Quantum Monte Carlo Simulation of van der Waals Systems

Anouar Benali¹, Luke Shulenburger², Nichols A. Romero¹, Jeongnim Kim³, O. Anatole von Lilienfeld¹

1- Argonne National Laboratory - Argonne, IL, USA
2- Sandia National Laboratories - Albuquerque, NM, USA
3- Oak Ridge National Laboratory - Oak Ridge, TN, USA

Contact: benali@anl.gov
QMCPACK Simulation Suite

- Jeongnim Kim. ORNL (Formerly UIUC)
- David Ceperley UIUC
- Luke Shulenburger SNL
- Ken Esler Stoneridge (Formerly UIUC)
- Miguel Morales LLNL
- Jeremy McMinis LLNL
- Nichols Romero ANL
- Anouar Benali ANL

IBM Blue Gene/Q
Mira @ Argonne National Laboratory
Mira: next-generation supercomputer

48 racks
1,024 nodes per rack
16 cores per node
64 threads per node
16GB memory/node
1.6GHz 16-way core processor
240 GB/s, 35 PB storage

786k cores
8.15 PF/s
BGQ - High Performance features

Quad FPU (QPU)
DMA unit
List-based prefetcher
TM (Transactional Memory)
SE (Speculative Execution)
Wakeup-Unit
Scalable Atomic Operations

Instruction Extensions (QPX) to PowerISA
4-wide double precision FPU SIMD (BG/L,P are 2-wide) usable as:
  scalar FPU
  4-wide FPU SIMD
  2-wide complex arithmetic SIMD
Attached to AXU port of A2 core – A2 issues one instruction/cycle to AXU
8 concurrent floating point operations (FMA) + load +store
  ▪ 6 stage pipeline
Permute instructions to reorganize vector data supports a multitude of data alignments
4R/2W register file
  32x32 bytes per thread
32B (256 bits) data path to/from L1 cache

Intrinsic:
TYPE: vector4double A;
Loads and stores
Unary operations
Binary operations
Multiply-add operations
Special functions
With four multiply-add units, the quad FPU can perform eight double-precision floating-point operations per cycle. The processor can execute up to eight double-precision floating-point operations, based on a fused multiply-add (FMA), along with an FP load and an FP store in a single cycle.

- The performance comes from the quad pipe Floating point unit.
- Each cycle, the quad FPU, can serve as a simple scalar FPU or a four wide SIMD FPU, or it can perform two complex arithmetic SIMD operations.
- All of these operations can be single or double precision.
Einspline

 Eval_z:
 Evaluation of spline coefficients (complex)

\[
\begin{align*}
&\text{for ( i = 0; i < 64; i++ )}
\text{ }\{ \\
&\quad s = d[i]; \\
&\quad p = (\text{double *})\text{coefs}[i]; \\
&\quad \text{for ( n = 0; n < M - rem ; n = n + 8)}\{ \\
&\quad\quad \text{double a0, a1,.., a7;} \\
&\quad\quad \text{double b0, b1, .., b7;} \\
&\quad\quad a0 = v[n+0]; \\
&\quad\quad //\text{code} \\
&\quad\quad a7 = v[n+7]; \\
&\quad\quad \text{b0 = p[n+0];} \\
&\quad\quad //\text{code} \\
&\quad\quad b7 = p[n+7]; \\
&\quad\quad \text{//operations} \\
&\quad\quad a0=a0+s*b0; \\
&\quad\quad \ldots \\
&\quad\quad v[n+0]=a0; \\
&\quad\quad \ldots \\
&\quad \}\}
\end{align*}
\]

