Multilevel Monte Carlo for the simulation of dilute polymers

Mike Giles

Oxford University Mathematical Institute

Joint work with Endre Süli, James Whittle and Shenghan Ye

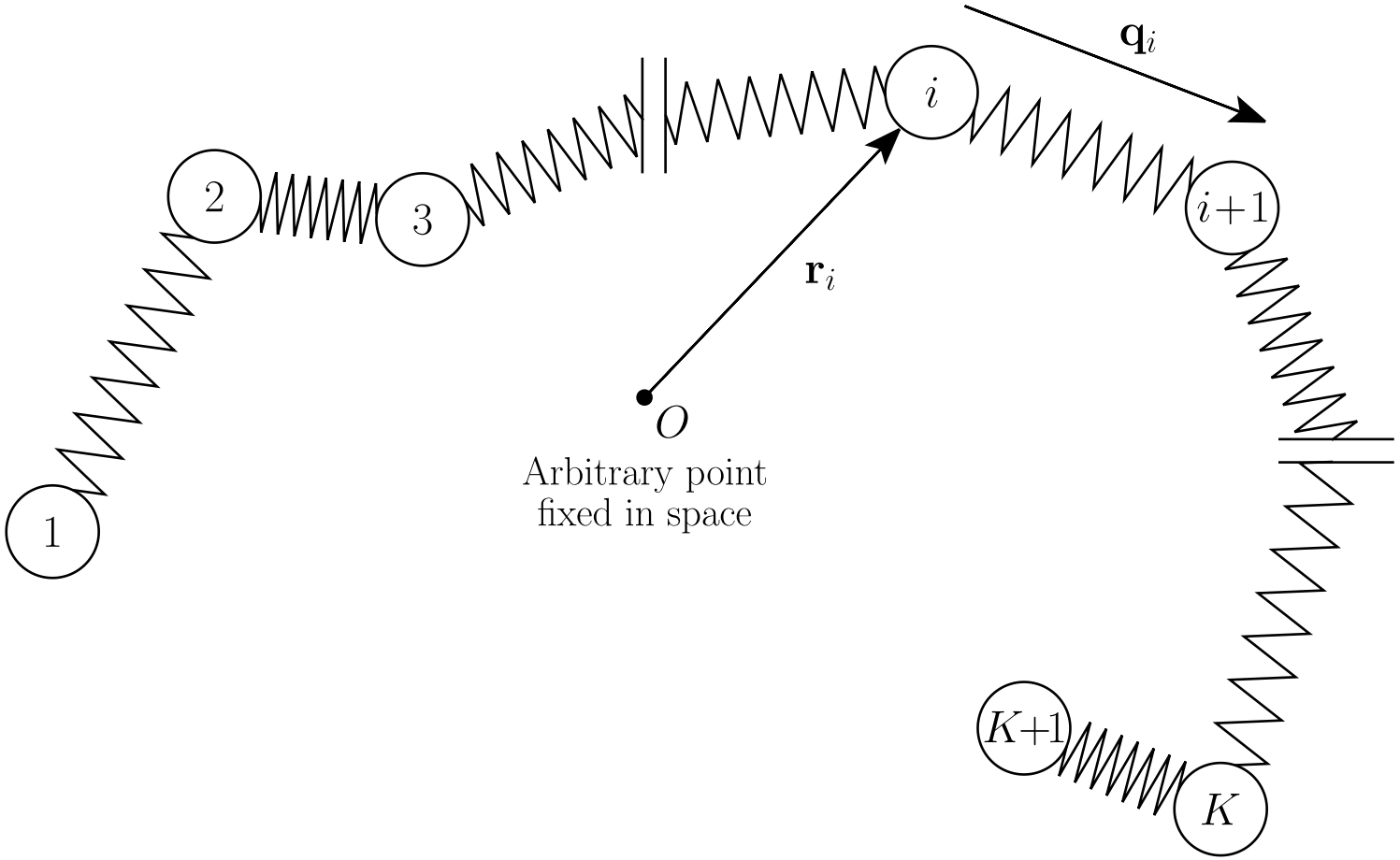
London Probability Seminar

April 24, 2015

Outline

- mathematical modelling
- properties of the SDE
- numerical approximation
- multilevel Monte Carlo method
- adaptive time-stepping in MLMC
- MLMC for equilibrium distribution
- preliminary results
- conclusions

Modelling



- long-chain molecules in a fluid
- modelled as ball-and-spring systems, subject to random forcing
- K bonds, $K+1$ "balls", separation q_i will be key variable

