Estimating Accuracy in Classical Molecular Simulation

Stephen Bond

University of Illinois Urbana-Champaign
Department of Computer Science

Institute for Mathematics and its Applications
July 2007

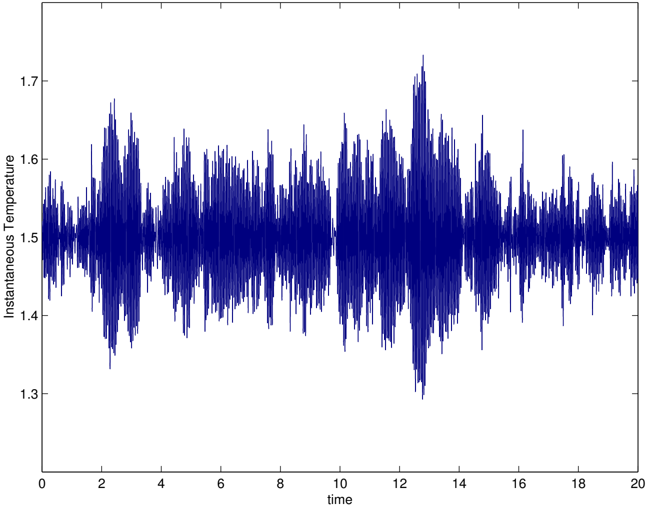
Acknowledgements

- Ben Leimkuhler (Edinburgh)
- Brian Laird (Kansas)
- Ruslan Davidchack (Leicester)
- Nana Arizumi (Illinois)

