Parallel Flat Histogram Simulations

Malek O. Khan

Dept. of Physical Chemistry

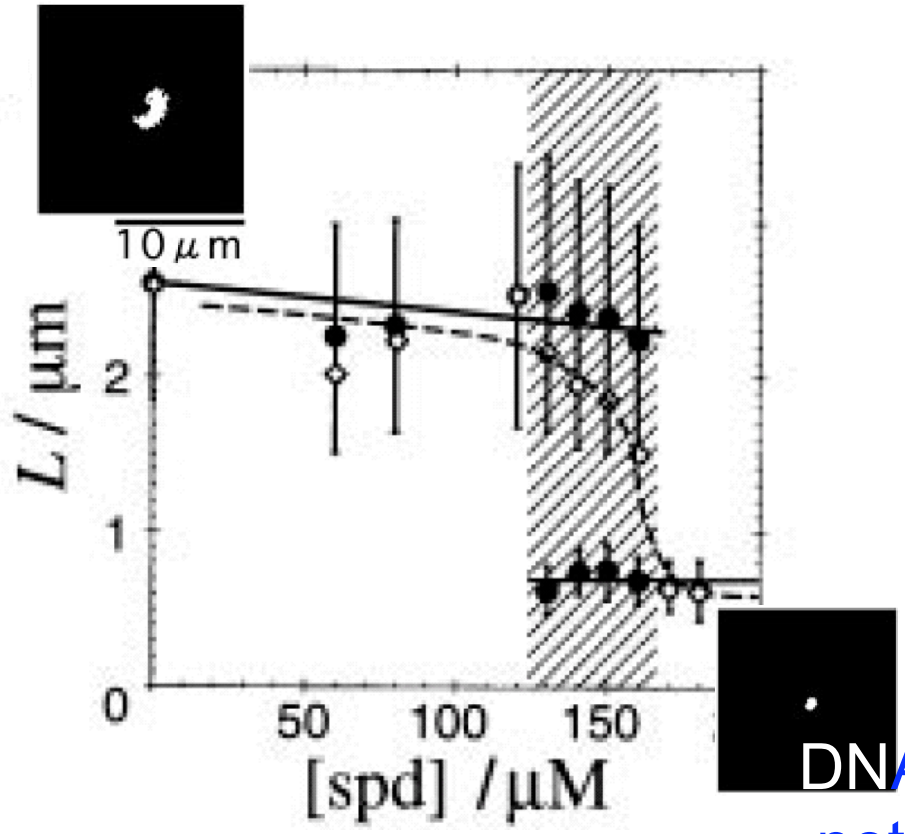
Uppsala University

Parallel Flat Histogram Simulations

Malek O. Khan

Dept. of Physical Chemistry, Uppsala University

Original motivation: DNA condensation



Fluorescence microscopy of DNA with multivalent ions

K. Yoshikawa et al, *Phys. Rev. Letts.,* 76, 3029, 1996

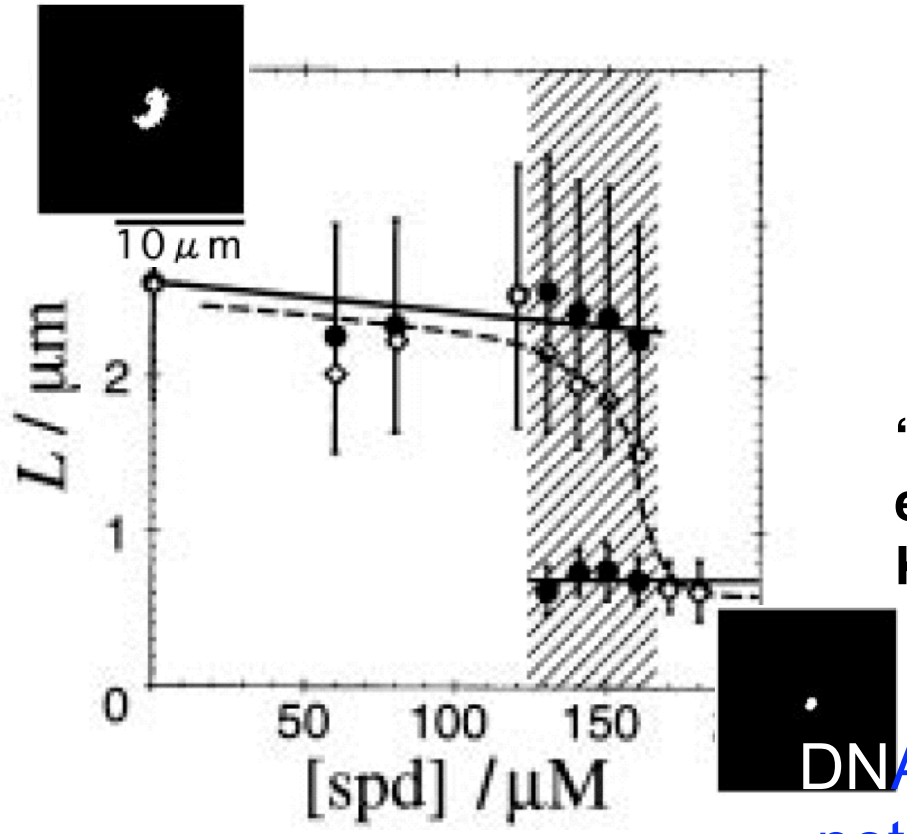
DNA either elongated or compact not in between