Using QPX

\[
\begin{align*}
&\text{for ( i = 0; i < 64; i++ )}\{ \\
&\quad s = d[i]; \\
&\quad p = (\text{double *})\text{coefs}[i]; \\
&\quad \text{vector4double t = \{ s, s, s, s \};} \\
&\quad \text{for ( n = 0; n < M - rem ; n = n + 8)}\{ \\
&\quad\quad \text{vector4double f0, f1;} \\
&\quad\quad \text{vector4double g0, g1;} \\
&\quad\quad g0 = \text{vec}_ld( j, \quad p ); \\
&\quad\quad g1 = \text{vec}_ld( j+32, \quad p ); \\
&\quad\quad f0 = \text{vec}_ld( j, \quad v ); \\
&\quad\quad f1 = \text{vec}_ld( j+32, \quad v ); \\
&\quad\quad f0 = \text{vec}_madd( t, \quad g0, \quad f0 ); \\
&\quad\quad f1 = \text{vec}_madd( t, \quad g1, \quad f1 ); \\
&\quad\quad \text{vec}_st( \quad f0, \quad j, \quad v ); \\
&\quad\quad \text{vec}_st( \quad f1, \quad j+32, \quad v ); \\
&\quad \}\}
\end{align*}
\]
QMC Modelization

The many-body trial wavefunction

\[ \Psi_T(R) = J(R) \Psi_{AS}(R) = e^{J_1+J_2+...} \sum_k C_k D_k^\dagger(\phi)D_k(\phi) \]

Correlation (Jastrow)

\[ J_1 = \sum_i^{N} \sum_{i \neq j}^{N_{ions}} u_1(|r_i - r_j|) \]

\[ J_2 = \sum_{i \neq j}^{N} u_2(|r_i - r_j|) \]

Anti-symmetric function (Pauli principle)

\[ D_k^\sigma = \left[ \begin{array}{ccc} \phi_1(r_1) & \cdots & \phi_1(r_{N^\sigma}) \\ \vdots & \ddots & \vdots \\ \phi_{N^\sigma}(r_1) & \cdots & \phi_{N^\sigma}(r_{N^\sigma}) \end{array} \right] \]

Single-particle orbitals

\[ \phi_i = \sum_l C_l^i \Phi_l \]

Many methods of approximating the plane-wave-represented single-particle orbitals with polynomials:

-B-spline approximation in QMC, report significant reduction in time of calculation while maintaining plane-wave-level accuracy
### Profiling

**System:**
- Ar Solid – 32 atoms – 256 electrons – B-splines representation of WF (1.9Gb):
- 256 nodes – 32 threads – 2 Walkers per thread

Profile with original version of QMCPACK

**Flat profile:**

**Total run time:** 53min40

<table>
<thead>
<tr>
<th>Name</th>
<th>Time</th>
<th>Total</th>
<th>Ts/call</th>
</tr>
</thead>
<tbody>
<tr>
<td>eval_multi_UBspline_3d_z_vgh</td>
<td>56.95s</td>
<td>58369.57s</td>
<td>58369.57s</td>
</tr>
<tr>
<td>eval_multi_UBspline_3d_z</td>
<td>14.02s</td>
<td>72738.82s</td>
<td>14369.25s</td>
</tr>
<tr>
<td>SymmetricDTD</td>
<td>2.11s</td>
<td>77918.51s</td>
<td>2161.01s</td>
</tr>
<tr>
<td>EinsplineSetExtended::evaluate</td>
<td>1.70s</td>
<td>79663.07s</td>
<td>1744.56s</td>
</tr>
</tbody>
</table>

71% of the application time spent in the Spline evaluation of the Wave Function
Optimization

-> 2 algorithms accessing memory horizontally or vertically over the \{Points in space; minicube around the point\}

    Important reduction of the number of arithmetic operations

-> Complete rewriting of the functions with QPX intrinsics

    - Increase of the number of floating point operations
    - Reduction of the number of cycles per operation