S. Bond and B. Leimkuhler *Acta Numerica* **16**, 2007.

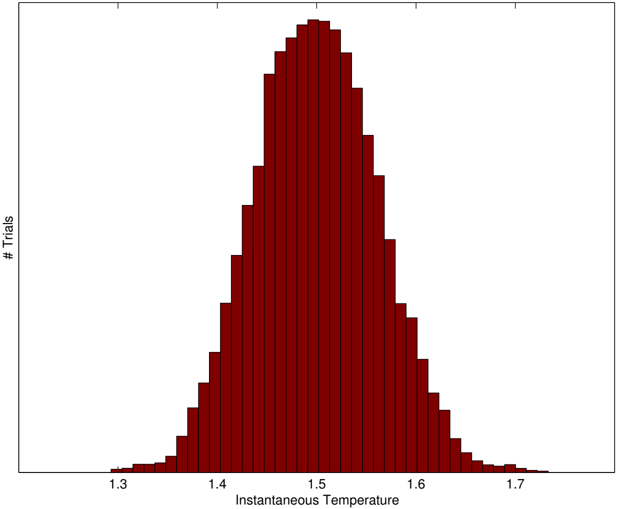
Motivation

Computation of Averages



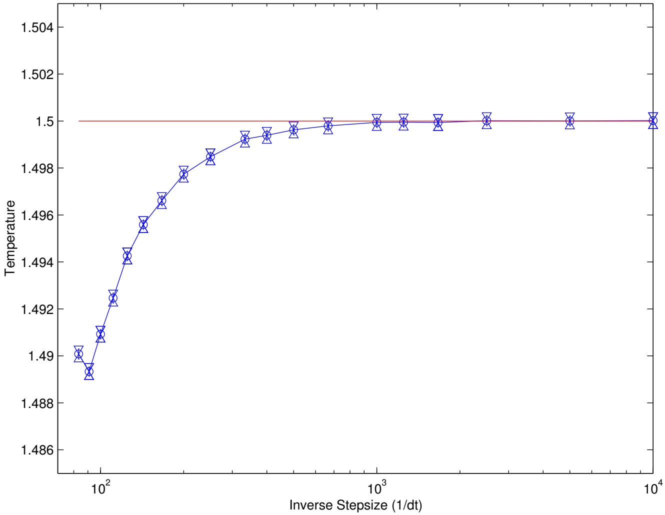
Motivation

Computation of Averages

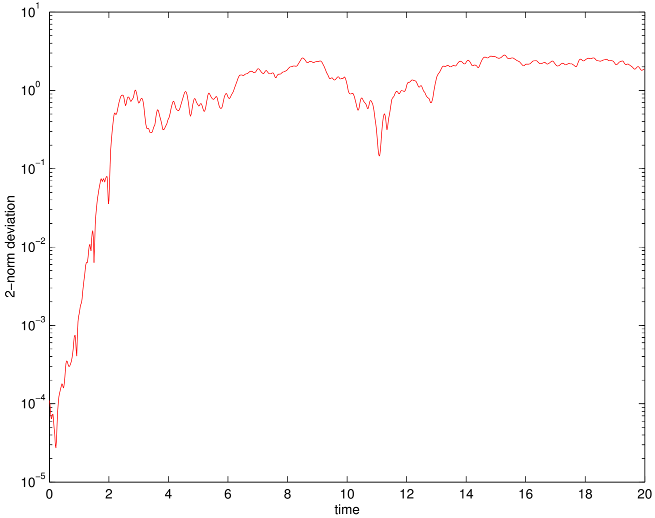


Motivation

Convergence of Averages



Divergence of Trajectories



Goal

- What is the error in an average from a MD trajectory?

$$\text{Error} = |\langle A \rangle_{\text{numerical}} - \langle A \rangle_{\text{exact}}|$$

- Any estimate must account for two factors:

$$\text{Error} \le \text{Statistical Error} + \text{Truncation Error}$$

- Asymptotic Bound:

$$\text{Error} \le C_1 \frac{1}{\sqrt{t}} + C_2 \, \Delta t^p$$

- Talk will focus on truncation error.

Paper on statistical error:

E. Cancès, F. Castella, P. Chartier, E. Faou, C. Le Bris, F. Legoll, G. Turinici, '04, '05.

System of Equations

- Newton's Equations: Force = mass × acceleration

$$\dot{q} = p/m \quad \text{and} \quad \dot{p} = -\nabla U(q)$$

q = position, m = mass, p = momenta

- First Order System

$$\dot{z} = F(z), \quad \text{where} \quad F : \mathbf{R}^n \to \mathbf{R}^n$$

- Exact Solution Map

$$z(t) = \Phi_t(t_0, z_0)$$

Ergodic

- Time Average:

$$\langle A\rangle_{\mathrm{time}} = \lim_{t\to\infty} \frac{1}{t}\int_0^t A\left(z\left(\tau\right)\right)\mathrm{d}\tau$$

- Ensemble Average:

$$\langle A\rangle_{\mathrm{ensemble}} = \int_\Omega A\left(z\right)\rho\left(z\right)\mathrm{d}z$$

- Ergodicity

$$\langle A\rangle_{\mathrm{time}} = \langle A\rangle_{\mathrm{ensemble}} \qquad (a.e.)$$

Liouville Equation

- Continuity Equation for the Probability Density:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho F) = 0$$

- Generalized Liouville Equation:

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho \nabla \cdot F = 0$$

or

$$\frac{\mathrm{D} \ln \rho}{\mathrm{D}t} = -\nabla \cdot F$$

Example: Nosé-Hoover

- Nosé-Hoover Vector Field

$$\begin{aligned}\frac{dq}{dt} &= M^{-1}p\\ \frac{dp}{dt} &= -\nabla U(q) - \frac{\xi}{\mu}p\\ \frac{d\xi}{dt} &= p^T M^{-1} p - g k_B T\end{aligned}$$

- Invariant Distribution

$$\rho \propto \exp\left\{-\frac{1}{k_B T}\left(\frac{1}{2}p^T M^{-1} p + U(q) + \frac{\xi^2}{2\mu}\right)\right\}$$

Error Analysis

- First Order System

$$\dot{z} = F(z)$$

- Forward Error:
  Is the numerical trajectory close to the exact trajectory?

$$\|\overline{z}_{\Delta t}(t) - z(t)\| \leq C\Delta t^p$$

- Backward Error:
  Is the numerical trajectory actually an exact trajectory, but for a different problem?

$$\|\overline{F}_{\Delta t}(z) - F(z)\| \leq C\Delta t^p$$

"Method of Modified Equations"

Error Analysis

- Ergodicity:
  Exact trajectories are sensitive (chaotic) to perturbations in the initial conditions
  $\longrightarrow$ Large Forward Error.