Parallel Flat Histogram Simulations

Malek O. Khan

Dept. of Physical Chemistry, Uppsala University

Original motivation: DNA condensation



Fluorescence microscopy of DNA with multivalent ions

K. Yoshikawa et al, *Phys. Rev. Letts.,* 76, 3029, 1996

"unambiguous interpretation of experimental observations" - Kenneth Ruud

DNA either elongated or compact not in between

Polyelectrolyte model

- Hamiltonian

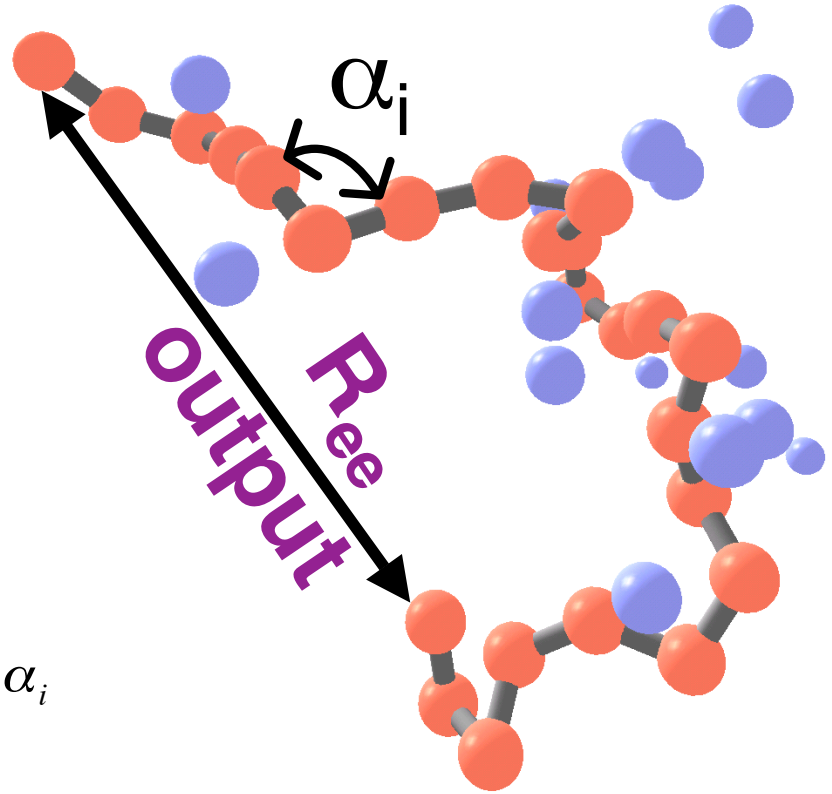
$$U_0 = U_{bond} + U_{es} + U_{hc} + U_{angle}$$

- Fixed bond length
- Electrostatics

$$U_{es} = \sum_{i<j}^{N+N_c} \frac{q_i q_j e^2}{4\pi\varepsilon_r\varepsilon_0 \left|\vec{r}_i - \vec{r}_j\right|}$$

- Hard sphere particles
- Intrinsic stiffness

$$U_{angle} = C\sum_{i=2}^{N-1}\cos\alpha_i$$



- Outer spherical cell
- Moves - clothed pivot & translation

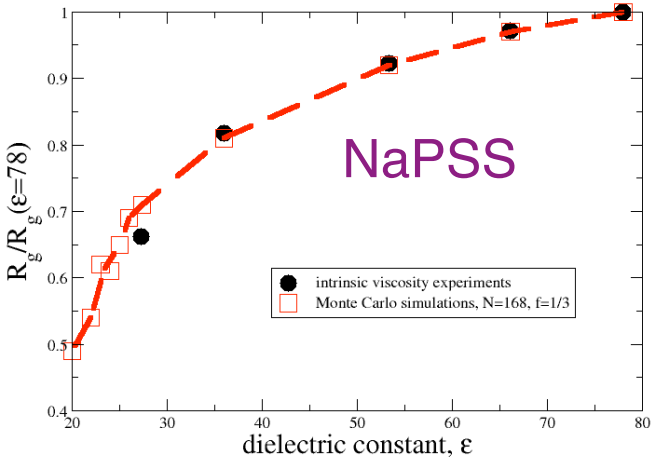
Condensation of polyelectrolytes - MC

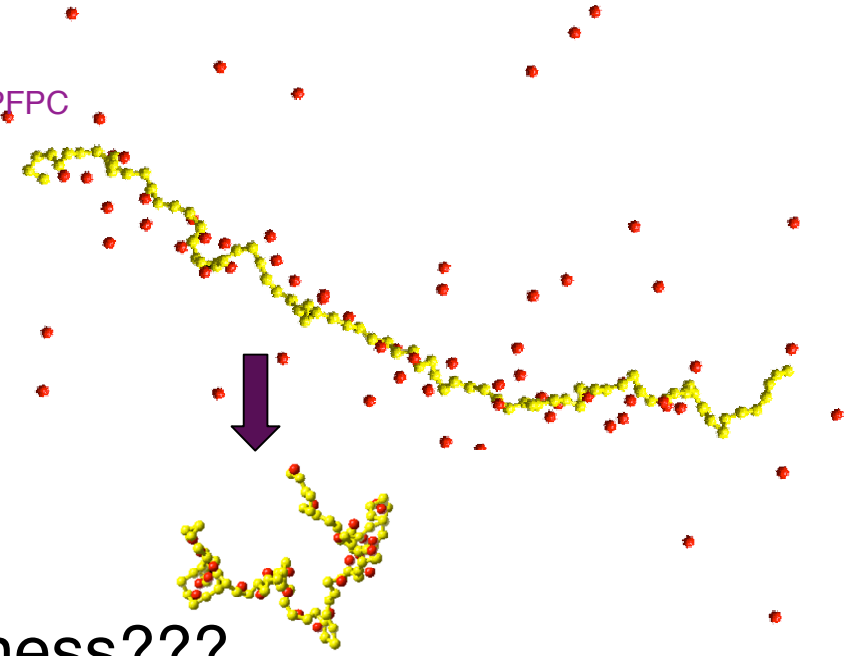
- Polyelectrolyte conformation
  - Stretched under normal conditions
  - Large electrostatic interactions lead to condensation

$$U_{es} = \frac{q_i q_j e^2}{4\pi\varepsilon_r\varepsilon_0\left|\vec{r}_i - \vec{r}_j\right|}$$