Modelling

- FENE (finitely extensible nonlinear elastic) model limits extension of molecular bonds
- motion of "balls" given by force balance:

$$\text{elastic force} + \text{random force} + \text{viscous drag} = 0$$

$$-\nabla V \quad + \quad R \quad - \quad k\,(\dot{r}_i - v(r_i)) \quad = \quad 0$$

where V is the elastic potential, and v is the velocity of the fluid

- shifting to a moving frame of reference, a local Taylor series expansion gives

$$v(x) \approx \kappa\, x$$

where κ is the local rate-of-strain tensor $\partial v/\partial x$

Modelling

This modelling leads to the following SDE for i^{th} "ball":

$$\mathrm{d}r_i = \left(\kappa\, r_i - \nabla_{r_i} V(r)\right)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_i$$

where $\mathrm{d}W_i$ is the Brownian forcing, assumed to be independent of the forcing of the others, and

$$V(r) = \sum_{i=1}^{K} U_i(\|q_i\|^2/2)$$

with U_i being the elastic potential for the i^{th} bond.

Hence,

$$\mathrm{d}r_i = \left(\kappa\, r_i - \left(U_{i-1}'\left(\|q_{i-1}\|^2/2\right)\, q_{i-1} - U_i'\left(\|q_i\|^2/2\right)\, q_i\right)\right)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_i$$

if we define $q_0 \equiv q_{K+1} \equiv 0$ to account for non-existent bonds on either end.

Modelling

Hence, the system of coupled SDEs for the bond vectors is

$$\mathrm{d}q_i = \left(\kappa\, q_i + U'_{i+1} q_{i+1} - 2\, U'_i q_i + U'_{i-1} q_{i-1}\right)\,\mathrm{d}t + \sqrt{2}\,\left(\mathrm{d}W_{i+1} - \mathrm{d}W_i\right)$$

which can be written collectively as

$$\mathrm{d}q = (K\, q - D\,\nabla V)\,\mathrm{d}t + \sqrt{2}\; L\;\mathrm{d}W$$

where $V(q) \equiv \sum_i U_i(\|q_i\|^2/2)$, and K, L and D are of the form

$$K = \begin{pmatrix} \kappa & & \\ & \kappa & \\ & & \kappa \end{pmatrix}, \qquad L = \begin{pmatrix} -I & I & \\ & -I & I \\ & & -I & I \end{pmatrix},$$

$$D = \begin{pmatrix} 2I & -I & \\ -I & 2I & -I \\ & -I & 2I \end{pmatrix} = L\,L^T.$$

Invariant distribution

The Fokker-Planck PDE for the probability density function $p(q,t)$ is

$$\frac{\partial p}{\partial t} = \nabla \cdot \Big((D\,\nabla V - K\,q)\,p + D\,\nabla p \Big),$$

so when $K=0$ the invariant (or equilibrium) distribution is

$$p_\infty(q) = C\,\exp(-V(q)).$$

In the particular case of the FENE model, we have

$$U_i(s) = -\beta\,\log(1-2s),$$

and therefore the invariant distribution for $K=0$ is

$$p_\infty(q) = C\,\prod_i (1-\|q_i\|^2)^\beta.$$

Contraction property

If κ is not too large, can prove the following result:

If two paths $q^{(1)}$, $q^{(2)}$ have

- different initial conditions
- same driving Brownian motion W_t

then

$$\left\| q^{(1)}(t) - q^{(2)}(t) \right\| \quad \longrightarrow \quad 0$$

exponentially as $t \to 0$.

Numerical approximation

The nonlinear SDE is approximated as

$$q_{n+1} = q_n + (K\, q_n - D\, \nabla V(q_n))\, h_n + \sqrt{2}\; L\; \Delta W_n$$

using an adaptive timestep h_n, and independent Brownian increments ΔW_n which can be expressed as

$$\Delta W_n = \sqrt{h_n}\; Z_n$$

where Z_n are vectors of independent $N(0,1)$ random variables.

Adaptive timestep

With the FENE model, no bond length can exceed 1 – numerical approximation should share this property.

Try to ensure this through the restrictions:

$$\begin{aligned} h_n \, U_i'(\|q_{i,n}\|^2/2) \, \|q_{i,n}\| &\leq 1 - \|q_{i,n}\| \\ 5\sqrt{2\,h_n} &\leq 1 - \|q_{i,n}\| \end{aligned}$$

where $q_{i,n}$ is the i^{th} bond vector at timestep n.

They can be combined to give

$$h_n = \frac{\min_i (1 - \|q_{i,n}\|)^2}{\max(2\beta, 50)}$$

This sets an upper bound on the timestep – smaller values need to be chosen for accuracy.

