Current Limits of \textit{ab initio} Molecular Dynamics Simulations

François Gygi

\textit{Center for Applied Scientific Computing}

\textit{Lawrence Livermore National Laboratory}

fgygi@llnl.gov

UCRL-PRES-203443 Work performed under the auspices of the U. S. Department of Energy by University of California Lawrence Livermore National Laboratory under Contract W-7405-Eng-48
Ab initio Molecular Dynamics: an accurate atomic-scale simulation method

- Solid state physics
- Surface physics
- Chemical Physics
- High pressure physics
- Nanotechnology
- Biochemistry

No empirical parameters.
No experimental input.

Predictive simulations.
The many scales of materials modeling

- **Continuum**
  - ALE3D
  - Plasticity of complex shapes

- **Mesoscale**
  - NIKE3D
  - Aggregate grain response, poly-crystal plasticity

- **Microscale**
  - Dislocation Dynamics
  - Collective behavior of defects, single crystal plasticity
  \[ \sigma = \sigma(\varepsilon, \dot{\varepsilon}, T, P, \ldots) \]

- **Atomic Scale**
  - Molecular Dynamics
  - Unit mechanisms of defect mobility and interaction

Length Scale:
- nm (10^{-9} m)
- \( \mu \text{m} \) (10^{-6} m)
- mm (10^{-3} m)

Time Scale:
- ps
- ns
- ms
- s
Relevance of ab initio MD in fission/fusion materials

- Vacancy and self-interstitial energy calculations
- He-vacancy interactions
- Validation of interatomic potentials
- Reactivity of SC water
- Erosion of first wall
Density Functional Theory: The Kohn-Sham equations

- Coupled, non-linear, integro-differential equations:

\[
- \Delta \varphi_i + V(\rho, \mathbf{r}) \varphi_i = \varepsilon_i \varphi_i \quad i = 1 \ldots N_{el}
\]

\[
V(\rho, \mathbf{r}) = V_{ion}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + V_{XC}(\rho(\mathbf{r}), \nabla \rho(\mathbf{r}))
\]

\[
\rho(\mathbf{r}) = \sum_{i=1}^{N_{el}} |\varphi_i(\mathbf{r})|^2
\]

\[
\int \varphi_i^*(\mathbf{r}) \varphi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}
\]

- Overall computational complexity: \(O(N^3)\) for \(N\) electrons
Plane-wave *ab initio* MD implementations rely on efficient algorithms

- **Key algorithm #1:** Complex 3-d FFT
  - Optimized FFT libraries are available
    - FFTW (M. Frigo, MIT)
    - TU Wien group (F. Franchetti)

- **Key algorithm #2:** Matrix multiplication
  - >90% of peak on BG/L single-node
  - Parallel implementation: ScaLAPACK
  - DGEMM is the asymptotic bottleneck of *ab initio* MD for very large sizes
Ab initio MD implementations at LLNL

- Plane-wave, pseudopotential method ($O(N^3)$)
- GP: production code used in ~15 projects at LLNL
  - Ported to Linux/x86, AIX, HP/Tru64
  - C++/MPI/OpenMP parallelism
- Qbox: a new implementation
  - MPI only
  - xlC, g++, icc
  - Web-aware, XML interface (Apache Xerces-C)
  - Component testing started 04/2003
  - First MD simulation: 09/2003
  - 40% of peak on 32 Power3 CPUs (measured with hpmcount)
Example of application: the structure of elemental Boron

- The exact structure of Boron in normal conditions is still debated
- Recent experiments suggest that Boron undergoes an amorphization under pressure (Sanz, Loubeyre, CEA/ESRF)
- First-Principles simulations are the tool of choice to investigate these issues

T. Ogitsu, G. Galli, LLNL
MD simulations of Boron under high pressure

- DFT simulations of Boron on MCR: a 2000 CPU Linux cluster
  - Over 2 weeks of full machine runs (T.Ogitsu, G.Galli, PAT/LLNL)
  - Computed static compressibility and electronic structure up to 1.8 Mbar
  - 1280 atoms and 3840 electrons, the largest First-Principles MD Materials Science simulation to date
  - Results show complex interplay between interstitial disorder and electronic structure
High-Pressure Physics: Two-phase simulation of Lithium Hydride

- **Two-phase** ab initio molecular dynamics
- Start from a solid-liquid interface
- Constant pressure, constant temperature (NPT) simulation
- Observe solidification or liquefaction

Parallel implementation: hierarchical structure of the electronic wavefunctions

Each single-particle wavefunction is represented as a Fourier series, or on a 3-d grid.
Specialized transform-and-interpolate FFT

- Many FFTs operate on sparse datasets
- ~2x speedup when working only on non-zero segments
Custom parallel 3-d FFT scaling up to 256 tasks

- Double precision complex transforms
- IBM power3 16 tasks/node, up to 16 nodes

FFT timings, N\times N\times N complex

19 ms / transform
Qbox scaling up to 1960 CPUs

- LLNL MCR Linux platform (2300 Pentium4 CPUs, Quadrics switch)
  - 1536 atoms (512 H₂O molecules)
  - Many additional tuning opportunities still to be explored

![Graph showing FPMD iteration time](image)

- 1.84x Speedup from 980 to 1960 tasks
- Note: strong scaling
New LLNL platforms

- Thunder (23TF)
  - 4k CPUs in 1k nodes
  - Moving to Itanium 64 bit architecture

- BlueGene/L (180/360TF)
  - 65536 nodes
  - 3D Torus and tree networks
  - Code must fit in 256MB/node
  - Limited operating system on compute nodes
  - Currently testing on a 512-node BG/L prototype
Porting *ab initio* MD to new architectures

- **Needs**
  - Availability of a standard ISO C++ compiler
  - *limited* MPI implementation (no MPI-2)
  - Libraries: BLAS, Lapack, BLACS, ScaLAPACK
  - Efficient FFTs (multiple 1-d complex)

- **Programming model:**
  - MPI

- **High-level C++ design reduces cost of porting**
What are the obstacles on the road to 10000 atoms?

