Energy Derivatives in Quantum Monte Carlo

Raymond Clay

University of Illinois at Urbana-Champaign
rcclay2@illinois.edu

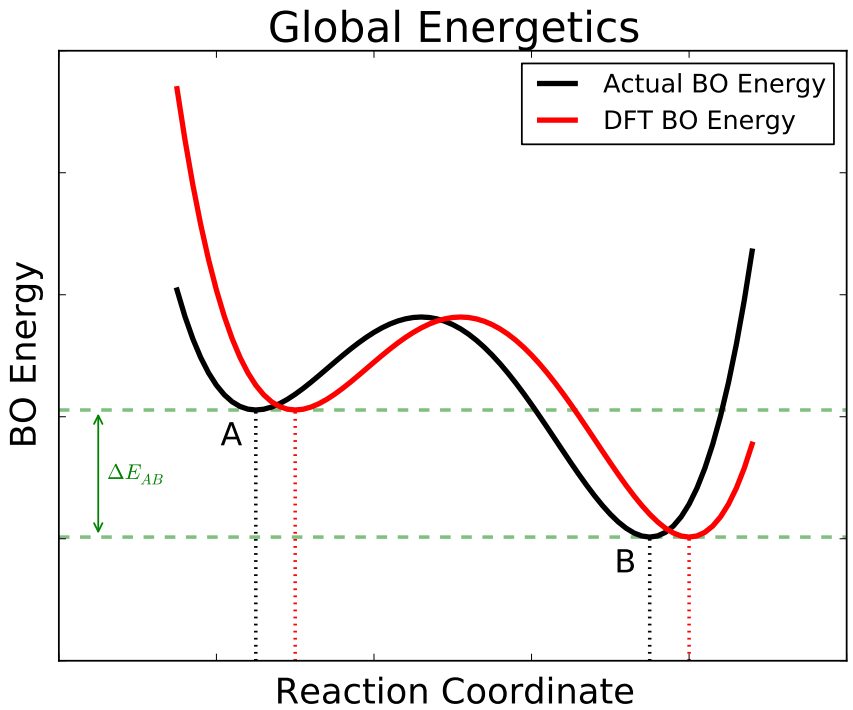
July 17, 2014

Importance of Energy Derivatives

Many important equilibrium properties are derivatives of the (free) energy.

- Forces, pressure, magnetization, polarization
- Bulk modulus, compressibility, elastic constants

Importance of Energy Derivatives



Energy derivatives help us:

- **Find local minima in the BO energy surface**
  - Structural optimization
  - Structure searching
- **Describe the shape of the local minima.**
  - Phonons
  - Elastic properties

Routine DFT Applications

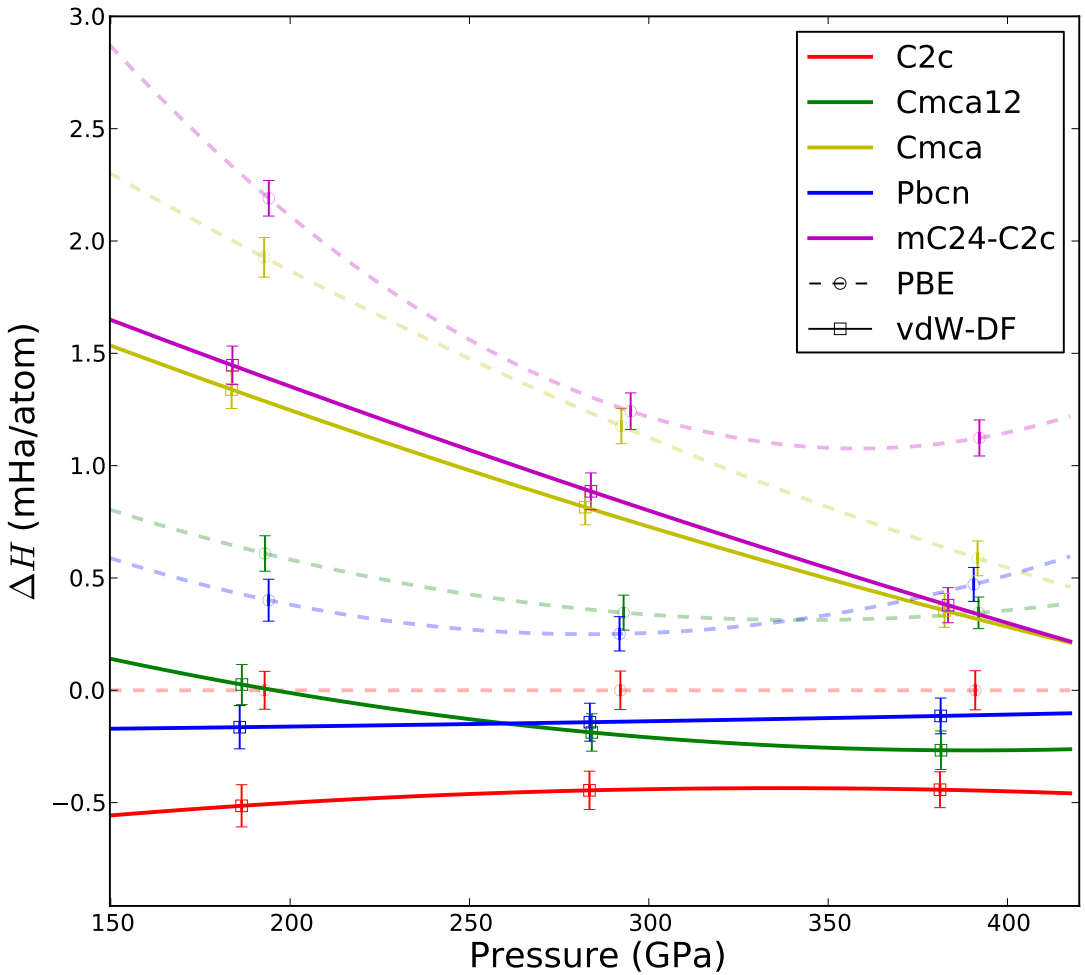
The following are common DFT applications which have very little competition from QMC.