-> Manual memory prefetching when possible

    Increase the availability of data in the L1 cache
# Profiling

## Coefficients (type)

<table>
<thead>
<tr>
<th></th>
<th>Eval_Z</th>
<th>Eval_D</th>
<th>Eval_S</th>
<th>Eval_Z_VGH</th>
<th>Eval_D_VGH</th>
<th>Eval_S_VGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eval_Z</td>
<td>Complex Double</td>
<td>Double</td>
<td>Float</td>
<td>Complex Double</td>
<td>Double</td>
<td>Float</td>
</tr>
<tr>
<td>Eval_D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eval_S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Coefficients, Gradients, Hessian (type)

<table>
<thead>
<tr>
<th></th>
<th>Eval_Z_VGH</th>
<th>Eval_D_VGH</th>
<th>Eval_S_VGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eval_Z_VGH</td>
<td>Complex Double</td>
<td>Double</td>
<td>Float</td>
</tr>
<tr>
<td>Eval_D_VGH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eval_S_VGH</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Speed up

<table>
<thead>
<tr>
<th></th>
<th>Eval_Z</th>
<th>Eval_D</th>
<th>Eval_S</th>
<th>Eval_Z_VGH</th>
<th>Eval_D_VGH</th>
<th>Eval_S_VGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm B</td>
<td>0.38</td>
<td>0.81</td>
<td>0.39</td>
<td>1.59</td>
<td>0.93</td>
<td>1.62</td>
</tr>
<tr>
<td>Algorithm M</td>
<td>2.48</td>
<td>0.91</td>
<td>1.02</td>
<td>2.15</td>
<td>1.01</td>
<td>0.95</td>
</tr>
</tbody>
</table>

## Algorithm (X) with QPX

<table>
<thead>
<tr>
<th></th>
<th>Eval_Z</th>
<th>Eval_D</th>
<th>Eval_S</th>
<th>Eval_Z_VGH</th>
<th>Eval_D_VGH</th>
<th>Eval_S_VGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm (X)</td>
<td>3.94</td>
<td>1.08</td>
<td>1.26</td>
<td>7.62</td>
<td>1.58</td>
<td>1.31</td>
</tr>
<tr>
<td>QPX + Prefetch</td>
<td>4.25</td>
<td>1.23</td>
<td>1.81</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### Profiling

**System:**
- Ar Solid – 32 atoms – 256 electrons – Bsplines WF (1.9Gb):
- 256 nodes – 32 threads – 2 Walkers per thread

Profile with QPX and Prefetch

Profile with Original Algorithm

<table>
<thead>
<tr>
<th>Flat profile:</th>
<th>Flat profile:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total run time:</strong> <strong>20min03</strong></td>
<td><strong>Total run time:</strong> <strong>53min40</strong></td>
</tr>
</tbody>
</table>

Each sample counts as 0.01 seconds.

<table>
<thead>
<tr>
<th>% cumulative self time</th>
<th>seconds</th>
<th>seconds</th>
<th>seconds</th>
<th>seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.08</td>
<td>5380.43</td>
<td>5380.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.25</td>
<td>12270.83</td>
<td>3152.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.68</td>
<td>14441.45</td>
<td>2170.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.85</td>
<td>16292.97</td>
<td>1851.52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- `.eval_multi_USB spline_3d_z_vgh`
- `.eval_multi_USB spline_3d_z`
- `.SymmetricDTD`
- `EinsplineSetExtended::evaluate`

<table>
<thead>
<tr>
<th>% cumulative self time</th>
<th>seconds</th>
<th>seconds</th>
<th>seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>56.95</td>
<td>58369.57</td>
<td>58369.5</td>
<td></td>
</tr>
<tr>
<td>14.02</td>
<td>72738.82</td>
<td>14369.25</td>
<td></td>
</tr>
<tr>
<td>2.11</td>
<td>77918.51</td>
<td>2161.01</td>
<td></td>
</tr>
<tr>
<td>1.70</td>
<td>79663.07</td>
<td>1744.5</td>
<td></td>
</tr>
</tbody>
</table>

