MOLECULAR INTEGRATION SIMULATION TOOLKIT -

INTERFACING NOVEL INTEGRATORS WITH MOLECULAR DYNAMICS CODES

Iain Bethune ibethune@epcc.ed.ac.uk
@ibethune #CCP5Conference

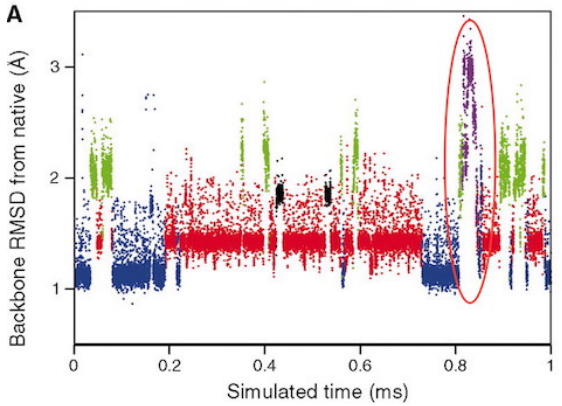
Elena Breitmoser
Ben J Leimkuhler

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

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 bio-molecules through sampling with molecular dynamics (MD) simulations."

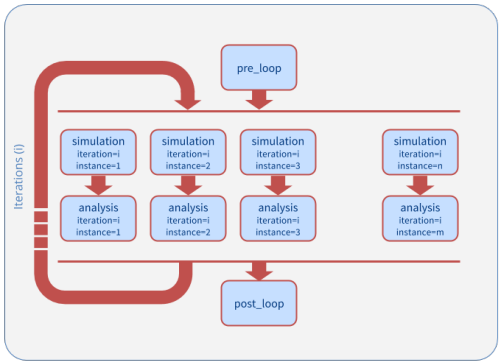
http://www.extasy-project.org

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

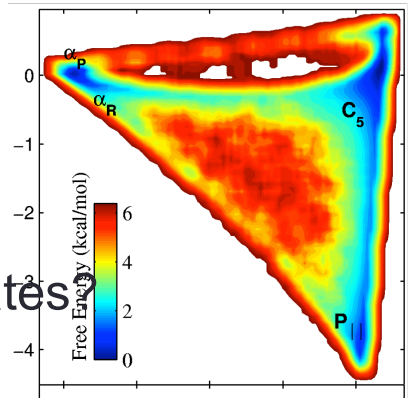
THE UNIVERSITY OF EDINBURGH

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?