- The cost of solving the Kohn-Sham equations is eventually dominated by orthogonalization ($O(N^3)$)

\[
\begin{align*}
-\Delta \varphi_i + V(\rho, \mathbf{r}) \varphi_i &= \varepsilon_i \varphi_i \quad i = 1 \ldots N_{el} \\
V(\rho, \mathbf{r}) &= V_{\text{ion}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + V_{\text{XC}}(\rho(\mathbf{r}), \nabla \rho(\mathbf{r})) \\
\rho(\mathbf{r}) &= \sum_{i=1}^{N_{el}} |\varphi_i(\mathbf{r})|^2 \\
\int \varphi_i^*(\mathbf{r}) \varphi_j(\mathbf{r}) d\mathbf{r} &= \delta_{ij}
\end{align*}
\]
Linear-scaling methods

- The “Holy Grail” of electronic structure methods: achieve *linear scaling*: $O(N)$ operations
- Introduce approximations to reduce the computational cost from $O(N^3)$ to $O(N)$.
- Several approaches proposed in the past 10 years
- Most successful approach: represent the solutions of the Kohn-Sham equations in terms of non-orthogonal, localized functions.
- We aim at a *controlled accuracy, $O(N)$ method*
  - Simple parameters (e.g. grid spacing) control numerical accuracy
  - As robust as $O(N^3)$ methods
Linear-scaling methods

- Goal: make all matrices sparse in

\[
E(Y) = \text{tr}(S^{-1}Y^T HY) \quad Y \in \mathbb{R}^{M \times N} \quad S = Y^T Y
\]

\[
S_{ij} = \langle \phi_i | \phi_j \rangle = \int_{\Omega} \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) \, d^3 \mathbf{r}
\]

\[
H_{ij} = \langle \phi_i | H \phi_j \rangle = \int_{\Omega} \phi_i^*(\mathbf{r}) H \phi_j(\mathbf{r}) \, d^3 \mathbf{r}
\]
Linear-scaling methods

- A real-space, finite-difference scheme meets the requirements for linear-scaling
- Spherical subdomains, with Dirichlet b.c.
- A new algorithm allows for adaptation of the localization centers (localization domains follow orbitals during MD simulations)
**O(N) with Controlled accuracy: localization radii**

- Errors in computed ionic forces are decaying exponentially for large localization radii.
- Errors are computed by comparison with $O(N^3)$ method with same numerical approximations.
$O(N)$ Molecular dynamics with adaptive localization centers

- Deuterium at 1000K
- Small energy drift (2x10^{-3} a.u./ps)
- Physical properties identical to those obtained in an $O(N^3)$ calculation
Coupling ab initio MD to other methods

- We are developing software for coupling
  - DFT MD – continuum
  - DFT MD – classical MD
  - DFT MD – Quantum Monte-Carlo
A coupled *ab initio* MD / polarizable continuum model for simulation in H$_2$O

- Compute the effective electrostatic potential by solving

  $$\nabla \cdot (\varepsilon(\rho(\mathbf{r}))\nabla \phi) = 4\pi\rho$$

  $$w = \varepsilon^{\frac{1}{2}} \phi$$

  $$-\nabla^2 w + pw = q$$


- Finite-difference, 4th order Mehrstellen operator, multigrid algorithm.

Polarizable medium

\[ \varepsilon \sim 80 \]

“Quantum mechanical” cavity

\[ \varepsilon \sim 1 \]
Ab initio MD coupled to a polarizable continuum model of water

Energy profile of the phosphoramidic mustard cyclization reaction
Reaction coordinate: distance N-C (constraint).
Comparison with explicit solvent simulation (70 water molecules)

The future: Numerical issues

- Discretization of PDEs
  - FE vs FD (compact higher order/Mehrstellen)
- Fast solvers (multigrid, etc.)
- Integration of stiff ODEs
- Coupling of models
  - QM-continuum interface
    - Poisson-Boltzmann solvers
  - QM-classical MD
  - DFT - QMC coupling
CS issues

- Establish data standards (mostly XML)
- Manage large datasets ($10^{10} - 10^{12}$ bytes)
- Data compression
- Parallel XML parsing
- Coupling to database SW
- Visualization
- Fault-tolerant SW
- Using variable partitions/process migration
- Identify the right parallel programming model for 65k CPUs
- Prepare for “architecture discontinuities”
Finding the right people: the ideal development team

- Applied mathematicians
  - Use the right model, numerical approach
- Computer scientists
  - Use computers efficiently
- Physicists
  - Build the right product
The “Simulation Facility” analogy

Advanced Photon Source, ANL

Synchrotron Beamline infrastructure

ASCi White, LLNL

Building the “computational beamlines” of future DOE supercomputers