Clamping

Even so, it is possible for the discrete approximation to cross the boundary $\|q_i\| = 1$.

This is avoided through "clamping" by setting

$$q_{i,n+1}^{clamped} := \frac{1-\delta}{\|q_{i,n+1}\|}\, q_{i,n+1}$$

if $\|q_{i,n+1}\| > 1-\delta$, where $\delta \ll 1$ (typically 10^{-5}).

We believe that the additional weak error due to clamping is negligible.

Objective

Our objective is to numerically estimate $\lim_{T\to\infty} \mathbb{E}\left[P(T)\right]$ where

$$P(T) \;\equiv\; U'(\|q\|^2/2)\, q\, q^T \Big|_{t=T}$$

This corresponds to the stress exerted by the molecule on the fluid. In the future this will be applied to the fluid in a coupled simulation.

We start with a computation for a fixed, large T, then address the challenge of letting $T \to \infty$.

Monte Carlo simulation

The standard Monte Carlo approach would be to perform N independent path simulations and use the estimate

$$N^{-1} \sum_{n=1}^{N} P^{(n)}$$

However, to get good accuracy we would need to use very small timesteps and lots of paths, so the cost would be high.

Multilevel Monte Carlo simulation

Instead, MLMC uses a sequence of levels with differing accuracies / costs, and relies on the telescoping summation

$$\mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{\ell=1}^{L} \mathbb{E}[P_\ell - P_{\ell-1}]$$

so we can use the estimator

$$N_0^{-1} \sum_{n=1}^{N_0} P_0^{(n)} \;+\; \sum_{\ell=1}^{L} \left\{ N_\ell^{-1} \sum_{n=1}^{N_\ell} \left(P_\ell^{(n)} - P_{\ell-1}^{(n)} \right) \right\}$$

with independent estimation for each level

Multilevel Monte Carlo simulation

Given a desired RMS accuracy ε, MLMC algorithm/theory tells us

- how many levels to use, to ensure that weak error due to discretisation is less than $\varepsilon/\sqrt{2}$
- how many samples to use on each level, to ensure that overall the statistical error is less than $\varepsilon/\sqrt{2}$

In standard SDE applications, use geometric sequence of levels with

$$h_\ell = 2^{-\ell}\, h_0$$

and for Milstein approximation (as used here) usually achieve ε accuracy at $O(\varepsilon^{-2})$ total cost, instead of $O(\varepsilon^{-3})$ with standard Monte Carlo method

Multilevel Monte Carlo simulation

Key point is using the same Brownian motion for each sample

$$P_\ell^{(n)} - P_{\ell-1}^{(n)}$$

in

$$N_\ell^{-1} \sum_{n=1}^{N_\ell} \left(P_\ell^{(n)} - P_{\ell-1}^{(n)} \right)$$

Strong convergence of the numerical approximation ensures that $P_\ell^{(n)} - P_{\ell-1}^{(n)}$ is small, so has a small variance, $O(h_\ell^2)$, and hence few samples are needed on finest levels.

With uniform timesteps, Brownian increments for coarse path are obtained by summing increments for fine path.

Potential problem: how does it work with adaptive time-stepping?

Multilevel Monte Carlo simulation

Actually, surprisingly easy, based on prior work with Chris Lester, Ruth Baker & Kit Yates (2014) for continuous-time Markov chains.