- Statistics:
  Thermodynamic properties (averages) are not a function of the details of the initial conditions
  $\longrightarrow$ Small Backward Error.

Backward Error Analysis: Modified Equations

- Given a pth-order numerical method, Ψ, we can always construct a modified vector field, $F_{\Delta t}$, such that the numerical method provides a qth-order approximation to the flow of the modified system.
- If the numerical method and vector field are time-reversible (symplectic/Hamiltonian), the modified vector field will be time-reversible (symplectic/Hamiltonian).
- Unfortunately, even if the vector field is analytic, the modified vector field does not converge as $q \to \infty$.
- Fortunately, it is still useful as a truncated series.

Big Picture

Flow Map	$\Phi_t \quad \approx \quad \overline{\Psi}_{\Delta t}$
$\uparrow$	$\uparrow \quad \nearrow \quad \downarrow$
Vector Field	$F \quad \approx \quad \overline{F}_{\Delta t}$
$\downarrow$	$\downarrow \qquad\quad \downarrow$
Ensemble	$\rho \quad \approx \quad \overline{\rho}_{\Delta t}$

Example: Verlet

- Hamiltonian

$$H(q,p) = \frac{1}{2} p^T M^{-1} p + U(q)$$

- Verlet

$$\begin{aligned} p^{n+1/2} &= p^n - \frac{\Delta t}{2} \nabla U(q^n) \\ q^{n+1} &= q^n + \Delta t M^{-1} p^{n+1/2} \\ p^{n+1} &= p^{n+1/2} - \frac{\Delta t}{2} \nabla U\left(q^{n+1}\right) \end{aligned}$$

- Splitting

$$H_1 = \frac{1}{2} p^T M^{-1} p, \qquad H_2 = U(q)$$

Example: Verlet

- Strang Splitting

$$\exp\left(\Delta t \mathcal{L}\right) = \exp\left(\frac{\Delta t}{2}\mathcal{L}_2\right)\exp\left(\Delta t \mathcal{L}_1\right)\exp\left(\frac{\Delta t}{2}\mathcal{L}_2\right) + \mathcal{O}\left[\Delta t^3\right]$$

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$$

$$\mathcal{L}_1 = M^{-1} p \cdot \nabla_q \qquad \mathcal{L}_2 = -\nabla_q U\left(q\right) \cdot \nabla_p$$

- Modified Equations

$$\exp\left(\Delta t \overline{\mathcal{L}}_{\Delta t}^{[r]}\right) = \exp\left(\frac{\Delta t}{2}\mathcal{L}_2\right)\exp\left(\Delta t \mathcal{L}_1\right)\exp\left(\frac{\Delta t}{2}\mathcal{L}_2\right) + \mathcal{O}\left[\Delta t^{r+1}\right]$$

Solve for $\overline{\mathcal{L}}_{\Delta t}^{[r]}$ using Baker-Campbell-Hausdorff formula

Example: Verlet

- Original Hamiltonian:

$$H(q,p) = \frac{1}{2} p^T M^{-1} p + U(q)$$

- Modified Hamiltonian:

$$H_{2,\Delta t}(q,p) = H(q,p) + \frac{\Delta t^2}{12}\left(p^T M^{-1} U'' M^{-1} p - \frac{1}{2} \nabla U^T M^{-1} \nabla U\right)$$

Verlet conserves $H_{2,\Delta t}$ to 4th order accuracy!