CoCo-MD, DM-d-MD

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

MIST library

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

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

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

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

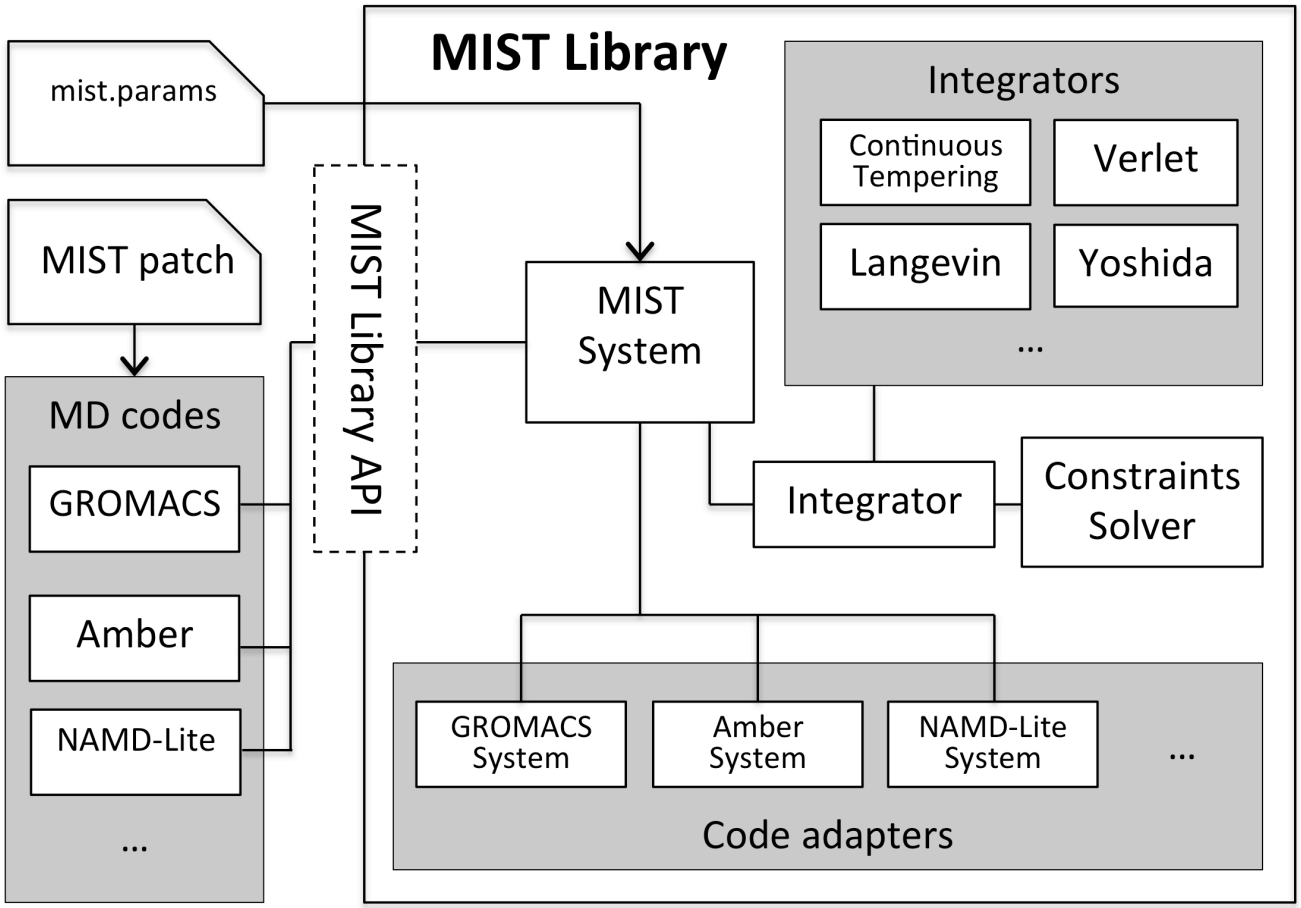
- Plug-in to existing MD codes
  - Use existing highly-tuned force evaluation code
  - System setup and output in well-known formats
  - Minimise loss of performance due to abstraction (overhead)
  - Ability to test algorithms on production-scale systems

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

MIST: Molecular Integration Simulation Toolkit



- C++ library
- Shared Memory
- C / Fortran interface
- Open Source (BSD licence)
- No external dependencies

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

MIST: Molecular Integration Simulation Toolkit

```
void VerletIntegrator::Step(double dt)
{
    VelocityStep(0.5 * dt);

    PositionStep(dt);

    system->UpdateForces();

    VelocityStep(0.5 * dt);

}
```

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

MIST: Molecular Integration Simulation Toolkit

```
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);
}
```