1. Structural Optimization
   - Atomic positions and box geometries are chosen to minimize enthalpy.
   - Ab initio random structure searching.
2. Phonon Spectra
   - Frozen-phonon technique.
3. Quantum Molecular Dynamics

Prohibitively expensive in QMC because we don't usually have access to forces or stresses.

When Isn't DFT Enough?

- Plot[1] of QMC enthalpy vs. QMC pressure for several ground state hydrogen structures (C2c, Cmca12, Cmca, Pbcn, and mC24-C2c)
- Structures optimized with PBE (dashed lines) and vdW-DF (solid lines) functionals.
- Up to 1mHa/atom enthalpy differences



[1] Phys. Rev. B 89, 184106 (2014); doi:10.1103/PhysRevB.89.184106

Definitions

We will focus on energy derivatives w.r.t. structural deformations.

Force

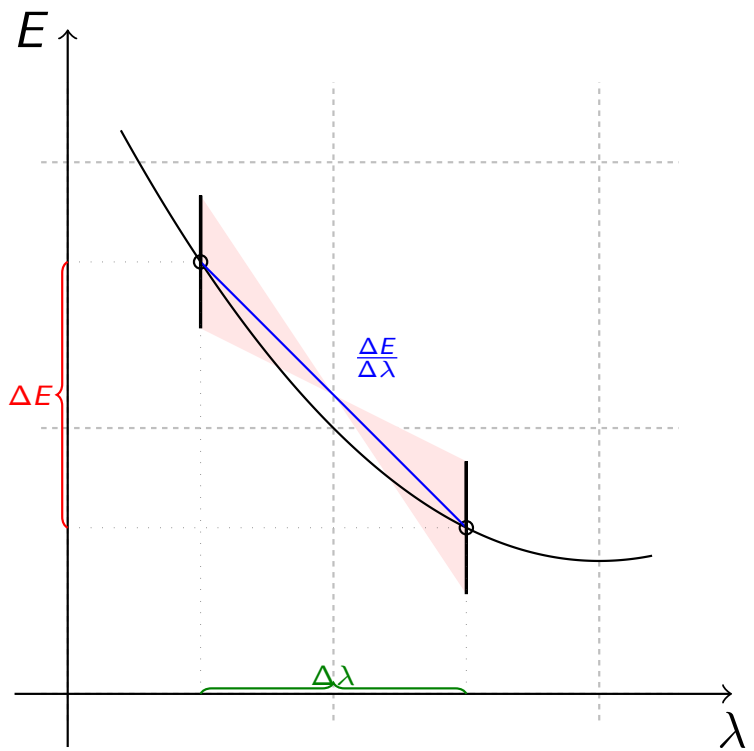
$$\mathbf{F}_\alpha = -\nabla_\alpha E(\{\mathbf{R}\}) \tag{1}$$

Stress

Consider an infinitesimal isotropic deformation $\mathbf{r}' = (\mathbb{I} + \epsilon)\mathbf{r}$

$$\sigma_{\alpha\beta} = -\frac{1}{\Omega}\frac{\partial E}{\partial \epsilon_{\alpha\beta}} \tag{2}$$

Energy Derivatives: General Problem



Finite-Differences

- We only know energy up to some error.
- Causes a trade-off between statistical and systematic error.