Example: Generalized Leapfrog

- Generalized Leapfrog

$$
\begin{aligned}
p^{n+1/2} &= p^n - \frac{\Delta t}{2}\nabla_q H\left(q^n, p^{n+1/2}\right)\\
q^{n+1/2} &= q^n + \frac{\Delta t}{2}\nabla_p H\left(q^n, p^{n+1/2}\right)\\
p^{n+1} &= p^{n+1/2} - \frac{\Delta t}{2}\nabla_q H\left(q^{n+1}, p^{n+1/2}\right)\\
q^{n+1} &= q^{n+1/2} + \frac{\Delta t}{2}\nabla_p H\left(q^{n+1}, p^{n+1/2}\right)
\end{aligned}
$$

- Generalized Leapfrog Modified Hamiltonian:

$$
\overline{H}_{\Delta t} = H + \frac{\Delta t^2}{24}\left(2H_{q_j q_k} H_{p_j} H_{p_k} + 2H_{q_j p_k} H_{p_j} H_{q_k} - H_{p_j p_k} H_{q_j} H_{q_k}\right)
$$

Liouville Equation for Modified Vector Field

- Modified Equations

$$\frac{d\overline{z}}{dt} = \overline{F}_{\Delta t}(\overline{z}) \qquad \text{where} \quad \overline{F}_{\Delta t} = F + \Delta t^p G$$

- Modified Liouville Equation

$$\frac{\overline{D}}{\overline{D}t}\overline{\rho}_{\Delta t} = -\overline{\rho}_{\Delta t}\nabla \cdot \overline{F}_{\Delta t}$$

- Weighting factor

$$\omega_{\Delta t} := \overline{\rho}_{\Delta t}/\rho, \quad \text{assuming} \quad \rho, \overline{\rho}_{\Delta t} > 0$$

implies

$$\frac{\overline{D}}{\overline{D}t}\ln(\omega_{\Delta t}) = -\Delta t^p (\nabla \cdot G + G \cdot \nabla \ln \rho)$$

Averages

- Truncation Error Estimate

$$\begin{aligned}\langle A\rangle_{\mathrm{Num}} - \langle A\rangle_{\mathrm{Exact}} &\approx \int_\Gamma A\,(q,p)\,\rho_{\Delta t}\,\mathrm{d}\Gamma - \int_\Gamma A\,(q,p)\,\rho\,\mathrm{d}\Gamma \\ &\approx \frac{\langle A\rangle_{\mathrm{Num}}\langle 1/\omega_{\Delta t}\rangle_{\mathrm{Num}} - \langle A/\omega_{\Delta t}\rangle_{\mathrm{Num}}}{\langle 1/\omega_{\Delta t}\rangle_{\mathrm{Num}}}\end{aligned}$$

- Reweighted Averages

$$\langle A\rangle_{\mathrm{Exact}} = \frac{\langle A/\omega_{\Delta t}\rangle_{\mathrm{Num}}}{\langle 1/\omega_{\Delta t}\rangle_{\mathrm{Num}}} + \mathcal{O}\left[\Delta t^r\right]$$

Example:

- Nosé-Poincaré Hamiltonian:

$$H(q, \tilde{p}, s, \pi_s) = s\left(\frac{1}{2\,s^2}\tilde{p}^T M^{-1}\tilde{p} + U(q) + \frac{\pi_s^2}{2\mu} + g\,k\,T\,\ln s - E_0\right)$$

- Nosé-Poincaré Modified Hamiltonian:

$$\begin{aligned}
\overline{H}_{\Delta t} &= H_{NP} + \frac{\Delta t^2}{12} s\left(\frac{\pi_s}{\mu\,s}\tilde{p}^T M^{-1}\nabla U\right. \\
&- \frac{1}{2}\nabla U^T M^{-1}\nabla U + \frac{1}{s^2}\tilde{p}^T M^{-1} U'' M^{-1}\tilde{p} \\
&- \left.\frac{1}{2\,\mu}\left(\frac{1}{s^2}\tilde{p}^T M^{-1}\tilde{p} - g\,k\,T\right)^2 + \frac{2\,g\,k\,T\,\pi_s^2}{\mu^2}\right)
\end{aligned}$$

Example:

- Modified marginal distribution:

$$\bar{\rho}_{\Delta t}(q,p)\,\mathrm{d}p\,\mathrm{d}q = \frac{1}{C}\int_s\int_{p_s}\delta\left[\overline{H}_{\Delta t}(q,s,\tilde{p},p_s)-\overline{E}_0\right]\mathrm{d}\tilde{p}\,\mathrm{d}q\,\mathrm{d}p_s\,\mathrm{d}s,$$

$$= \frac{1}{C}\int_s\int_{p_s}\delta\left[s\left(H_\mathrm{N}-H_\mathrm{N}^0+\Delta t^2 G\right)\right]\mathrm{d}\tilde{p}\,\mathrm{d}q\,\mathrm{d}p_s\,\mathrm{d}s.$$

