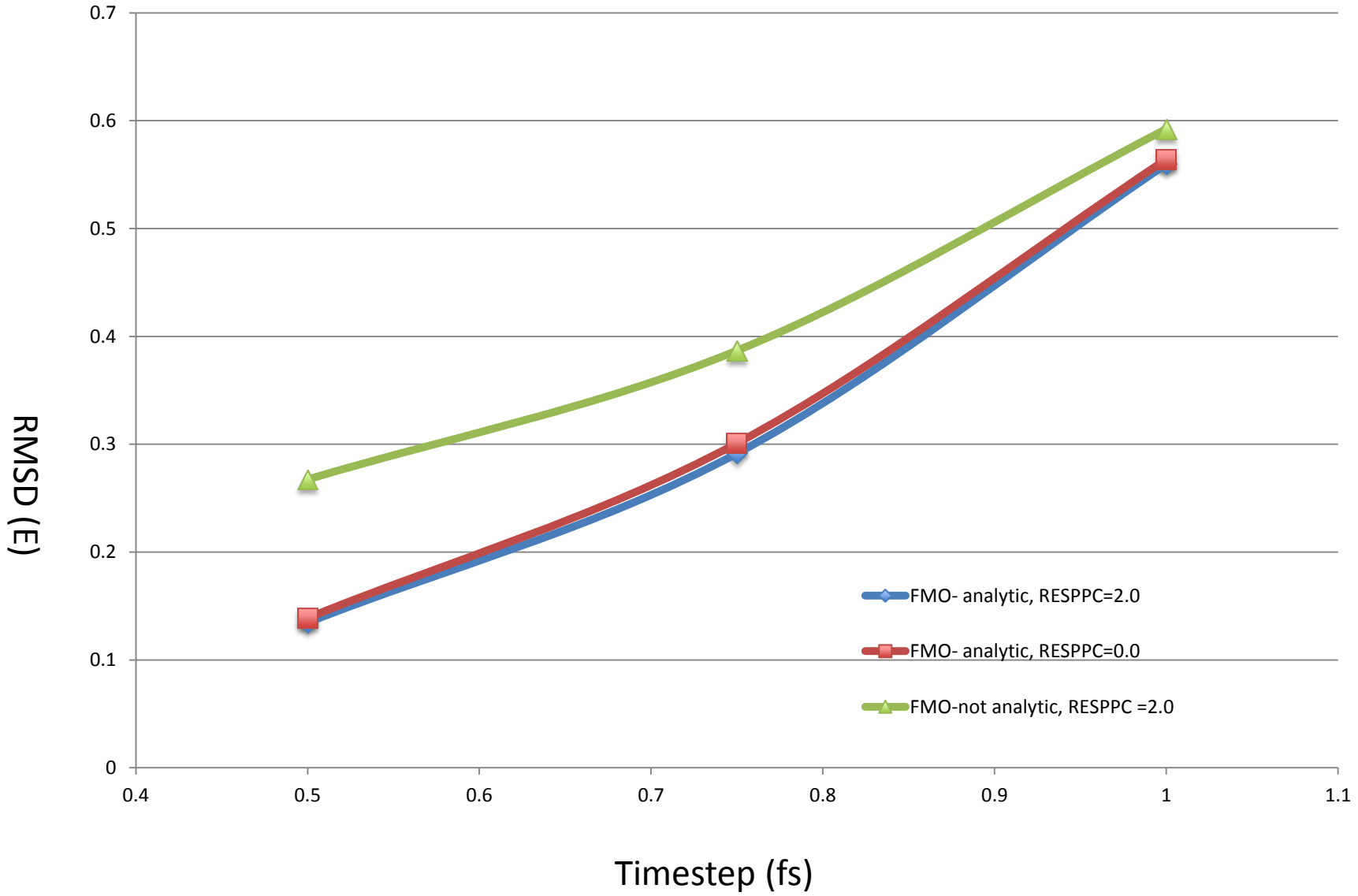
and

$$V_{\mu\nu}^{IJ} = \sum_{K \neq IJ} \left(u_{\mu\nu}^K + v_{\mu\nu}^K \right) \qquad \Delta \mathbf{D}^{IJ} = \mathbf{D}^{IJ} - (\mathbf{D}^I \oplus \mathbf{D}^J).$$

The contribution from the electrostatic potential contains both one and two electron terms

$$u_{\mu\nu}^K = \sum_{\mathrm{A} \in \mathrm{K}} \langle \mu | \frac{-\mathrm{Z_A}}{|r - R_\mathrm{A}|} | \nu \rangle, \qquad v_{\mu\nu}^K = \sum_{\lambda\sigma \notin \mathrm{K}} D_{\lambda\sigma}^{\mathrm{K}} (\mu\nu \mid \lambda\sigma),$$

FMO-RHF/3-21G* 17 H_2O, 200 fs



FMO Gradient Cont.