**Total run time Speedup of 2.68 times**
HPM PROFILING

Original Code

27.644.290.379.027  All XU Instruction
22.786.190.220.714  All AXU Instruction
43.043.218.198.088  FP Operations Group 1

Derived metrics for code block "mpiAll" averaged over process(es) on node <0,0,0,0,0>:
Instruction mix:  FPU = 45.18 %,  FXU = 54.82 %
Instructions per cycle completed per core = 0.6138
Per cent of max issue rate per core = 33.65 %
Total weighted GFlops for this node = 13.412
Loads that hit in L1 d-cache = 94.03 %
  L1P buffer = 5.36 %
  L2 cache = 0.35 %
  DDR = 0.26 %
DDR traffic for the node:  ld = 1.508,  st = 0.540,  total = 2.049 (Bytes/cycle)

Percentage of peak= 6.55%

Total run time Speedup of 2.68 times

BGQ optimized Code

8.581.366.867.332  All XU Instruction
4.896.512.230.816  All AXU Instruction
13.017.533.928.058  FP Operations Group 1

Derived metrics for code block "mpiAll" averaged over process(es) on node <0,0,0,0,0>:
Instruction mix:  FPU = 36.33 %,  FXU = 63.67 %
Instructions per cycle completed per core = 0.4417
Per cent of max issue rate per core = 28.12 %
Total weighted GFlops for this node = 10.922
Loads that hit in L1 d-cache = 88.60 %
  L1P buffer = 5.92 %
  L2 cache = 4.50 %
  DDR = 0.98 %
DDR traffic for the node:  ld = 3.503,  st = 1.101,  total = 4.604 (Bytes/cycle)

Percentage of peak= 5.33%
Application speedup using QPX and prefetching is **2.68** folds from original Algorithm.
Overall Scaling on LCF Architectures

(a) Weak Scaling
- **Mira (256)**
- **Titan/CPU (64)**
- **Titan/GPU (64)**

(b) Strong Scaling
- **Mira (1024)**
- **Titan/CPU (512)**
- **Titan/GPU (256)**
Applications on van der Waals dominated systems
Van der Waals forces are important
-> Noble gases are proto-typical! We use Ar as a case of principle.
- London \((C_6/R^6)\) widely used in force fields (Lennard-Jones tail)
- Axilrod-Teller-Muto \((C_9/R^9)\) is 3-body analogue

Dispersion Coefficients \((C_6, C_9)\)

London\(^1,\)\(^2\)

Axilrod-Teller-Muto\(^3\)

\[
W(R) = -C_6/R^6
\]

\[
W(R) = C_9 \left( 3 \cos[\varphi] \cos[\varphi] \cos[\varphi] + 1 \right)/R^9
\]

\(^1\) W. Heitler and F. London, Z. Phys. 44, 455 (1927)
\(^2\) R. Eisenschitz and F. London, Z. Phys. 60, 491 (1930)
\(^3\) B. M. Axilrod and E. Teller, J. Chem. Phys. 11, 299 (1943)
Objectives

Van der Waals forces are important
- Noble gases are proto-typical! We use Ar as a case of principle.
  - London ($C_6/R^6$) widely used in force fields (Lennard-Jones tail)
  - Axilrod-Teller-Muto ($C_9/R^9$) is 3-body analogue

- Argon EOS
  - Evaluation of the 2body, 3 body and MBC to the crystal solid (in progress)

- We apply the method to Ellipticine and DNA
  - Binding Energy of the drug Ellipticine to DNA
QMC Modelization

- We Solve the many-body Schrödinger equation and we express the wavefunction as follow;

\[
\Psi_T(x_1, x_1, \ldots, x_N) = J(x_1, x_1, \ldots, x_N) \Psi_{AS}(x_1, x_1, \ldots, x_N)
\]

DFT Calculation (LDA functional)

\[
J_1(\mathbf{R}) = \prod_{i} \exp \left[ \sum_{k} (b_{k} r_{i} + c_{k}) v_{k}(r_{i}) \right]
\]