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

MIST: Molecular Integration Simulation Toolkit

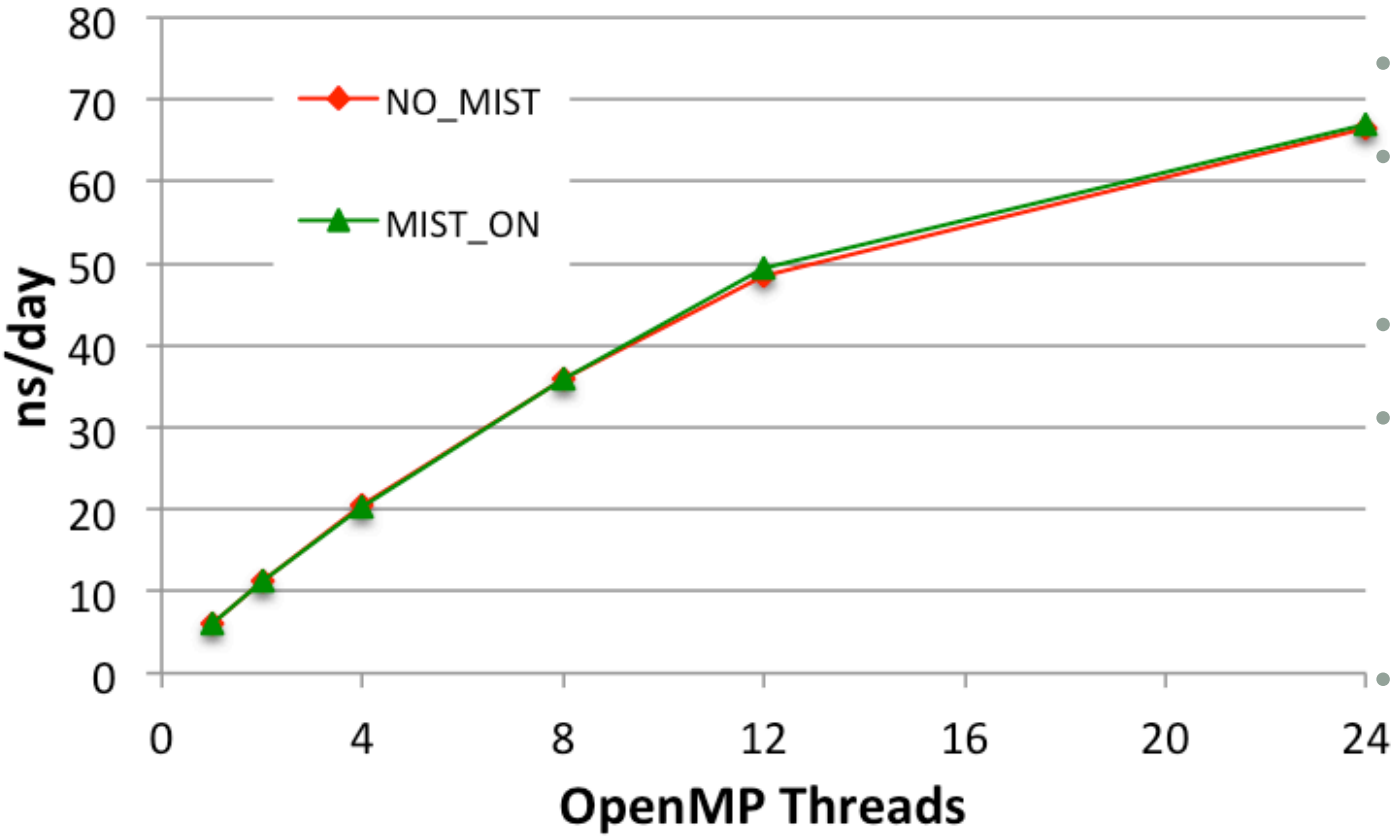
```
void ContinuousTempering::Step(double dt)
{
    ...
    for (i = 0; i < system->GetNumParticles(); i++)
    {
        m_inv = system->GetInverseMass(i);
        v = system->GetVelocity(i);
        f = system->GetForce(i);
        f = Vector3::Scale(1 - coupl, f);
        v = v + Vector3::Scale(dt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

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

MIST: Molecular Integration Simulation Toolkit