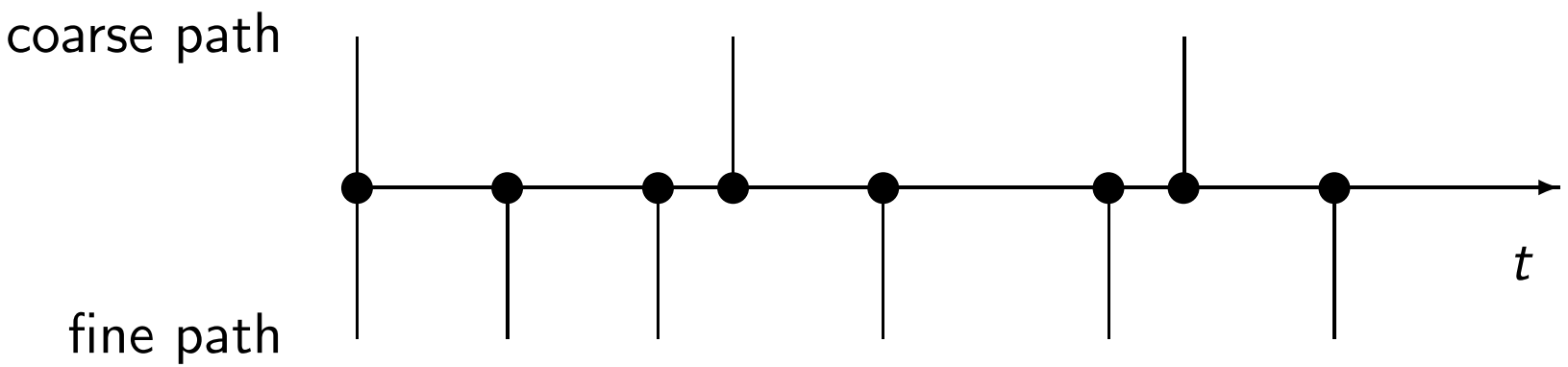
On level ℓ use

$$h_{\ell,n} = 2^{-\ell}\ \frac{\min_i(1-\|q_{i,n}\|)^2}{\max(2\beta, 50)}$$

Coarse and fine paths each compute their own adaptive timesteps independently – this ensures the telescoping sum works correctly

Multilevel Monte Carlo simulation

As time proceeds, Brownian increments are generated as needed at discretisation times which are a union of coarse and fine path times:



The fact that the timesteps are not nested is not a problem – strong convergence still ensures a strong coupling between the coarse and fine paths.

Multilevel Monte Carlo simulation

Final challenge: how to obtain expectation as $T \to \infty$?

Key idea here comes from Chang-han Rhee and Peter Glynn (2014) who consider contracting Markov chains:

$$X_0 = x, \quad X_{n+1} = \phi_n(X_n), \quad n \geq 0$$

where $\{\phi_n\}$ is a sequence of iid random functions such that

$$\sup_{x \neq y} \mathbb{E}\left[\left(\frac{d(\phi_n(x), \phi_n(y))}{d(x,y)}\right)^{2\gamma}\right] < 1$$

for some distance metric d, and some $\gamma \in (0,1)$.

Multilevel Monte Carlo simulation

They are interested in

$$\lim_{M\to\infty} \left\{ \mathbb{E}[f(X_M)] \mid X_0 = x \right\}$$

which can be re-expressed as

$$\lim_{M\to\infty} \left\{ \mathbb{E}[f(X_0)] \mid X_{-M} = x \right\}$$

and they use multilevel with $M_\ell \to \infty$ as $\ell \to \infty$ and same random ϕ_n for *coarse* and *fine* paths for $-M_{\ell-1} \le n < 0$.

This works because contraction property leads to effect of difference in values at $-M_{\ell-1}$ decaying exponentially, so

$$\left\| X_0^f - X_0^c \right\| \;\sim\; \exp(-c\, M_{\ell-1})$$

Multilevel Monte Carlo simulation

Back to our polymer application, instead of estimating

$$\lim_{T\to\infty} \Big\{ \mathbb{E}[P(q(T))] \mid q(0) = q_0 \Big\}$$

we use the same idea and estimate

$$\lim_{T\to\infty} \Big\{ \mathbb{E}[P(q(0))] \mid q(-T) = q_0 \Big\}$$

and use multilevel with $h_\ell \to 0$, $T_\ell \to \infty$ as $\ell \to \infty$ and the same Brownian motion $W(t)$ for coarse and fine paths for $-T_{\ell-1} < t < 0$.

This again works because of the contraction property which leads to effect of difference in q values at time $-T_{\ell-1}$ decaying exponentially, so

$$\Big\| q^f(0) - q^c(0) \Big\| \;=\; O(h_\ell) \;+\; O(\exp(-c\, T_{\ell-1}))$$

Implementation

Two versions of the MLMC algorithm have been implemented:

- MATLAB for development purposes
- CUDA C for execution on NVIDIA GPUs – offers 100× speedup over MATLAB implementation