One-Body + Two-body Jastrow

\[
J_2(\mathbf{R}) = \prod_{i<j} \exp \left[ \sum_{k} (b_{k} r_{ij} + c_{k}) v_{k}(r_{ij}) \right]
\]

Variational Monte Carlo

\[
E_{VMC} = \min_{\alpha} \langle \Psi_T(\mathbf{R};\alpha) | \hat{H} | \Psi_T(\mathbf{R};\alpha) \rangle
\]

Diffusion Monte Carlo

\[
E_{DMC} = \langle \phi_0 | \hat{H} | \Psi_T \rangle, \phi_0 = \lim_{\beta \rightarrow \infty} \exp^{-\beta \hat{H}} \Psi_T
\]

Trial Wavefunction (PWSCF)

Optimization of the factors (convergence using VMC)

New Trial Wavefunction

- Solid: Corrections to finite sizes effects, Kinetic and MPC, twists averaging
Argon Systems

<table>
<thead>
<tr>
<th>$d_{\text{Ar-Ar}}$ (Å)</th>
<th>$E_2$ (meV)</th>
<th>$E_3$ (meV)</th>
<th>$C_6$</th>
<th>$C_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>This Work</td>
<td>3.757</td>
<td>-12.232</td>
<td>0.289</td>
<td>63.1</td>
</tr>
<tr>
<td></td>
<td>±0.987</td>
<td>±0.567</td>
<td></td>
<td>517.6</td>
</tr>
<tr>
<td>Ref.</td>
<td>3.76$^a$</td>
<td>-12.3$^a$</td>
<td>0.3$^b$</td>
<td>64.3$^c$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>518$^c$</td>
</tr>
</tbody>
</table>


Argon Solid

QMC FCC energies for 108 atom supercell of Ar

Binding energy (from fit) = 88.1, with qha = 80.3 +/- 1.0 meV
Binding energy (from isolated) = 111.6, with qha = 103.8 +/- 1.0 meV
experiment = 80.1 meV

Bulk Modulus = 3.8, with qha = 3.3 +/- 0.1 GPa
experiment = 2.7 GPa

Lattice Constant = 5.280, with qha = 5.340 +/- 0.014 Angstrom
experiment = 5.311 Angstrom

Vinet Fit Goodness (reduced chisquare) = 3.24
Argon Solid

Experimental value 80.1meV
- Ellipticine is a planar polycyclic aromatic molecule
- Bind to DNA by non-covalent pi-pi-stacking with the nucleic acid Watson-Crick base pairs
- Binding energy is directly correlated to biological activity of the molecule in cancer treatment
Ellipticin

Predicting Noncovalent Interactions between Aromatic Biomolecules with London-Dispersion-Corrected DFT

I-Chun Lin, O. Anatole von Lilienfeld,† Maurício D. Coutinho-Neto,‡ Ivano Tavernelli, and Ursula Rothlisberger*  

Laboratory of Computational Chemistry and Biochemistry, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

Received: June 27, 2007; In Final Form: September 17, 2007
### Collective many-body van der Waals interactions in molecular systems

Robert A. DiStasio, Jr.*; O. Anatole von Lilienfeld,* and Alexandre Tkatchenko†

*Department of Chemistry, Princeton University, Princeton, NJ 08544; "Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439; and "Fritz-Haber-Institut der Max-Planck-Gesellschaft", Faradayweg 4-6, 14195 Berlin, Germany

Edited by Peter J. Rossky, The University of Texas at Austin, Austin, TX, and approved July 27, 2012 (received for review May 22, 2012)

Van der Waals (vdW) interactions are ubiquitous in molecules and condensed matter, and play a crucial role in determining the structure, stability, and function for a wide variety of systems. The accurate prediction of these interactions from first principles is a substantial challenge because they are inherently quantum mechanical phenomena that arise from correlations between many interacting quantum subsystems.