- Change of variables, integrating

$$\bar{\rho} = \frac{1}{C}\int_{p_s} \mathrm{e}^{N_f\eta_0}\left|gk_BT+h^2\frac{\partial}{\partial\eta}G(q,\mathrm{e}^{\eta},p,p_s)\right|_{\eta=\eta_0}^{-1}\mathrm{d}p_s.$$

$$\eta_0 = \frac{-1}{g\,k_B\,T}\left(H(q,p)+\frac{p_s^2}{2\,\mu}+h^2\,G(q,\mathrm{e}^{\eta_0},p,p_s)-H_\mathrm{N}^0\right),$$

- More mathematical manipulations

$$\bar{\rho} = \frac{\rho_c}{\bar{C}}\exp\left\{-\frac{\Delta t^2}{24k_BT}\left[\sum_{j,k}\frac{2p_jp_kU_{q_jq_k}}{m_jm_k}-\sum_j\frac{U_{q_j}^2}{m_j}-\frac{1}{\mu}\left(\sum_j\frac{p_j^2}{m_j}-gk_BT\right)^2\right]\right\},$$

Example:

- Weighting Factor:

$$\omega_{\Delta t} \approx \exp\left\{\frac{-\Delta t^2}{24\, k_{\rm B}\, T}\left[2p^T M^{-1} U''(q)\, M^{-1} p - \nabla U(q)^T\, M^{-1} \nabla U(q) - \frac{1}{\mu}\left(p^T M^{-1} p - g\, k_{\rm B}\, T\right)^2\right]\right\}$$

- Reweighted Averages:

$$\langle A\rangle_{\rm Exact} \approx \frac{\langle A/\omega_{\Delta t}\rangle_{\rm Num}}{\langle 1/\omega_{\Delta t}\rangle_{\rm Num}}$$

- Hybrid Monte Carlo:
  J. Izaguirre and S. Hampton, *J. Comput. Phys.* **200**, 2004.
  E. Akhmatskaya and S. Reich, *LNCSE* **49**, 2006.

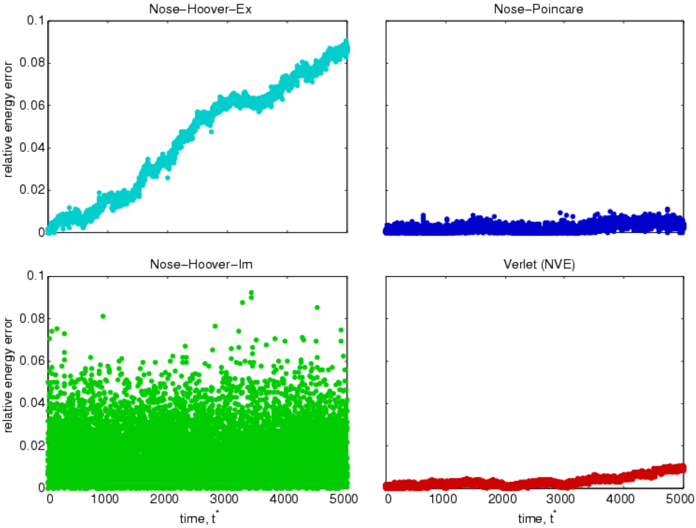
- Time correlation functions:
  R. D. Skeel, Preprint, 2007.

Numerical Experiment:

- System:
  - 256 Particle Gas
  - Lennard-Jones Potential
  - $T = 1.5\epsilon/k$, $\rho = 0.95r_0^3$, $t = 20r_0\sqrt{m/\epsilon}$
- Method:
  - Nosé-Poincaré (Symplectic, Time-Reversible)
  - $\Delta t = 0.012r_0\sqrt{m/\epsilon}$ to $0.0001r_0\sqrt{m/\epsilon}$
- Reference:
  - Bond, Laird, and Leimkuhler *J. Comput. Phys.* **151** 1999.
  - S. Bond and B. Leimkuhler *Acta Numerica* **16**, 2007.