The differentiation of $E_X^{'}$ with respect to nuclear coordinate *a* leads to

$$\frac{\partial E_X^{'}}{\partial a} = \sum_{\mu\nu\in X} D_{\mu\nu}^X \frac{\partial h_{\mu\nu}^X}{\partial \alpha} + \frac{1}{2}\sum_{\mu\nu\lambda\sigma\in X}\left[D_{\mu\nu}^X D_{\sigma\lambda}^X - \frac{1}{2}D_{\mu\lambda}^X D_{\nu\sigma}^X\right]\frac{\partial(\mu\nu\,|\,\lambda\sigma)}{\partial\alpha}$$

$$-2\sum_{i,j\in X}^{occ} S_{ji}^{a,X} F_{ji}^{\prime X} - 4\sum_{i\in X}^{occ}\sum_{r\in X}^{vir} U_{ri}^{a,X} V_{ri}^X + \frac{\partial \mathrm{E}_X^{\mathrm{N}R}}{\partial\alpha},$$

where

$$F_{ij}^{'X} = h_{ij}^X + \sum_{k\in X}^{occ}\left[2(ij\,|\,kk) - (ik\,|\,jk)\right] \qquad S_{ij}^{a,X} = \sum_{\mu\nu\in X} C_{\mu i}^{X*}\frac{\partial S_{\mu\nu}^X}{\partial\alpha}C_{\nu j}^X,$$

and

$$\overline{U}^{a,X,Y} = 4\sum_{i\in X}^{occ}\sum_{r\in X}^{vir} U_{ri}^{a,X} V_{ri}^Y. \qquad \frac{\partial C_{\mu i}^X}{\partial a} = \sum_{m\in X}^{occ+vir} U_{ri}^{a,X} C_{\mu m}^X.$$

FMO Gradient Cont.

The differentiation of the ESP contribution with respect to nuclear coordinate *a* leads to

$$\frac{\partial}{\partial a} Tr(\Delta \mathbf{D}^{IJ} \mathbf{V}^{IJ}) = \sum_{\mu\nu \in IJ} \Delta D^{\mathrm{IJ}}_{\mu\nu} \sum_{\mathrm{K} \neq IJ}^{all} \left[\frac{\partial u^{\mathrm{K}}_{\mu\nu}}{\partial \alpha} + \sum_{\mu\nu \in \mathrm{K}} D^{\mathrm{K}}_{\lambda\sigma} \frac{\partial (\mu\nu \mid \lambda\sigma)}{\partial a} \right]$$

$$-2 \sum_{\mu\nu \in IJ} W^{IJ}_{\mu\nu} \frac{\partial S^{IJ}_{\mu\nu}}{\partial a} + 2 \sum_{\mu\nu \in I} W^{I}_{\mu\nu} \frac{\partial S^{I}_{\mu\nu}}{\partial a} + 2 \sum_{\mu\nu \in J} W^{J}_{\mu\nu} \frac{\partial S^{J}_{\mu\nu}}{\partial a} + \bar{U}^{a,IJ,IJ} - \bar{U}^{a,I,IJ} - \bar{U}^{a,J,IJ}$$

$$-2 \sum_{K \in IJ}^{all} \sum_{\mu\nu \in \mathrm{K}} \Delta X^{\mathrm{K(IJ)}}_{\mu\nu} S^{a,\mathrm{K}}_{\mu\nu} + 4 \sum_{\mathrm{K} \neq \mathrm{IJ}}^{all} \sum_{\mu\nu \in IJ} \sum_{r \in K}^{vir} \sum_{i \in K}^{occ} \Delta D^{\mathrm{IJ}}_{\mu\nu} U^{a,\mathrm{K}}_{ri} (\mu\nu \mid ri),$$