Finite Differencing Methods

2-point Finite Difference Formula

$$\frac{dE}{d\lambda} = \frac{E(\lambda + \Delta) - E(\lambda)}{\Delta} + \mathcal{O}(\Delta^2) \tag{3}$$

If there is statistical uncertainty in E, then

$$\begin{aligned} \mathrm{Var}\left[\frac{dE}{d\lambda}\right] = \quad & \frac{1}{\Delta^2}\{\ \mathrm{Var}\left[E(\lambda + \Delta)\right] + \mathrm{Var}\left[E(\lambda)\right] \\ & -2\mathrm{Cov}\left[E(\lambda + \Delta) \mid E(\lambda)\right]\ \} \end{aligned} \tag{4}$$

Notes:

- There is a trade-off between statistical and systematic error.
- Improve efficiency by maximizing $\mathrm{Cov}\left[E(\lambda + \Delta) \mid E(\lambda)\right]$

Correlated Sampling with VMC

Consider systems A and B, described by hamiltonians $\hat{H}_A$ and $\hat{H}_B$.
Non-Correlated Sampling:

$$\Delta E_{AB} = \left(\frac{1}{\int \Psi_A^2} \int \Psi_A^2 E_L^A \right) - \left(\frac{1}{\int \Psi_B^2} \int \Psi_B^2 E_L^B \right) \tag{5}$$

- Run two independent VMC simulations for the systems A & B.
- Calculate E_A and E_B in post processing.
- As $\Psi_A \to \Psi_B$, $\mathrm{Var}\left[\Delta E_{AB}\right] \to 2\mathrm{Var}\left[E_A\right]$

Correlated Sampling with VMC

Correlated Sampling

$$\Delta E_{AB} = \frac{1}{\int \Pi} \int \Pi \left(\frac{\Psi_A^2/\Pi}{\langle \Psi_A^2/\Pi \rangle} E_L^A - \frac{\Psi_B^2/\Pi}{\langle \Psi_B^2/\Pi \rangle} E_L^B \right) \tag{6}$$

- Run a single VMC simulation over the distribution Π
  - Chosen to minimize $\mathrm{Var}\left[\Delta E_{AB}\right]$
  - "Umbrella sampling": $\Pi = \Psi_A^2 + \Psi_B^2$
  - "Space-Warp[1] "
- As $\Psi_A \to \Psi_B$, $\mathrm{Var}\left[\Delta E_{AB}\right] \to 0$!

[1] PRB 61, 16291 (2000); doi: 10.1103/PhysRevB.61.R16291

Advantages

- It's a very general technique (keep it in your bag of tricks)
- Works rigorously with VMC and RMC (and DMC with approximations)

Disadvantages

- Need a different trial wavefunction for each perturbation.
- Need a minimum of 7 trial wavefunctions for stresses, and 3N+1 for forces.

Hellmann-Feynman Theorem

Theorem

If $|\Psi\rangle$ is an eigenstate of $\hat{H}$, or if $|\Psi\rangle$ is a variational minimum that doesn't explicitly depend on λ , then:

$$\frac{dE_\lambda}{d\lambda} = \langle\Psi|\frac{d\hat{H}}{d\lambda}|\Psi\rangle \tag{7}$$

$\hat{O} = \frac{d\hat{H}}{d\lambda}$ we take as our "Hellman-Feynman Estimator"

WARNING:

- Remember the mean and variance must exist.
- Subject to "mixed estimator" problem.

Nielsen & Martin Stress Estimator[2]

Stress Estimator

$$\hat{\sigma}_{\alpha\beta} = -\sum_k \frac{\hbar^2}{2m_k} \nabla_{k\alpha} \nabla_{k\beta} + \frac{1}{2} \sum_{k \neq k'} \frac{(\mathbf{x}_{kk'})_\alpha (\mathbf{x}_{kk'})_\beta}{x_{kk'}} \left(\frac{d}{dx_{kk'}} \hat{V} \right)$$

- Advanced feature in QMCPACK
- Finite variance.
- Mixed estimator.
- Currently works for all-electron calculations bulk calculations. Pseudopotentials later.