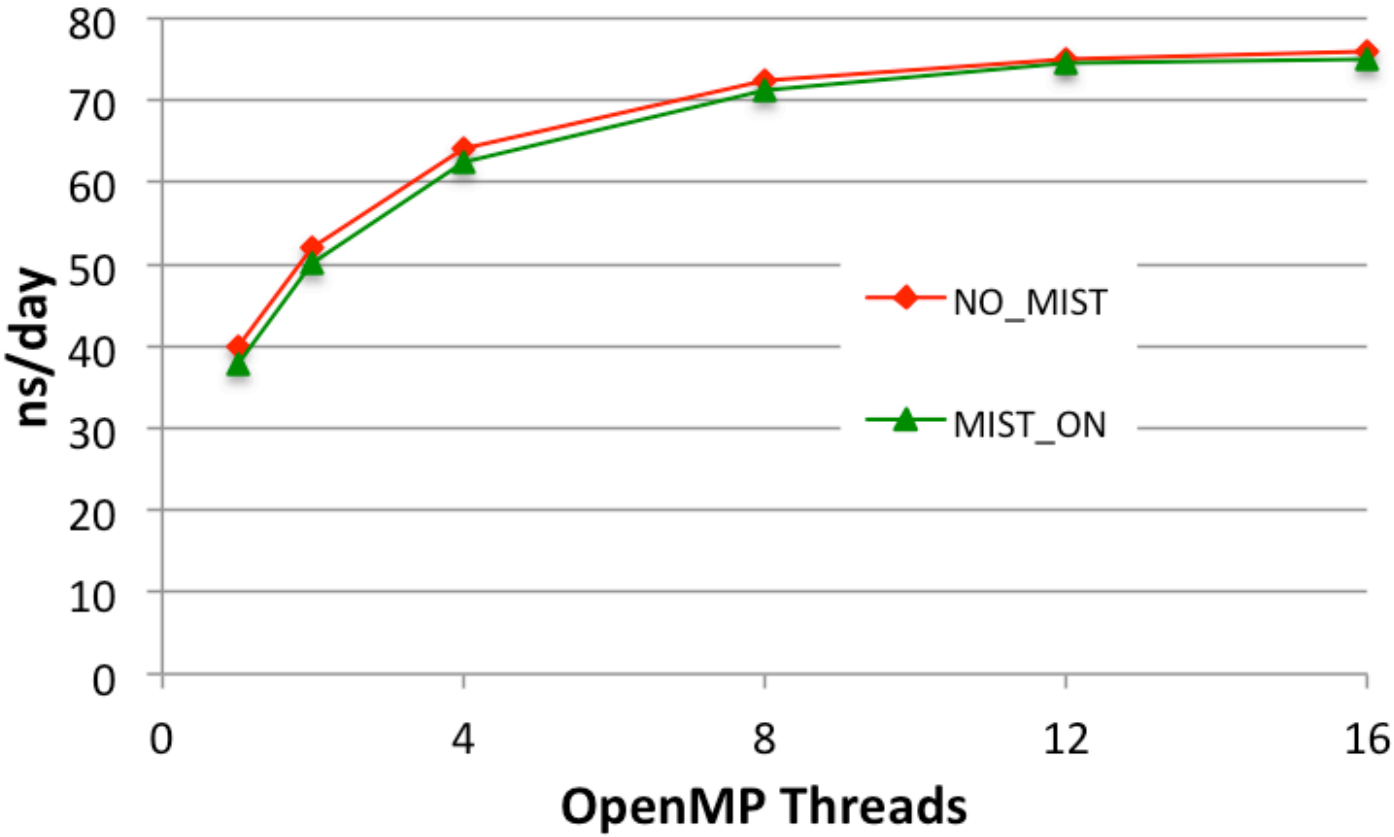
- 12,000 atoms
- Verlet NVE
- CHARMM22 all-atom flexible
- Gromacs 5.0.2
- Intel Ivy Bridge (Cray XC30)
- Overhead < 1%

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

MIST: Molecular Integration Simulation Toolkit



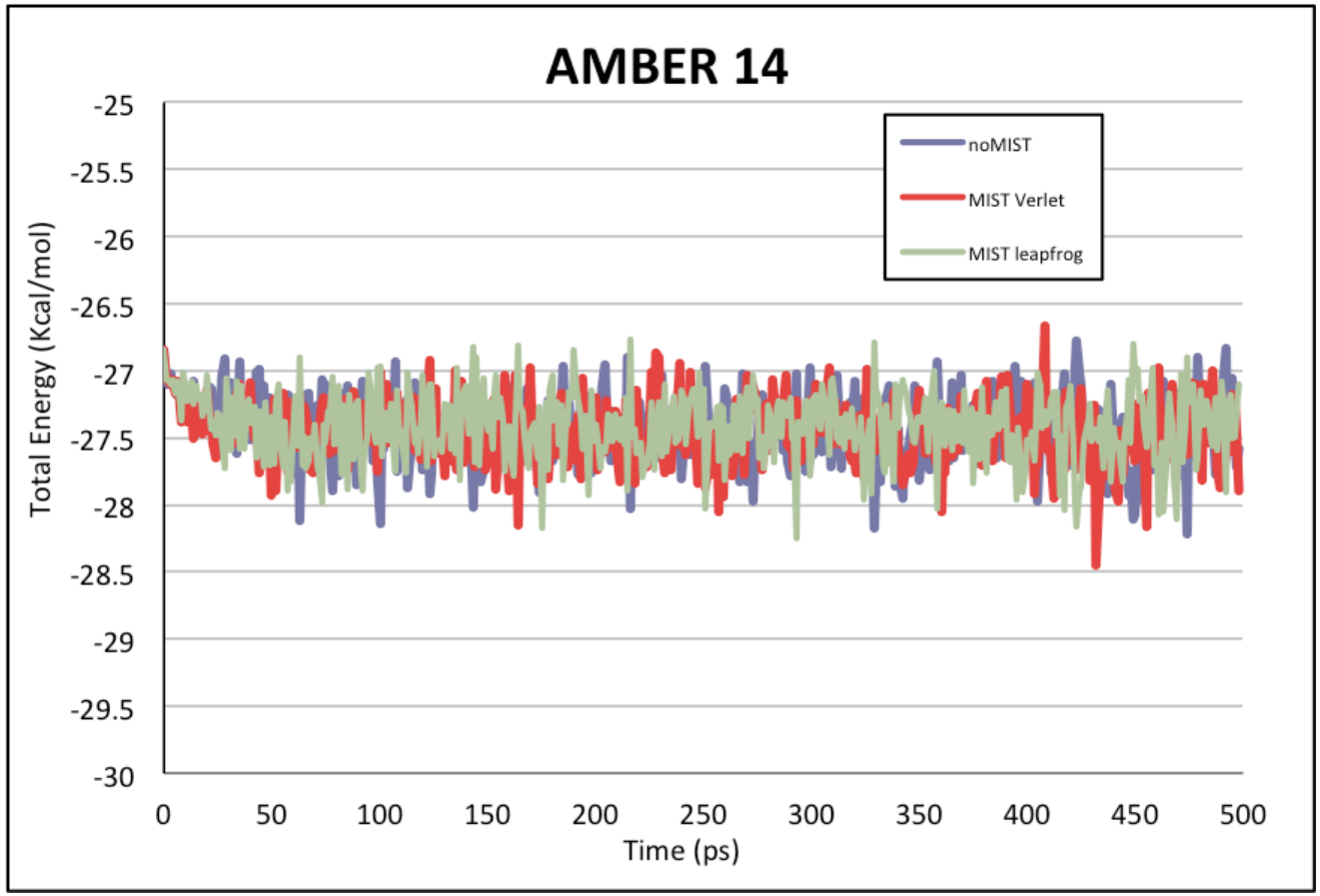
- Intel Sandy Bridge
- Nvidia K20x GPU
- Gromacs 5.0.2
- Overhead 1-5%

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

MIST: Molecular Integration Simulation Toolkit



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

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

Application