- Flexible polyelectrolytes

Experiments by R. Watson, J. Cooper-White & V. Tirtaatmadja, PFPC





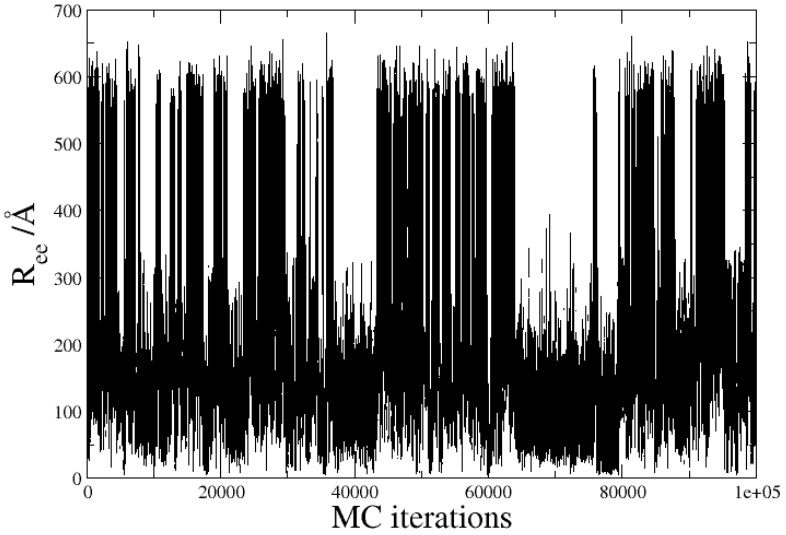
- Effect of intrinsic chains stiffness???

Problem 1: Mixture of length scales - bonds and Coulomb lead to slow convergence --> parallel calculations

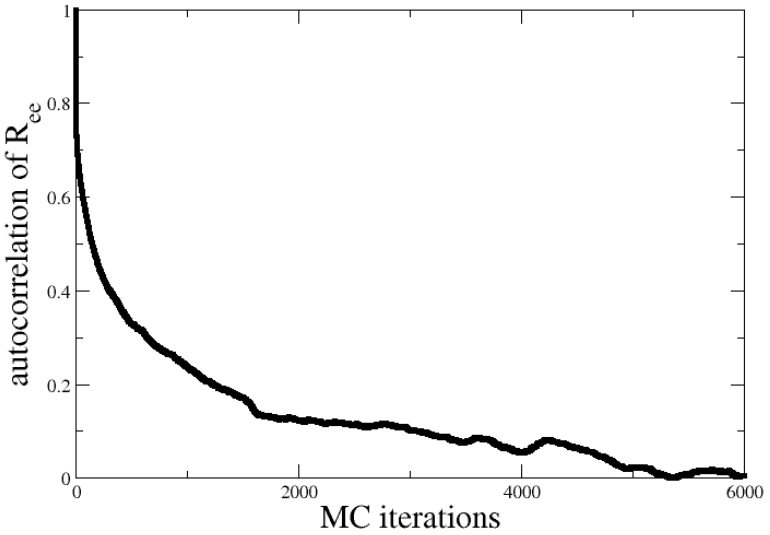
Problem 2: Long range Coulomb interaction - every particle interacts with every other particle

Convergence for stiff PE (N=128)

Total simulation time = 6 days



Autocorrelation time ~ hours



Months of computer time needed for stiff polyelectrolytes

Problem 1: Mixture of length scales - bonds and Coulomb lead to slow convergence --> parallel calculations

Problem 2: Long range Coulomb interaction - every particle interacts with every other particle in the Monte Carlo Simulation

Solutions

- Cluster moves - clothed pivot
- Parallel expended ensembles
- Parallel flat histogram techniques

Parallel flat histogram simulations

- Our implementation is a parallel implementation of a serial algorithm introduced by Engkvist & Karlström and Wang & Landau
- Instead of importance sampling create a flat distribution of the quantity of interest
- Correctly done this gives the potential of mean force (POMF) as a function of the quantity of interest

Engkvist & Karlström, Chem. Phys. 213 (1996), Wang & Landau, PRL 86 (2001)

Potential of mean force, *w*

$$p(\xi_0) = \frac{\int \exp\left[-\beta U(\vec{r})\right]\delta\left[\xi - \xi_0\right]d\vec{r}}{\int \exp\left[-\beta U(\vec{r})\right]d\vec{r}} \qquad w(\xi) = -k_B T \ln\left(p(\xi)\right)$$

