MOLECULAR INTEGRATION SIMULATION TOOLKIT - INTERFACING NOVEL INTEGRATORS WITH MOLECULAR DYNAMICS CODES

Iain Bethune  (ibethune@epcc.ed.ac.uk, @ibethune)

Elena Breitmoser
Ben J Leimkuhler
Advertisement Interlude!

CP2K Summer School
23-26 August 2016
King’s College London

Register at:
https://www.epcc.ed.ac.uk/cp2k-summer-school-2016
The ExTASY project

- UK-US Collaborative research project funded by NSF (US) & EPSRC (UK)
  - 2013 – 2017
  - Partners: Rice, Rutgers, Duke, Edinburgh, Nottingham, Imperial

Grand Challenges in the Chemical Sciences

“to enhance our ability to understand the behaviour and function of complex macromolecules such as proteins, DNA, and other biomolecules through sampling with molecular dynamics (MD) simulations.”

http://www.extasy-project.org
The ExTASY project

- **More** sampling via more simulations
  - Hundreds or thousands of concurrent MD jobs
  - Manage execution, data movement, efficient HPC resource utilisation

**Ensemble Toolkit**

- **Better** sampling via biased simulations
  - Don’t waste time sampling behaviour already observed
  - Drive systems towards unexplored regions
  - ... and still obtain true thermodynamic distribution of states

- **Faster** sampling via better algorithms
  - Couple the latest integration algorithms with existing MD codes
  - Increased timestep without loss of accuracy

CoCo-MD, DM-d-MD

MIST library
Molecular Dynamics Software

• Many highly-tuned MD codes developed
  • GROMACS, AMBER, LAMMPS, NAMD …
  • 1000s of person-years of effort
  • Support for multicore desktop, GPUs, HPC, custom hardware
• USPs
  • force-fields, special features (REMD, TAMD …), optimised for speed or scalability, file formats…

• Leads to code complexity!
  • New developments mainly by core developer groups
  • Require knowledge of parallelisation, low level optimisations, internal data structures
Molecular Dynamics Software

• Result:
  • Community stuck with small number of widely implemented algorithms:
  • Verlet / Leapfrog + Berendsen / Nosé-Hoover Thermostat, Barostats, Langevin Dynamics, SHAKE/RATTLE/LINCS
  • Force-biasing algorithms e.g. Metadynamics (PLUMED)

• What is missing? – examples:
  • Langevin BAOAB (Leimkuhler & Matthews, JCP, 2013)
    • Stable at 25% longer timesteps, 10x smaller KE, PE error
  • DLM rotational integration (Dullweber et al, JCP, 1997)
    • Symplectic, time-reversible -> long term stable CG-MD
Bridging the implementation gap

- High-level abstraction of system state
  - Allow direct modification of state variables (positions, momenta…)
  - Clean & simple API -> easy to write new integrators
  - No knowledge of parallelisation required
  - Independent of a particular MD code -> portability

- Plug-in to existing MD codes
  - Use existing highly-tuned force evaluation code
  - System setup and output in well-known formats
  - Minimise loss of performance due to abstraction (overhead)
  - Ability to test algorithms on production-scale systems
MIST: Molecular Integration Simulation Toolkit

- C++ library
- Shared Memory
- C / Fortran interface
- Open Source (BSD licence)
- No external dependencies
void VerletIntegrator::Step(double dt)
{
    VelocityStep(0.5 * dt);
    PositionStep(dt);
    system->UpdateForces();
    VelocityStep(0.5 * dt);
}
void VerletIntegrator::Step(double dt) {
    VelocityStep(0.5 * dt);

    constraintSolver->StorePositions();

    PositionStep(dt);

    ResolveConstraints_pos(dt);

    system->UpdateForces();

    VelocityStep(0.5 * dt);

    ResolveConstraints_vel(dt);
}
void ContinuousTempering::Step(double dt)
{
    ... 
    for (i = 0; i < system->GetNumParticles(); i++)
    {
        m_inv = system->GetInverseMass(i);
        v = system->GetVelocity(i);
        f = system->GetForce(i);
        f = Vector3::Scale(1 - coupl, f);
        v = v + Vector3::Scale(dt * 0.5 * m_inv, f);
        system->SetVelocity(i, v);
    }
    ...
}
MIST: Molecular Integration Simulation Toolkit

- 12,000 atoms
- Verlet NVE
- CHARMM22 all-atom flexible
- Gromacs 5.0.2
- Intel Ivy Bridge (Cray XC30)
- Overhead < 1%
MIST: Molecular Integration Simulation Toolkit

- Intel Sandy Bridge
- Nvidia K20x GPU
- Gromacs 5.0.2
- Overhead 1-5%
MIST: Molecular Integration Simulation Toolkit

- Water box in AMBER 14
- Stable integration
- MIST-leapfrog trajectory analytically identical
Application


\[ \hat{H}(q,p,\xi,p_\xi) = H(q,p) - f(\xi)U(q) + \frac{p_\xi^2}{2m_\xi} + \phi(\xi) \]

- Extended system, \( \xi \) acts as an effective temperature
- When \( f(\xi) = 0 \), recover unperturbed system
- When \( f(\xi) > 0 \), higher temperature promotes phase space exploration
- Perform metadynamics on \( \Phi(\xi) \), integrate with Langevin BAOAB algorithm
Application

- Deca-alanine
- Using NAMD-Lite + MIST
- Many transitions observed in 400ns simulation (black points)
- Canonical distribution of configurations at 300K (red points) used to construct free energy profile
MIST: Roadmap

• Improved constraint solver

• Support for multiple timestepping (RESPA etc.)

• MPI Parallelisation
  • Support for domain decomposition in Gromacs
  • Plug-in for NAMD

• New integrators
  • Boxed MD (Booth *et al*, *Phil. Trans. Roy. Soc. A*, 2014)
  • Force-biasing from on-the-fly Diffusion Maps (Clementi Group)
Summary

- MIST library freely available from:
  - http://www.extasy-project.org/mist
  - Support for Gromacs, AMBER, NAMD-Lite
  - 8 integrators currently implemented

- Try out existing algorithms…
- … or implement your own

- Your feedback is very welcome!
Questions?