**Results and Discussion**

To accurately compute the nonadditive many-body vdW energy, we begin by performing a self-consistent quantum mechanical calculation to generate the molecular electron density using semilocal density-functional theory (DFT) (23)—a method which accurately predicts many of the fundamental properties of matter. We then use the resulting electron density as input to a nonempirical many-body method, the nonlocal dispersion corrected density functional theory (DFT-D), to compute the dispersion energy. The DFT-D results are compared to results generated by other methods that are based on semiempirical potentials (19) and empirical potentials (20). The comparison demonstrates the limits of these methods, which fail to accurately capture the many-body nature of vdW interactions.

**Level of theory**

<table>
<thead>
<tr>
<th>Level of theory</th>
<th>$\Delta E_{\text{Bind}}$ (Kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFT$^1$</td>
<td>+5.2</td>
</tr>
<tr>
<td>vDW-TS$^1$</td>
<td>-46.6</td>
</tr>
<tr>
<td>vDW-TB$^1$</td>
<td>-39.1</td>
</tr>
<tr>
<td>PBE-D3/QZVP$^2$</td>
<td>-35.68 (D2) ; -32.84 (D2+D3)</td>
</tr>
<tr>
<td>PBE-NL/QZVP$^2$</td>
<td>-39.11 (D2) ; -36.27 (D2+D3)</td>
</tr>
<tr>
<td>dDSC-PBE/QZ4P$^2$</td>
<td>-40.91 (D2) ; -38.07 (D2+D3)</td>
</tr>
<tr>
<td>vDW-MB$^1$</td>
<td>-34</td>
</tr>
</tbody>
</table>

**References**

2. S. Grimme – Private communication

---

**Ellipticine**

![Ellipticine Structure](image)

| Base stacking | Between AMBER and MP2 data, similar to our preceding studies of intermolecular correlation (dispersion) effects compared with diffuse sp shells. Such a basis set still covers a smaller fraction and they compared their data with those obtained with the Moeller monocation intercalator amiloride using the second-order nucleobases and intercalator. Bondarev and co-workers thor-complexes, and there is so far only a single paper reporting presently available force fields, (5) Pullman, B.


---

**Collective many-body van der Waals interactions in molecular systems**

Robert A. DiStasio, Jr.*; O. Anatole von Lilienfeld,* and Alexandre Tkatchenko†

*Department of Chemistry, Princeton University, Princeton, NJ 08544; "Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439; and "Fritz-Haber-Institut der Max-Planck-Gesellschaft", Faradayweg 4-6, 14195 Berlin, Germany

Edited by Peter J. Rossky, The University of Texas at Austin, Austin, TX, and approved July 27, 2012 (received for review May 22, 2012)

Van der Waals (vdW) interactions are ubiquitous in molecules and condensed matter, and play a crucial role in determining the structure, stability, and function for a wide variety of systems. The accurate prediction of these interactions from first principles is a substantial challenge because they are inherently quantum mechanical phenomena that arise from correlations between many interacting quantum subsystems.

**Results and Discussion**

To accurately compute the nonadditive many-body vdW energy, we begin by performing a self-consistent quantum mechanical calculation to generate the molecular electron density using semilocal density-functional theory (DFT) (23)—a method which accurately predicts many of the fundamental properties of matter. We then use the resulting electron density as input to a nonempirical many-body method, the nonlocal dispersion corrected density functional theory (DFT-D), to compute the dispersion energy. The DFT-D results are compared to results generated by other methods that are based on semiempirical potentials (19) and empirical potentials (20). The comparison demonstrates the limits of these methods, which fail to accurately capture the many-body nature of vdW interactions.