Add *U** $$p^*(\xi_0) = \frac{\int \exp\left[-\beta U(\vec{r}) + U^*(\xi_0)\right]\delta\left[\xi - \xi_0\right]d\vec{r}}{\int \exp\left[-\beta U(\vec{r}) + U^*(\xi_0)\right]d\vec{r}}$$

$$p(\xi) = p^*(\xi)\exp\left[\beta U^*(\xi)\right]C_1$$

If *p*(ξ)* = constant

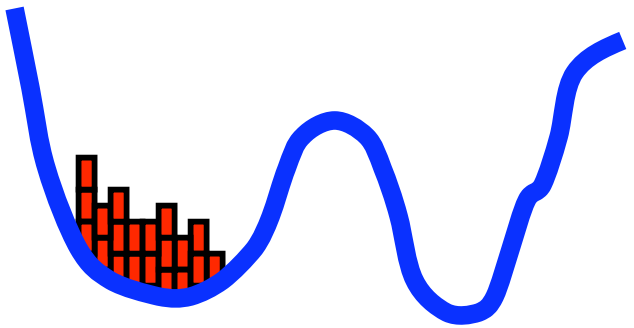
$$p(\xi) = \exp\left[\beta U^*(\xi)\right]C_2$$

$$U^*(\xi) = -w(\xi) + C_3$$

New formulation of the problem: construct a “flat” *p*

Implementation

- Discretize $U^*(\xi)$ and set to zero (here ξ is R_{ee})
- For every ξ visited, update $U^*(\xi)$ with δ_{pen}
- Repeat until $p^*(\xi)$ is "flat"
- Decrease δ_{pen} ----> $\delta_{pen}/2$
- Repeat until δ_{pen} is small
- Parameters:
  - Number of bins (~10^2-10^3)
  - Initial choice of δ_{pen} (0.001-1k_BT)
  - What is "flat" (max[$|p^*(\xi) - \langle p^*(\xi)\rangle|$] < (0.1-0.35)
  - Finish when δ_{pen} < (10^{-8} - 10^{-5})

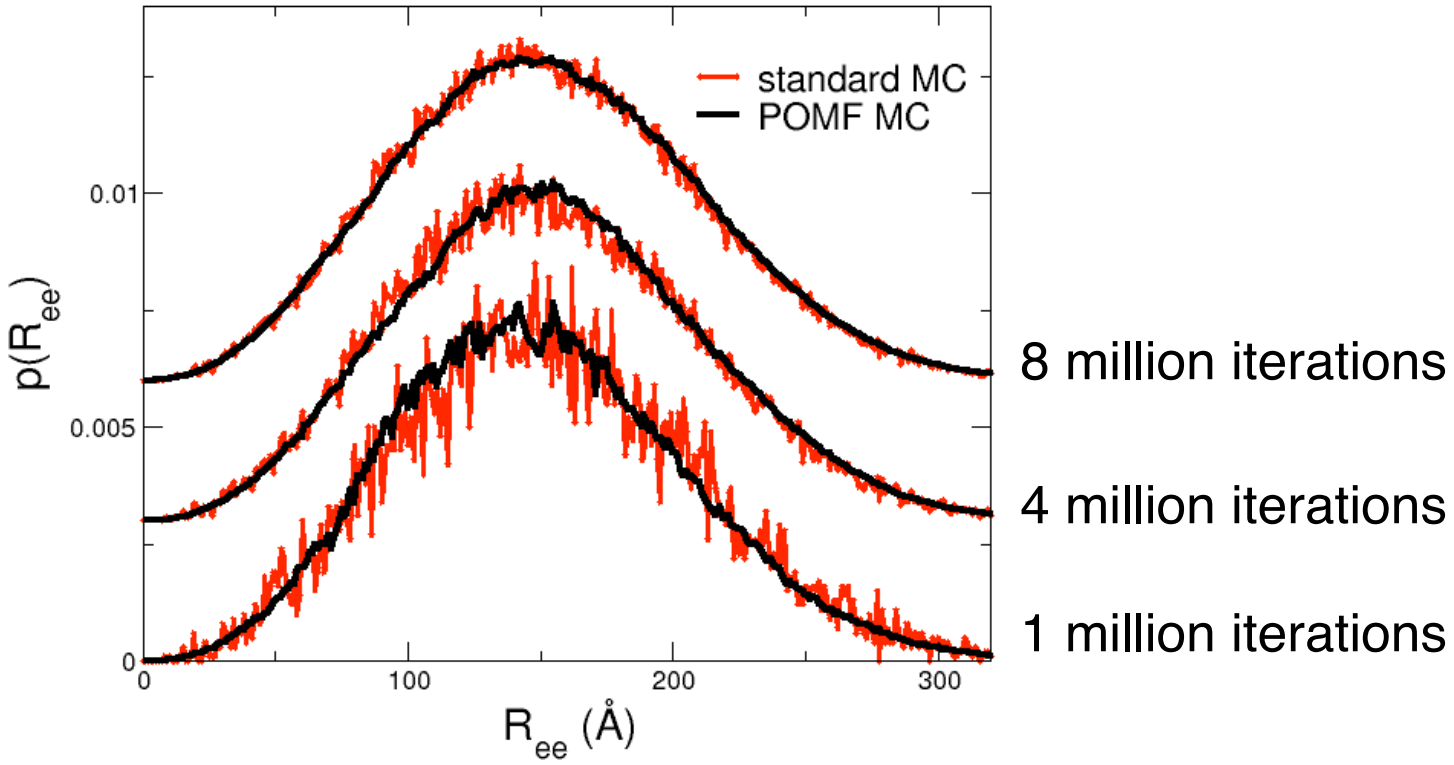


Parallel implementation

- Run copies on N_{cpu} processors with different random number seeds
- Calculate individual U^* and p^* on every CPU
- During simulation sum U^* and p^* from all processors
- Distribute $\langle U^* \rangle_{cpu}$ to all processors
- Check averaged $\langle p^* \rangle_{cpu}$
- Each processor does not have a constant p^* but the sum over N_{cpu}