- Continuous Tempering (Gobbo & Leimkuhler, *Phys. Rev. E*, 2015)

$$\hat{\mathrm{H}}(\mathrm{q},\mathrm{p},\xi,\mathrm{p}_{\xi}) = H(q,p) - f(\xi)U(q) + p_{\xi}^{2} / 2m_{\xi} + \phi(\xi)$$

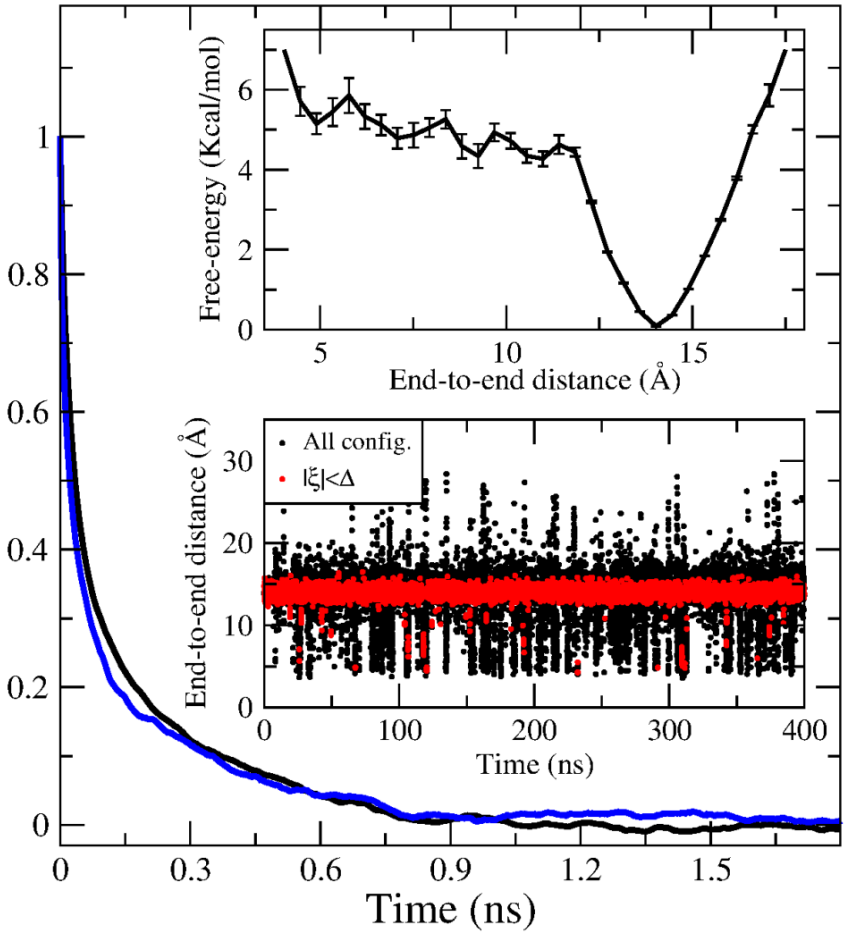
- Extended system, ξ acts as an effective temperature
- When f(ξ) = 0, recover unperturbed system
- When f(ξ) > 0, higher temperature promotes phase space exploration
- Perform metadynamics on Φ(ξ), integrate with Langevin BAOAB algorithm

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

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

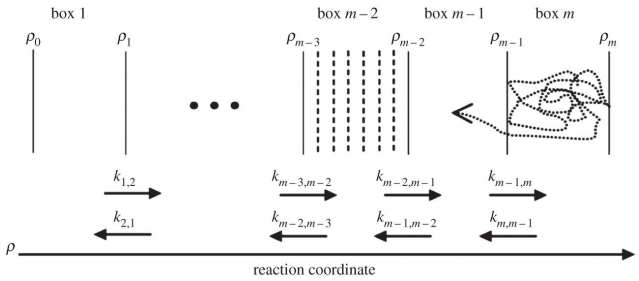
ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

MIST: Roadmap

- Improved constraint solver
  - Symmetric Newton iterative method (see Leimkuhler & Matthews, *Proc. Roy. Soc. A*, 2016)
- 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)
  - Simulated Tempering (Nguyen *et al, JCP*, 2013)



ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

The University of Edinburgh

Summary

- MIST library freely available:
  - 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!

ExTASY
Extensible Tools for Advanced Sampling and Analysis

epcc

THE UNIVERSITY OF EDINBURGH

Questions?