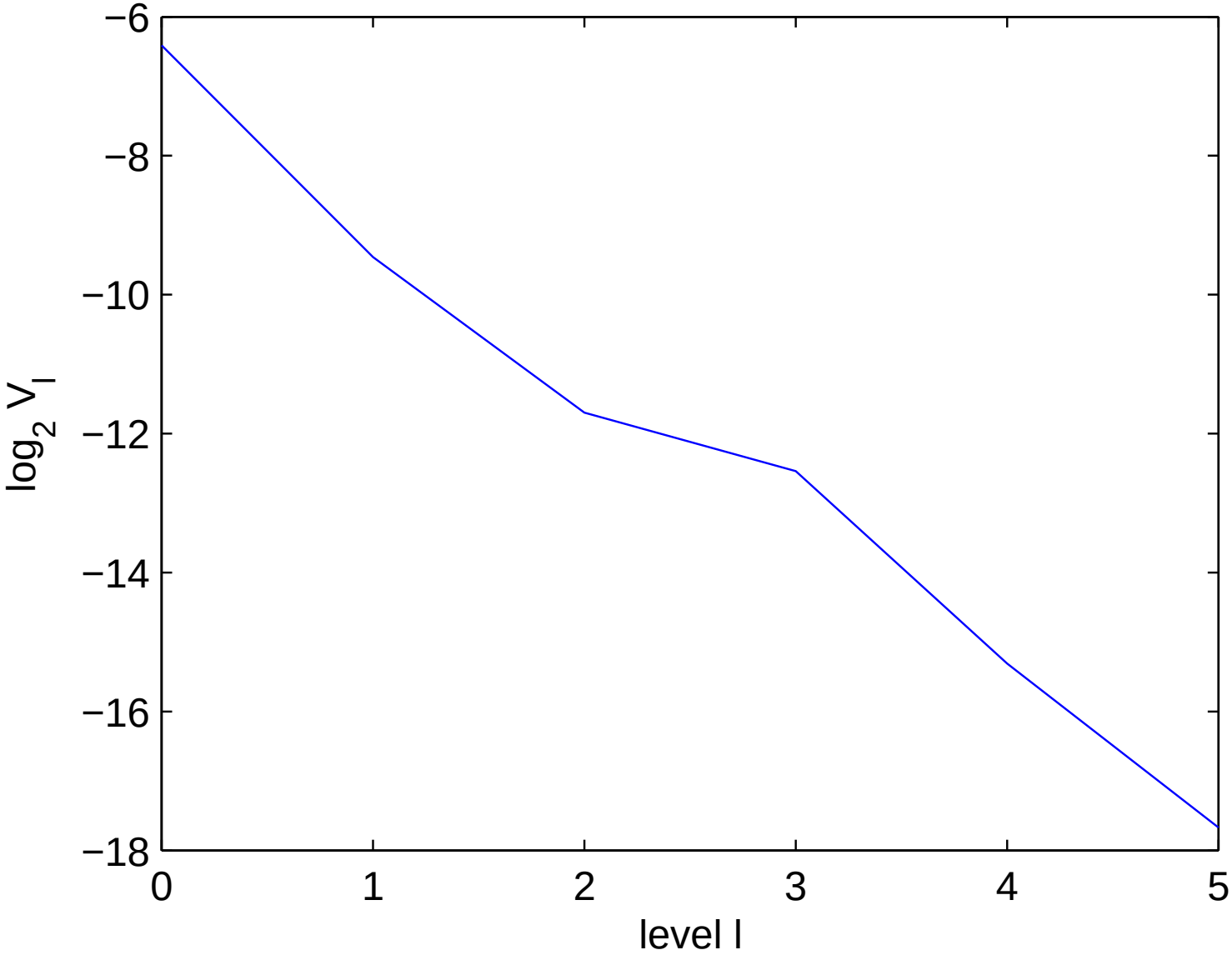
The latter handles polymers with up to 32 bonds

When there is a single bond, the code produces the correct value for

$$\lim_{T\to\infty} \mathbb{E}\left[q^{(0)\,T}\, q^{(0)} \,\Big|_{t=T}\right]$$

in comparison to the analytic value which can be deduced from the invariant distribution.

Preliminary results



Conclusions

- other research by Süli and Ye has used Fokker-Planck approach when there is just 1 bond, but SDE approach is only viable method when the number of bonds is large
- multilevel Monte Carlo greatly reduces the computational cost
- adaptive timestepping is necessary, and not difficult
- new idea due to Rhee and Glynn is crucial for estimating quantities associated with invariant distribution
- more validation required
- current research is investigating coupling to fluid calculation, and inclusion of history effects

References

John W. Barrett and Endre Süli. "Existence of global weak solutions to some regularized kinetic models for dilute polymers". *SIAM Multiscale Modelling and Simulation*, 6(2):506-546, 2007

M.B. Giles. "Multilevel Monte Carlo path simulation", *Operations Research*, 56(3):607-617, 2008

M.B. Giles. "Multilevel Monte Carlo methods", *Acta Numerica*, 2015

Chang-han Rhee and Peter W. Glynn. "Exact estimation for the equilibrium of Markov chains", *Journal of Applied Probability*, 51A:377-389, 2014

`http://people.maths.ox.ac.uk/gilesm/mlmc.html`

`http://people.maths.ox.ac.uk/gilesm/mlmc_community.html`