Distribution functions

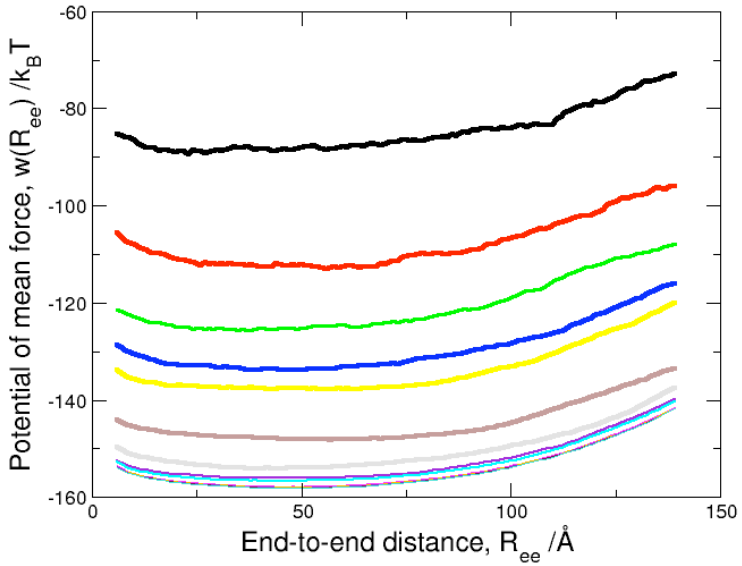
Neutral polymer, N=240

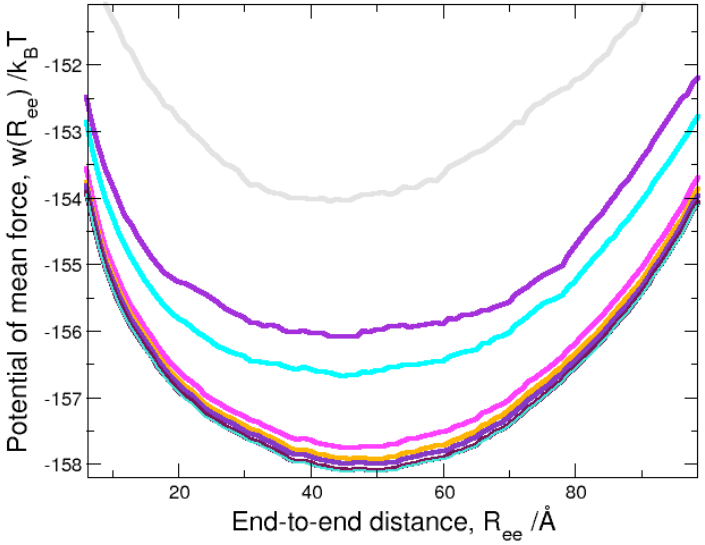


Flat histogram method at least of same quality

Evolution of the potential of mean force

Polyelectrolyte, N=64, tetravalent counterions
POMF at every update of δ_{pen} shown below
The right graph only shows the last 8