[2] PRB 32, 3780 (1985); doi: 10.1103/PhysRevB.32.3780

Using the Stress Estimator

```
<hamiltonian name="h0" type="generic" target="e">
    ...
    <estimator name="S" type="Force" mode="stress"
     source="ion0" target="e"/>
    ...
</hamiltonian>
```

σ_{ij} will appear as "S_i_j" in the scalar.dat file.

Stress Estimator Test

Stress	QMC (GPa)	LDA % Er	PBE % Er	vdW-DF % Er	vdW-DF2 % Er	HSE % Er
σ_{XX}	76.59 ± 0.16	6.31	13.32	21.64	25.43	4.42
σ_{YY}	73.79 ± 0.16	6.61	14.24	23.33	27.67	2.98
σ_{ZZ}	130.61 ± 0.14	-2.05	-2.70	-2.50	-2.64	-3.38
σ_{XY}	6.14 ± 0.12	0.94	-22.47	-45.95	-57.70	9.36
σ_{XZ}	-2.24 ± 0.11	-8.90	26.90	71.54	97.28	-28.97
σ_{YZ}	-2.84 ± 0.11	-40.72	-45.11	-51.22	-54.65	-0.64

- Tested on a pure hydrogen system with $N_e = 54$ and a density of $r_s = 1.60$.
- QMC stresses are finite-size corrected and extrapolated to reduce mixed-estimator bias.
- DFT errors with stresses are consistent with previous benchmarking studies.

Hellmann-Feynman Forces

We can try to evaluate the following estimator in QMC:

$$\hat{\mathbf{F}_\alpha} = -\overrightarrow{\nabla}_{\mathbf{R}_\alpha} V(\{\hat{\mathbf{r}}\}, \{\hat{\mathbf{R}}\}) \tag{8}$$

Problem: Estimator has a well defined mean, but infinite variance for $1/r$ potentials.

$$\langle \hat{\mathbf{F}} \rangle = -Z \int r^2 dr \, d\Omega \, \rho(\mathbf{r}) \frac{\hat{\mathbf{r}}}{r^2} \tag{9}$$

$$\langle \hat{\mathbf{F}^2} \rangle = Z^2 \int r^2 dr \, d\Omega \, \rho(\mathbf{r}) \frac{1}{r^4} \tag{10}$$

Ceperley-Chiesa-Zhang Estimator[3]

Eliminates divergence of the Hellmann-Feynman estimator by filtering out s-wave component of the force-density. How is this done?
First, create a sphere of radius $\mathcal{R}$ around ion.

$$\langle F_z \rangle = F_z^O + -Z \int_{in} d^3\mathbf{r}\rho(\mathbf{r})\frac{z}{r^2} \tag{11}$$

Define a force density as follows:

$$f_z(r) = -Z \int d\Omega \rho(r,\theta,\phi)\cos(\theta) \tag{12}$$

[3] PRL 94, 036404 (2005); doi: 10.1103/PhysRevLett.94.036404

Ceperley-Chiesa-Zhang Estimator

Physically, $f_z(r) \to 0$ linearly as $r \to 0$.

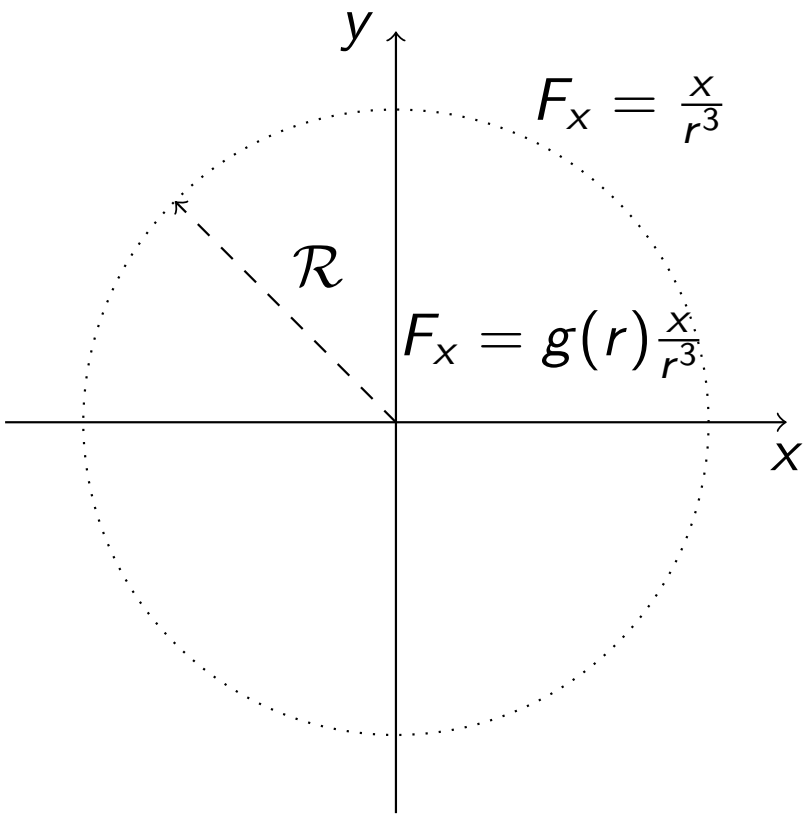
$$f_z(r) = -Z \int d\Omega \rho(r, \theta, \phi) \cos(\theta) \tag{13}$$