**Level of theory**

<table>
<thead>
<tr>
<th>Level of theory</th>
<th>$\Delta E_{\text{Bind}}$ (Kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFT$^1$</td>
<td>+5.2</td>
</tr>
<tr>
<td>vDW-TS$^1$</td>
<td>-46.6</td>
</tr>
<tr>
<td>vDW-TB$^1$</td>
<td>-39.1</td>
</tr>
<tr>
<td>PBE-D3/QZVP$^2$</td>
<td>-35.68 (D2) ; -32.84 (D2+D3)</td>
</tr>
<tr>
<td>PBE-NL/QZVP$^2$</td>
<td>-39.11 (D2) ; -36.27 (D2+D3)</td>
</tr>
<tr>
<td>dDSC-PBE/QZ4P$^2$</td>
<td>-40.91 (D2) ; -38.07 (D2+D3)</td>
</tr>
<tr>
<td>vDW-MB$^1$</td>
<td>-34</td>
</tr>
</tbody>
</table>

**References**

2. S. Grimme – Private communication
Collective many-body van der Waals interactions in molecular systems

Robert A. DiStasio, Jr.,* O. Anatole von Lilienfeld*, and Alexandre Tkatchenko†

*Department of Chemistry, Princeton University, Princeton, NJ 08544; †Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439; and ‡Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

Edited by Peter J. Rossky, The University of Texas at Austin, Austin, TX, and approved July 27, 2012 (received for review May 22, 2012)

Van der Waals (vdW) interactions are ubiquitous in molecules and condensed matter, and play a crucial role in determining the structure, stability, and function for a wide variety of systems. The accurate prediction of these interactions from first principles is a substantial challenge because they are inherently quantum mechanical.

Results and Discussion
To accurately compute the nonadditive many-body vdW energy, we begin by performing a self-consistent quantum mechanical calculation to generate the molecular electron density using semilocal density-functional theory (DFT) (23)—a method which accurately accounts for electron correlation.
Collective many-body van der Waals interactions in molecular systems

Robert A. DiStasio, Jr.*, O. Anatole von Lilienfeld*, and Alexandre Tkatchenko*

Van der Waals (vdW) interactions are ubiquitous in molecules and condensed matter, and play a crucial role in determining the structure, stability, and function for a wide variety of systems. The accurate prediction of these interactions from first principles is a substantial challenge because they are inherently quantum mechanical. The nondimensional many-body vdW energy, 

\[ E_{\text{vdW}} = \sum_{i<j} \frac{C_{6}}{r_{ij}^6} - \sum_{i<j} \frac{C_{12}}{r_{ij}^{12}} \]

where \( r_{ij} \) is the interatomic distance and \( C_{6} \) and \( C_{12} \) are the dispersion and exchange-correlation coefficients, respectively.

Results and Discussion

To accurately compute the nonadditive many-body vdW energy, we begin by performing a self-consistent quantum mechanical calculation to generate the molecular electron density using semilocal density-functional theory (DFT) (23)—a method which accurately}

## Table 1. Binding energies for the DNA–ellipticine complex in kcal/mol

<table>
<thead>
<tr>
<th>Level of theory</th>
<th>( \Delta E_{A^T-T}^{\text{DFT}} )</th>
<th>( \Delta E_{G-G}^{\text{DFT}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFT</td>
<td>+4.2</td>
<td>+1.9</td>
</tr>
<tr>
<td>vdW-TS</td>
<td>+2.5</td>
<td>-3.7</td>
</tr>
<tr>
<td>vdW-TB</td>
<td>+2.6</td>
<td>-3.5</td>
</tr>
<tr>
<td>vdW-MB</td>
<td>-0.1</td>
<td>-8.2</td>
</tr>
</tbody>
</table>

(Left) Binding energies (\( \Delta E_{\text{bind}}^{\text{DFT}} \)) for the DNA–ellipticine complex in kcal/mol. (Right) Relative conformational energies of A-DNA and B-DNA (\( \Delta E_{B-A} = E_{B} - E_{A} \)) consisting of pure adenine–thymine (A:T) and cytosine–guanine (C:G) sequences in kcal/mol per bp. All DFT calculations were performed using the PBE functional (37).