where

$$W^{X}_{\mu\nu} = \frac{1}{4} \sum_{\mu\nu \in X} D^{X}_{\mu\lambda} V^{\mathrm{IJ}}_{\lambda\sigma} D^{X}_{\sigma\nu}, \qquad \Delta X^{K(IJ)}_{\mu\nu} = \frac{1}{4} \sum_{\lambda\sigma \in \mathrm{K}} D^{\mathrm{K}}_{\mu\lambda} \left[\sum_{\zeta\eta \in IJ} \Delta D^{IJ}_{\zeta\eta} (\zeta\eta \mid \lambda\sigma) \right] D^{\mathrm{K}}_{\sigma\nu}.$$

FMO Gradient Cont.

The collection of all the $\bar{U}^a$ terms forms the following equation

$$\bar{U}^a = -\sum_{I}^{N} \bar{U}^{a,I,I} - \sum_{I>J}^{N} \left(\bar{U}^{a,IJ,IJ} - \bar{U}^{a,I,I} - \bar{U}^{a,J,J} \right) + \sum_{I>J}^{N} \left(\bar{U}^{a,IJ,IJ} - \bar{U}^{a,I,IJ} - \bar{U}^{a,J,IJ} \right)$$

$\bar{U}^a$ is zero when no ESP approximations are applied or all the ESPs are approximated uniformly.

In the case of $\bar{U}^a = 0$ only the last term in the ESP derivative includes the values $U_{ri}^{a,K}$ and one must solve the Coupled Perturbed Hartree Fock equations to find them

Coupled Perturbed Hartree Fock-FMO

For monomer *I* the corresponding Fock matrix rewritten in terms of MOs is

$$F_{ij}^{I} = F_{ij}^{'I} + V_{ij}^{I}$$
$$= \tilde{h}_{ij}^{I} + \sum_{k \in \mathrm{I}}^{occ} \left[2(ij \mid kk) - (ik \mid jk) \right]$$

Where the one electron Hamiltonian in FMO is

$$\tilde{h}_{ij}^{I} = h_{ij}^{I} + V_{ij}^{I}.$$

Differentiating the Fock matrix

$$\frac{\partial F_{ij}^{I}}{\partial a} = \frac{\partial}{\partial a}\left(\tilde{h}_{ij}^{I} + \sum_{k \in \mathrm{I}}^{occ} \left[2(ij \mid kk) - (ik \mid jk) \right] \right)$$

CPHF-FMO

After some algebra this leads to

$$\frac{\partial F_{ij}^{I}}{\partial a} = F_{ij}^{a,I} + \sum_{k\in \mathrm{I}}^{occ+vir} \left(U_{ki}^{a,\mathrm{I}} F_{kj}^{\mathrm{I}} + U_{kj}^{a,\mathrm{I}} F_{ik}^{\mathrm{I}} \right) + \sum_{k\in \mathrm{I}}^{occ+vir} \sum_{l\in \mathrm{I}}^{occ} U_{kl}^{a,\mathrm{I}} \mathrm{A}'_{ij,kl} + \sum_{K\neq I} \sum_{k\in \mathrm{K}}^{occ+vir} \sum_{l\in \mathrm{K}}^{occ} U_{kl}^{a,\mathrm{K}} 4(ij \mid kl),$$

Where the orbital Hessian contribution and Fock derivative is respectively

$$A'_{ij,kl} = 4(ij \mid kl) - (ik \mid jl) - (il \mid jk), \qquad F_{ij}^{a,I} = h_{ij}^{a,I} + V_{ij}^{a,I} + \sum_{k\in \mathrm{I}}^{occ} \left[2(ij \mid kk)^{\alpha} - (ik \mid jk)^{\alpha} \right]$$

The derivative of the one electron Hamiltonian and the two electron integral are

$$h_{ij}^{a,I} = \sum_{\mu\nu\in I} C_{\mu i}^{\mathrm{I}*} C_{\nu j}^{\mathrm{I}} \frac{\partial h_{\mu\nu}^{\mathrm{I}}}{\partial \alpha}, \qquad (ij \mid kl)^{a} = \sum_{\mu\nu\rho\sigma} C_{\mu i}^{\mathrm{I}*} C_{\nu j}^{\mathrm{I}} C_{\rho k}^{\mathrm{I}*} C_{\sigma l}^{\mathrm{I}} \frac{\partial (\mu\nu \mid \rho\sigma)}{\partial \alpha},$$