Given the general expansion of $f_z(r)$ in spherical harmonics:

$$f_z(r) = -Z \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} r^\ell f_m^\ell \int d\Omega Y_\ell^m(\theta, \phi) = -Z \sum_{\ell=0}^{\infty} a_\ell r^\ell \tag{14}$$

If we set the $\ell = 0$ term to zero, we filter out the s-wave component.

Chiesa-Ceperley-Zhang Estimator



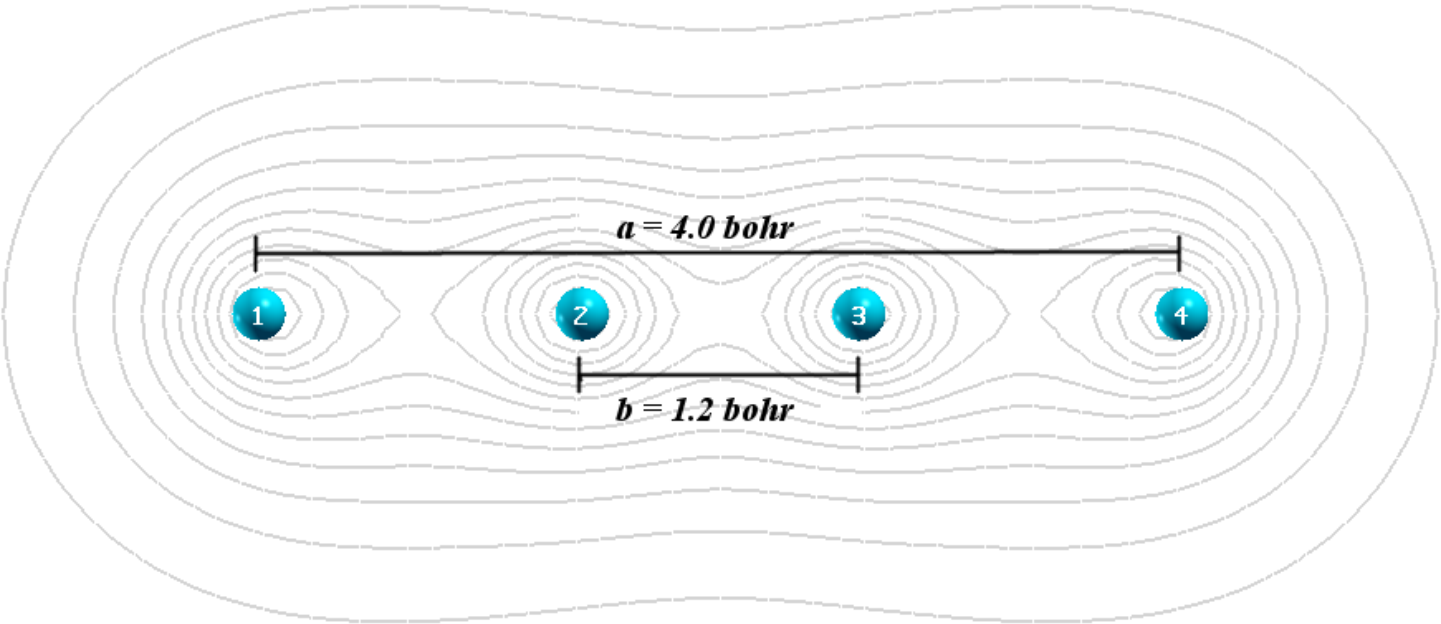
- $g(r) = \sum_{k=1}^{n} a_k r^{k+m}$
- On physical grounds, we can fix $g(r)$ to cancel off divergence.
- Not zero-variance, not zero-bias, but systematically improvable.
- Choose m, n, and $\mathcal{R}$ to minimize variance and bias.