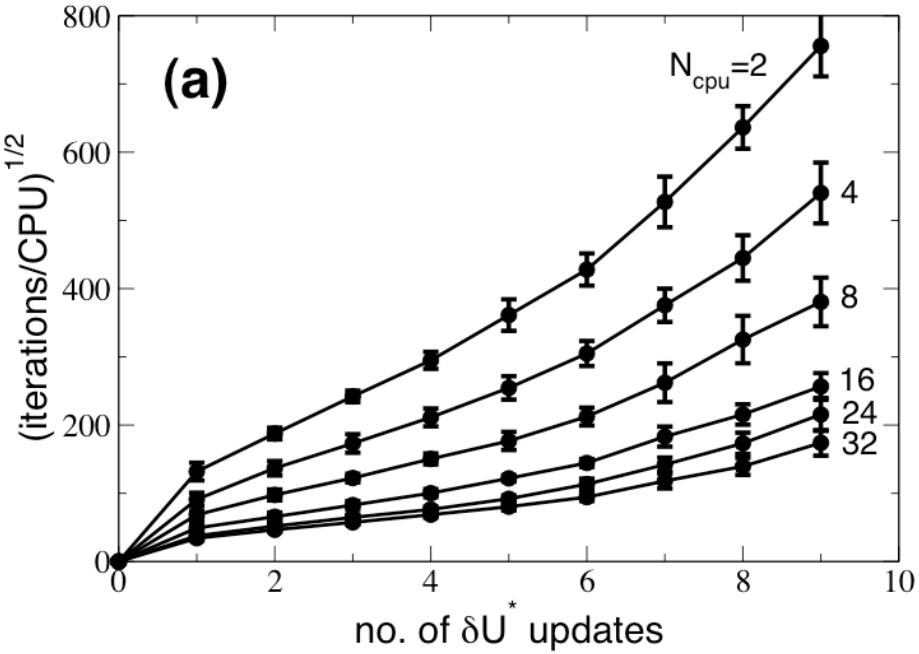


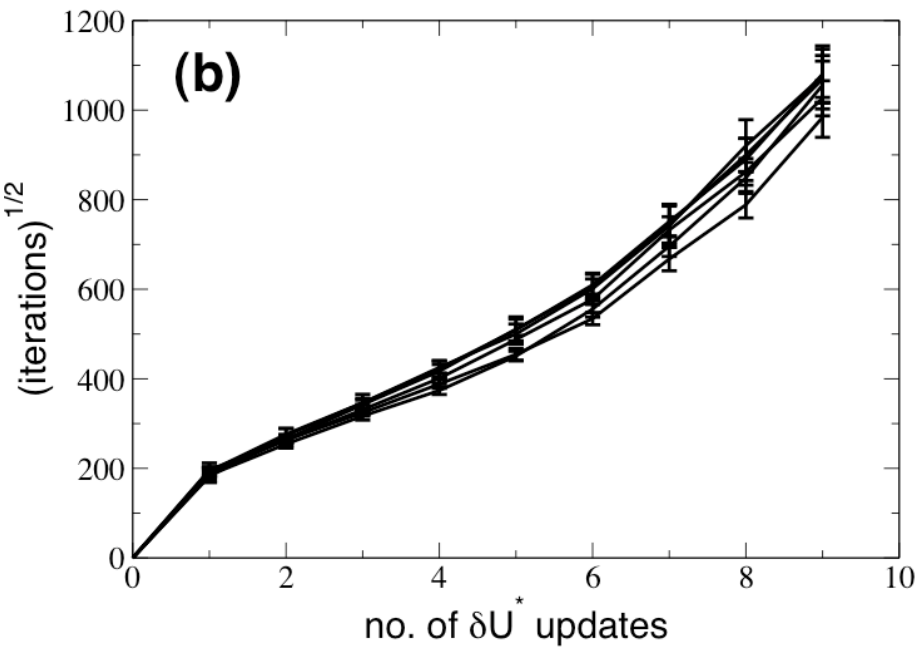


The POMF converges to a solution. There is no way of knowing if it is the correct solution. Experimental approach has to be taken.