The ESP derivative is

$$V_{ij}^{a,I} = \sum_{K\_I} \left(u_{ij}^{a,K} + \sum_{k\in \mathrm{K}}^{occ} 2(ij \mid kk)^{\alpha} \right).$$

CPHF-FMO

More algebra lead to

$$\frac{\partial F_{ij}^{I}}{\partial a} = F_{ij}^{a,I} - \left(\varepsilon_j^I - \varepsilon_i^I\right)U_{ij}^{a,I} - S_{ij}^{a,I}\varepsilon_j^I - \sum_{k,l\in \mathrm{I}}^{occ} S_{kl}^{a,\mathrm{I}}\left[2(ij\,|\,kl) - (ik\,|\,jk)\right] + \sum_{k\in \mathrm{I}}^{vir}\sum_{l\in \mathrm{I}}^{occ} U^{a,I}A'_{ij,kl}$$
$$- \sum_{K\neq I}\sum_{k,l\in \mathrm{K}}^{occ} 2S_{kl}^{a,\mathrm{K}}(ij\,|\,kl) + \sum_{\mathrm{K}\neq \mathrm{I}}\sum_{k\in \mathrm{K}}^{vir}\sum_{l\in \mathrm{K}}^{occ} U_{kl}^{\alpha,\mathrm{K}}4(ij\,|\,kl),$$

Where the following relationship is used

$$S_{ij}^{a,X} + U_{ji}^{a,X} + U_{ij}^{a,X} = 0.$$

And ε_i^I is the orbital energy of MO *i* on fragment *I*

CPHF-FMO

We can rewrite the CPHF-FMO equations in matrix for as

$$\mathbf{A}\mathbf{U}^a = \mathbf{B}_0^a,$$

Where the diagonal and off diagonal blocks of matrix **A** are respectively

$$A_{ij,kl}^{I,I} = \delta_{ik}\delta_{jl}\left(\varepsilon_j^I - \varepsilon_i^I\right) - \left[4(ij\,|\,kl) - (ik\,|\,jl) - (il\,|\,jk)\right], \qquad A_{ij,kl}^{I,K} = -4(ij\,|\,kl),$$

And the IJ element of vector $\mathbf{B}_0^a$ is

$$B_{0,ij}^{a,I} = F_{ij}^{a,I} - S_{ij}^{a,I}\varepsilon_j^I - \sum_{kl\in \mathrm{I}}^{occ} S_{kl}^{a,\mathrm{I}}\left[2(ij\,|\,kl) - (ik\,|\,jl)\right] - \sum_{\mathrm{K}\neq \mathrm{I}}\sum_{kl\in \mathrm{K}}^{occ} 2S_{kl}^{a,\mathrm{K}}(ij\,|\,kl).$$

Z-Vector Method in FMO

It is still impractical to solve the CPHF equations directly, so the Z-Vector method is applied.

Collecting $U_{ri}^{a,K}$ terms for all dimer fragments IJ leads to a gradient contribution of

$$\Re = 4\sum_{I>J}\sum_{K\neq IJ}^{all}\sum_{\mu\nu\in IJ}\sum_{r\in K}^{vir}\sum_{i\in K}^{occ}\Delta D_{\mu\nu}^{IJ}U_{ri}^{a,K}(\mu\nu|ri) = \sum_{K}^{all}\sum_{r\in \mathrm{K}}^{vir}\sum_{i\in \mathrm{K}}^{occ}U_{ri}^{\alpha,\mathrm{K}}X_{ri}^{\mathrm{K}},$$

where $X_{ri}^{K} = 4\sum_{(I>J)\neq K}\sum_{\mu\nu\in IJ}\Delta D_{\mu\nu}^{\mathrm{IJ}}(\mu\nu|ri).$

$\Re$ can be expressed in vector form as

$$\Re = \sum_{K}^{all}\sum_{r\in \mathrm{K}}^{vir}\sum_{i\in \mathrm{K}}^{occ}U_{ri}^{a,\mathrm{K}}X_{ri}^{\mathrm{K}} = X^{\mathrm{T}}U^{\alpha} \qquad = \mathbf{X}^{T}\mathbf{A}^{-1}\mathbf{B}_{0}^{a} \qquad = \mathbf{Z}^{T}\mathbf{B}_{0}^{a}.$$