Using the Force Estimator

```
<hamiltonian name="h0" type="generic" target="e">
   ...
   <estimator name="F" type="Force" mode="cep">
        <parameter name="rcut">1.0</parameter>
        <parameter name="nbasis">4</parameter>
        <parameter name="weightexp">2</parameter>
   </estimator>
   ...
</hamiltonian>
```

σ_{ij} will appear as "F_i_j" in the scalar.dat file.

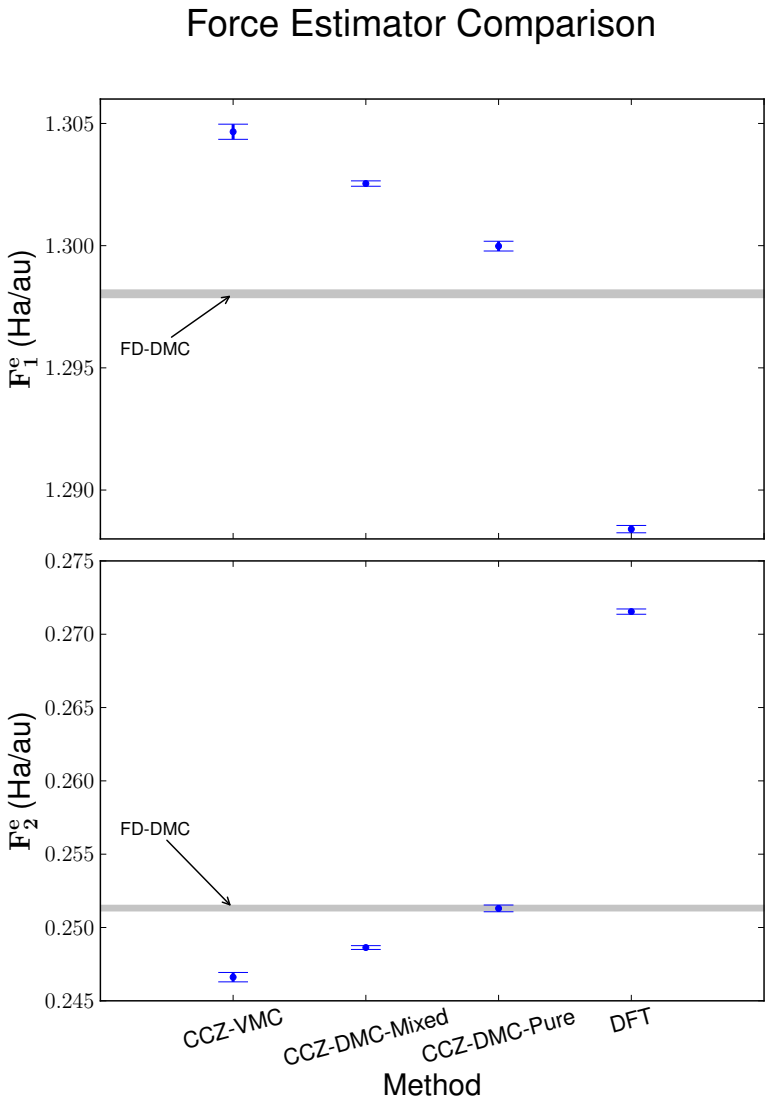
Chiesa-Ceperley-Zhang: Open Boundary Conditions



- DMC finite-difference forces used as reference.
- Tested Chiesa-Ceperley-Zhang Estimator ($\mathcal{R} = 0.4, n = 5$)

Force Estimator Tests: Results

- All QMC Chiesa estimates outperform PBE.
- Dramatically more efficient than finite-differences for large-scale calculations.
- Accuracy better than 0.15% compared to finite-differencing.