Time between updates

Experimental approach: Do it 11 times and collect statistics

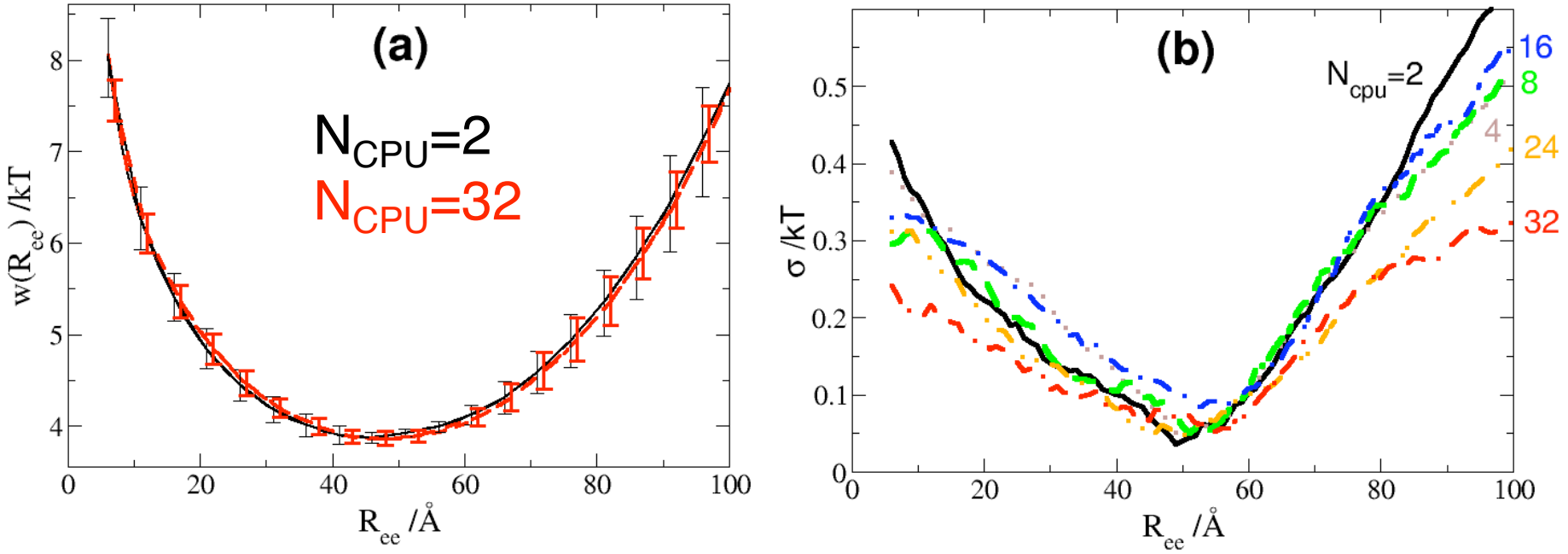




Time between updates is independent of N_{cpu}

Errors in the POMFs

Experimental approach: Do it 11 times and collect statistics

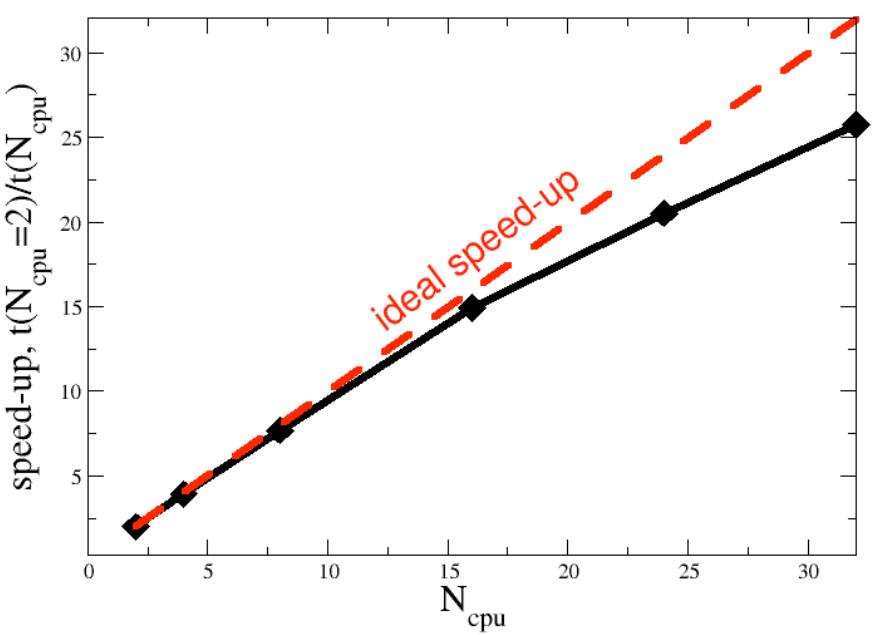


Error is independent of N_{cpu}

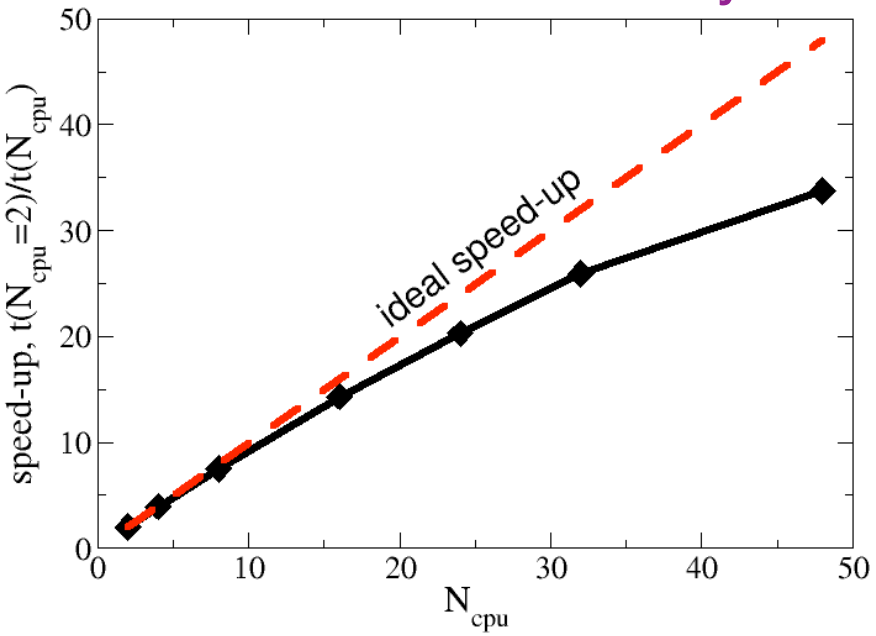
Parallel efficiency

Polyelectrolyte, N=64, tetravalent counterions