This reduces the CPHF problem to a set of simultaneous equations

$$\mathbf{A}^{T}\mathbf{Z} = \mathbf{X}.$$

Implementation of CPHF-FMO

The Z-Vector method is still time consuming. Taking advantage of the local nature of FMO and separating the diagonal and off-diagonal blocks of the previous equation we can rewrite the previous matrix equation as

$$\sum_{k\in \mathrm{I}}^{vir}\sum_{l\in \mathrm{I}}^{occ} \mathrm{A}^{\mathrm{I,I}}_{kl,ri}\mathrm{Z}^{\mathrm{I}}_{kl} = X^{\mathrm{I}}_{ri} - \sum_{\mathrm{K}\neq \mathrm{I}}^{all}\sum_{k\in \mathrm{K}}^{vir}\sum_{l\in \mathrm{K}}^{occ} \mathrm{A}^{\mathrm{K,I}}_{kl,ri}\mathrm{Z}^{\mathrm{K}}_{kl},$$

in matrix form

$$\left(\mathbf{A}^{I,I}\right)^{T}\mathbf{Z}^{I} = \mathbf{X}'^{I},$$

where

$$\mathbf{X}'^{I} = \mathbf{X}^{I} - \sum_{K\neq I}^{all}\left(\mathbf{A}^{K,I}\right)^{T}\mathbf{Z}^{K}.$$

The solution is then accomplished iteratively and is termed the self-consistent Z-vector (SCZV) method due the similarity of it to the self-consistent charge procedure

Self-Consistent Z-Vector (SCZV) method

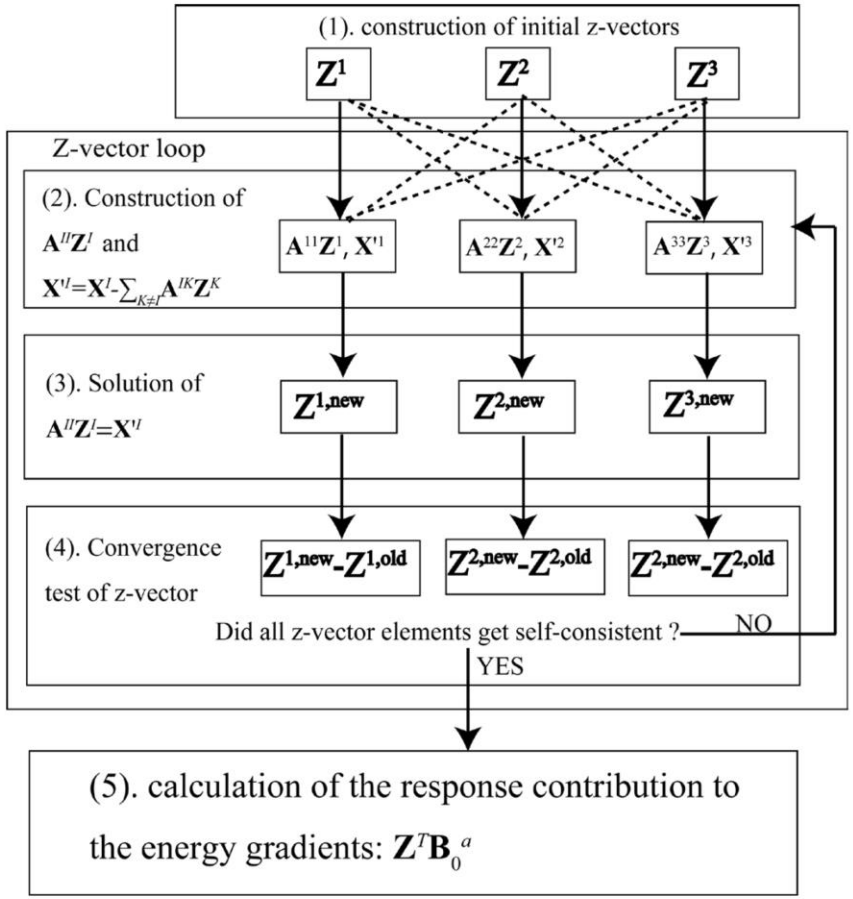


Figure 1