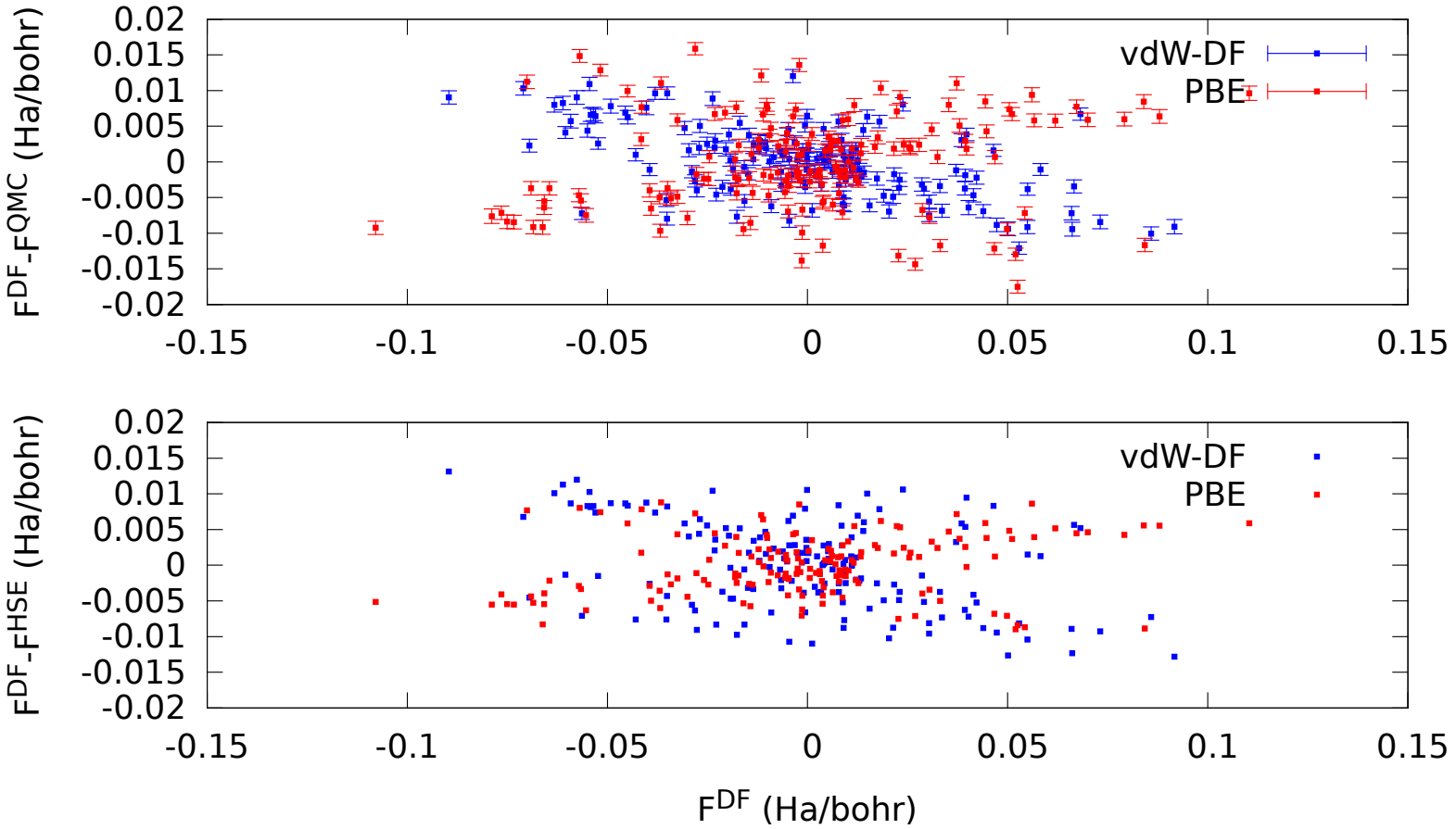


Chiesa-Ceperley-Zhang: Bulk Calculations

This framework also works for bulk calculations! Tested on solid and liquid H systems, as well as H+He mixtures.

- Uses the "optimized breakup" method of to handle the long-range force contributions.
- S-wave filtering on the short-range component of the force.

Chiesa-Ceperley-Zhang: Bulk Calculations



- Pure liquid hydrogen system. $N_e = 54$, $r_s = 1.60$.
- Benchmarking application. Force errors relative to QMC and HSE respectively.

Chiesa-Ceperley-Zhang: Summary

Advantages

- It's simple (two parameters in the input file).
- It's accurate (much better than 1%).
- It's efficient for light elements.

Disadvantages

- It's a mixed estimator.
- Error bar scales like Z^3 with atomic number.
- Estimator is not zero-variance.