Fig. 2. Percentagewise convergence of the individual vdW-NB contributions with respect to the vdW-MB energy. Displayed cases include the binding energy of the DNA-ellipticine complex (blue circles) and the relative binding energies of a single base pair in A-DNA and B-DNA consisting of pure adenine-thymine (black triangles) and pure cytosine-guanine (red squares) sequences. The unfilled markers at \( N = 2 \) correspond to the predictions of the vdW-TB effective pairwise model for each of the aforementioned systems.
Conclusion

- QMC is a great method but expensive. Requires tuning on supercomputers.

- Rare gas study confirms that the method goes below the Kcal/mol accuracy and reproduces CCSD (T) results.

- Van der Waals corrected DFT methods have improved greatly by the inclusion of Manybody effects. However, are still predicting energies widely spread. With a QMC benchmark, the order of the manybody-vdw correction can be controlled to reproduce DMC energies.
Acknowledgments


- This research used resources of the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357.

- Sandia National Laboratories is a multiprogram laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-AC04-94AL85000.
INCITE: Innovative and Novel Computational Impact on Theory and Experiment

INCITE promotes transformational advances in science and technology through large allocations of computer time, supporting resources, and data storage at the Argonne and Oak Ridge Leadership Computing Facilities (LCF[s]) for computationally intensive, large-scale research projects.
## Allocation Programs at the LCFs

<table>
<thead>
<tr>
<th></th>
<th>INCITE</th>
<th>ALCC</th>
<th>Director’s Discretionary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mission</strong></td>
<td>High-risk, high-payoff science that requires LCF-scale resources*</td>
<td>High-risk, high-payoff science aligned with DOE mission</td>
<td>Strategic LCF goals</td>
</tr>
<tr>
<td><strong>Call</strong></td>
<td>1x/year – (Closes June)</td>
<td>1x/year (Closes February)</td>
<td>Rolling</td>
</tr>
<tr>
<td><strong>Duration</strong></td>
<td>1-3 years, yearly renewal</td>
<td>1 year</td>
<td>3m, 6m, 1 year</td>
</tr>
<tr>
<td><strong>Typical Size</strong></td>
<td>30 - 40 projects</td>
<td>10M - 100M core-hours/yr.</td>
<td>100s of projects</td>
</tr>
<tr>
<td><strong>Review Process</strong></td>
<td>Scientific Peer-Review</td>
<td>Computational Readiness</td>
<td>10K – 1M core-hours</td>
</tr>
<tr>
<td><strong>Managed By</strong></td>
<td>INCITE management committee (ALCF &amp; OLCF)</td>
<td>DOE Office of Science</td>
<td>LCF management</td>
</tr>
<tr>
<td><strong>Availability</strong></td>
<td>Open to all scientific researchers and organizations</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Capability >20% of cores
## Twofold review process

<table>
<thead>
<tr>
<th>Step</th>
<th>New proposal assessment</th>
<th>Renewal assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Peer review: INCITE panels</strong></td>
<td><strong>Renewal assessment</strong></td>
</tr>
<tr>
<td></td>
<td>• Scientific and/or technical merit</td>
<td>• Change in scope</td>
</tr>
<tr>
<td></td>
<td>• Appropriateness of proposal method, milestones given</td>
<td>• Met milestones</td>
</tr>
<tr>
<td></td>
<td>• Team qualifications</td>
<td>• On track to meet future milestones</td>
</tr>
<tr>
<td></td>
<td>• Reasonableness of requested resources</td>
<td>• Scientific and/or technical merit</td>
</tr>
<tr>
<td>2</td>
<td><strong>Computational readiness review: LCF centers</strong></td>
<td><strong>Award Decisions</strong></td>
</tr>
<tr>
<td></td>
<td>• Technical readiness</td>
<td>• INCITE Awards Committee comprised of LCF directors, INCITE program manager, LCF directors of science, sr. management</td>
</tr>
<tr>
<td></td>
<td>• Appropriateness for requested resources</td>
<td>• Met technical/computational milestones</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• On track to meet future milestones</td>
</tr>
</tbody>
</table>