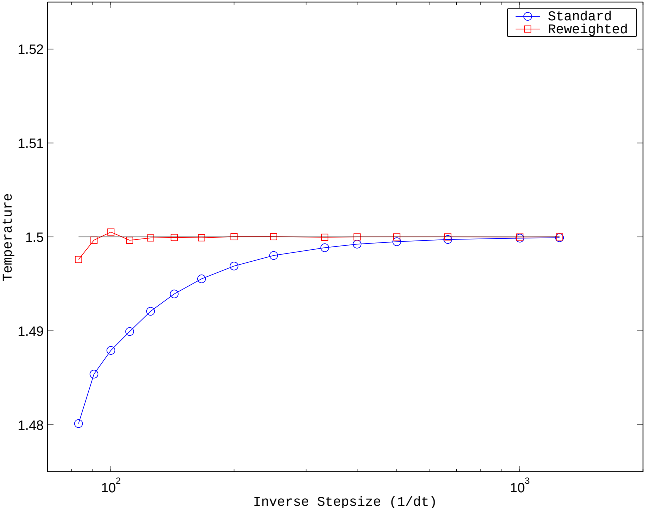
Numerical Experiment:

- "Extended" Energy Conservation:

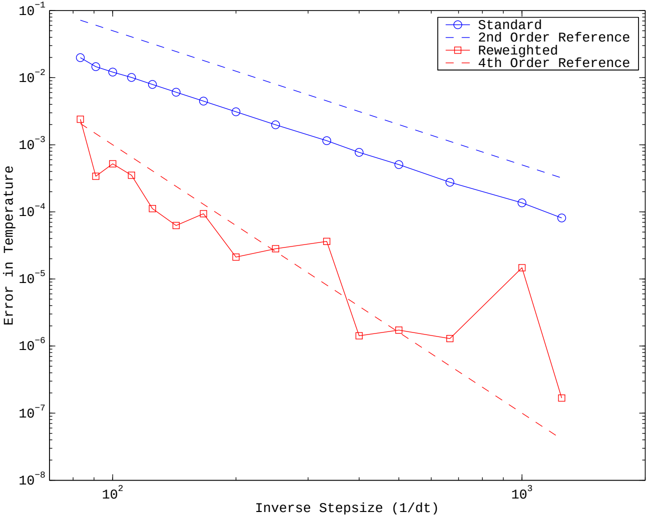


Numerical Experiment:

- Improved Estimator



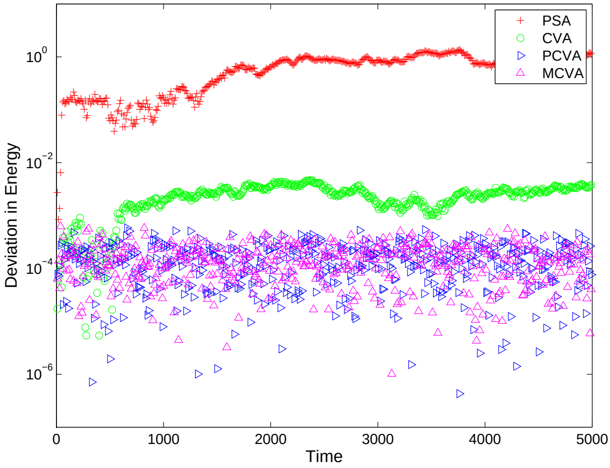
- Improved Estimator Error

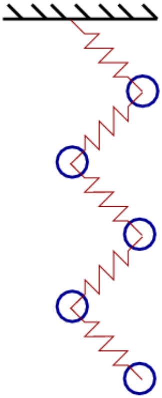


Future Directions:

- Further testing with more systems and averages
- Extensions to reduce computational cost
- Other ensembles and numerical methods

Stabilizing Hard Sphere Algorithms:





S. Bond and B. Leimkuhler, SIAM J. Sci. Comput, 2007, In press.