Zero-Variance Zero-Bias Forces: Introduction

Consider the local energy estimator, $E_L(\mathbf{R}) = \frac{\hat{H}\Psi_T}{\Psi_T}$. It has the following desirable properties:

1. **Zero-variance property:**
   - $\mathrm{Var}\left[E_L\right] \to 0$ as $\Psi_T \to \Phi_0$.
2. **Zero-bias property:**
   - $\langle E_L \rangle_{\Psi_T^2} \to E_0$ as $\Psi_T \to \Phi_0$

Can we make other estimators that behave like this?

Assaraf-Caffarel ZVZB Estimator[4]

$$\frac{d\langle \hat{H} \rangle_{\Psi_T \Phi_0}}{d\lambda} = \left\langle \frac{d\hat{H}}{d\lambda} + \frac{(\hat{H} - E_L)\Psi_\lambda}{\Psi_T} + 2(E_L - \langle E_L \rangle)\frac{\Psi_\lambda}{\Psi_T} \right\rangle_{\Psi_T \Phi_0} \tag{15}$$

Ingredients:

1. Bare Hellman-Feynman estimator: $\frac{d\hat{H}}{d\lambda}$
2. Zero-variance term: $\frac{(\hat{H}-E_L)\Psi_\lambda}{\Psi_T}$
   - Cancels divergences in $\frac{d\hat{H}}{d\lambda}$
   - Reduces statistical noise.
3. Zero-bias term: $2(E_L - \langle E_L \rangle)\frac{\Psi_\lambda}{\Psi_T}$
4. Trial wavefunction Ψ_T and trial wavefunction derivative Ψ_λ

[4] J. Chem. Phys. 119, 10536 (2003); doi: 10.1063/1.1621615

ZVZB Advantages

As $\Psi_T \to \Phi_0$ and $\Psi_\lambda \to \Phi_\lambda$,

- Variance goes to 0
- Estimator approaches the true $\frac{dE_0}{d\lambda}$

Advantages over Chiesa-Ceperley-Zhang:

- Efficiency tied to quality of trial wavefunctions, not underlying estimator or distribution.
- Errors might scale much better than Z^3.

ZVZB Disadvantages

$$\hat{O} = \frac{d\hat{H}}{d\lambda} + \frac{(\hat{H} - E_L)\Psi_\lambda}{\Psi_T} + 2(E_L - \langle E_L \rangle)\frac{\Psi_\lambda}{\Psi_T} \tag{16}$$

1. Complexity
   - Need machinery to store, evaluate, and optimize trial wavefunction derivatives.
   - Fixed-node calculations require special techniques to handle nodal divergence.
2. What is a good trial wavefunction derivative?
   - Delicate treatment of nodal divergences needed.
   - What about core electrons?
3. What's the best way to optimize this estimator?
   - Explore cost functions
   - Simultaneous optimization of many estimators.

Current Research

- Sandro Sorella
  - Quantum Molecular Dynamics in Hydrogen.
    PRL 100, 114501 (2008); doi: 10.1103/PhysRevLett.100.114501
    Nat. Comm. 5, 3487 (2014); doi: 10.1038/ncomms4487
  - Algorithmic Differentiation:
    J. Chem. Phys. 133, 234111 (2010); doi: 10.1063/1.3516208
  - Assaraf-Caffarel estimator with VMC and DMC.
    J. Chem. Phys. 119, 10536 (2003); doi: 10.1063/1.1621615
- Poole, Foulkes, Spencer, Haynes:
  - Algorithmic differentiation and molecular dynamics in DMC.
    APS March Meeting 2014 http://meetings.aps.org/link/BAPS.2014.MAR.S27.3

Current Research

- Saccani, Filippi, Moroni:
  - NEB calculations of molecules using QMC.
    J. Chem. Phys. 138, 084109 (2013); doi: 10.1063/1.4792717
  - Improved ZV terms.
    ES2013 http://es13.wm.edu/talks/Moroni.pdf
- UIUC:
  - Benchmarking forces and stresses in H and H+He mixtures.
  - Using the Chiesa-Ceperley-Zhang estimator.
    PRL 94, 036404 (2005); doi: 10.1103/PhysRevLett.94.036404

Conclusions

- Improved estimators make it possible to bring QMC levels of accuracy to:
  - Structural optimization
  - Phonon calculations
  - Molecular Dynamics & Classical Monte Carlo
- QMCPACK supports stresses and Chiesa-Ceperley-Zhang estimators in isolated and bulk systems.
- Research is ongoing to extend this to realizing previously mentioned applications.

Anyone who is interested in discussions, examples, etc., feel free to talk to me!