alpha SC

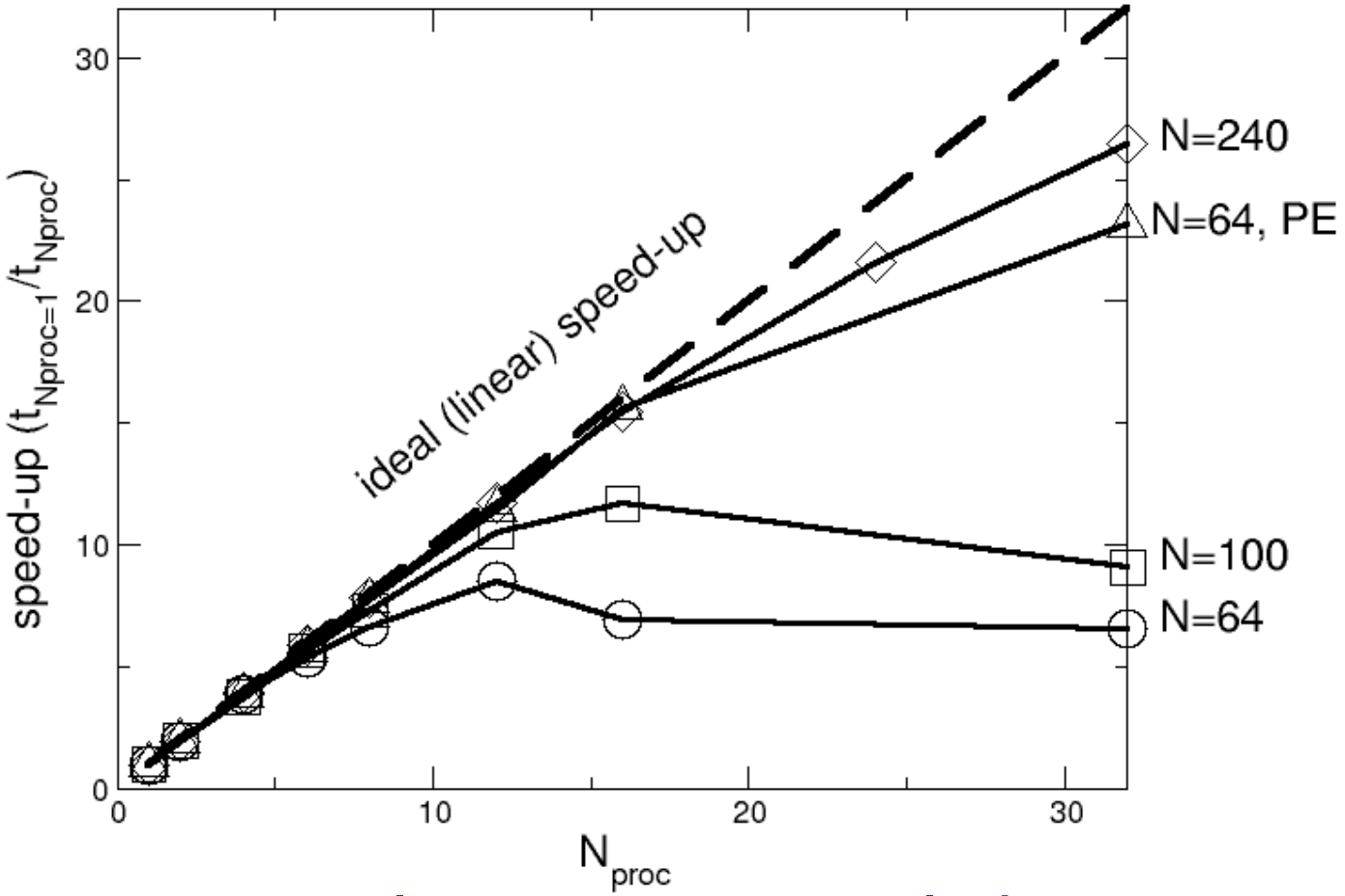


Brecca Beowulf - 194 CPU
Xeon 2.8 GHz with Myrinet



Extra time for communication is small up to N_{cpu}=32

Parallel efficiency (effect of system size)

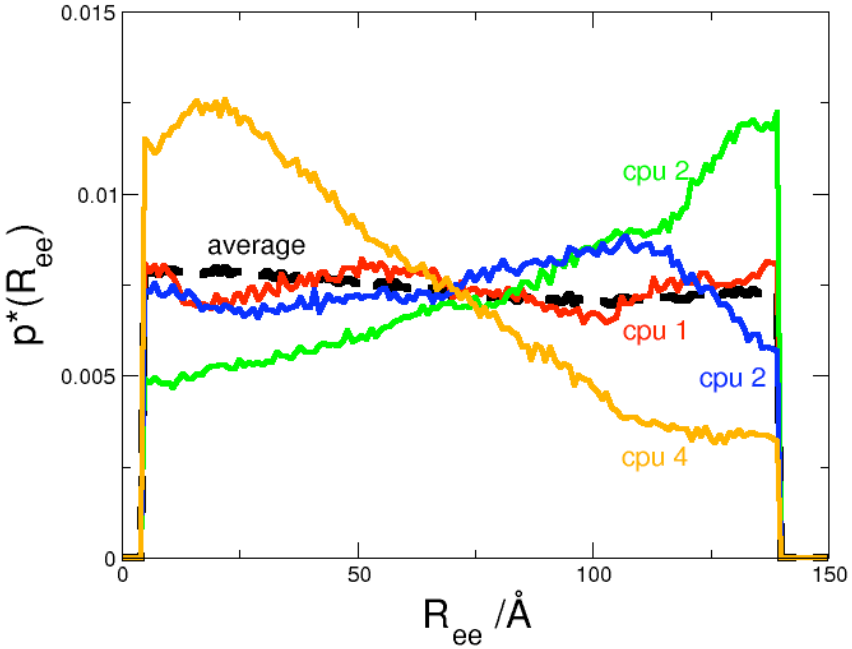


Larger, more complex, systems scale better

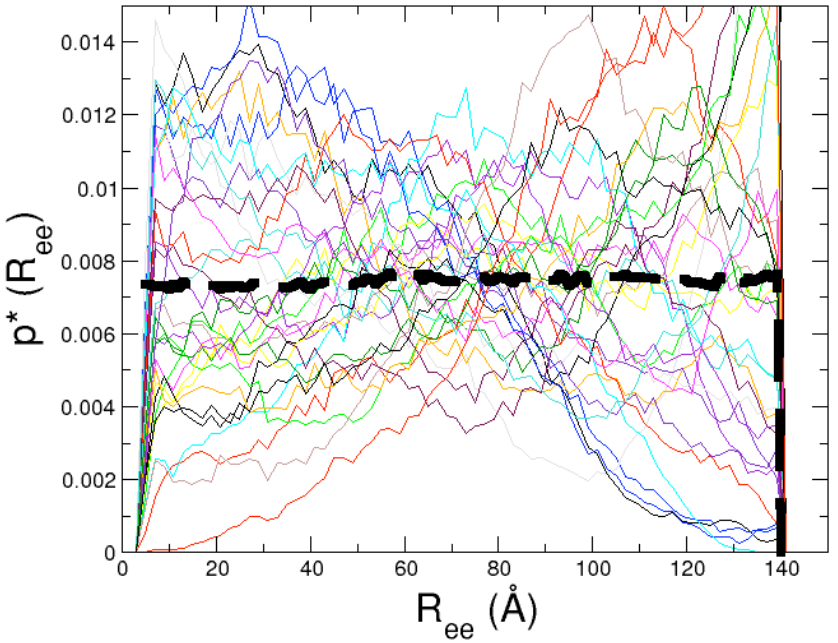
Individual processors

Polyelectrolyte, N=64, tetravalent counterions

$N_{cpu} = 4$



$N_{cpu} = 32$

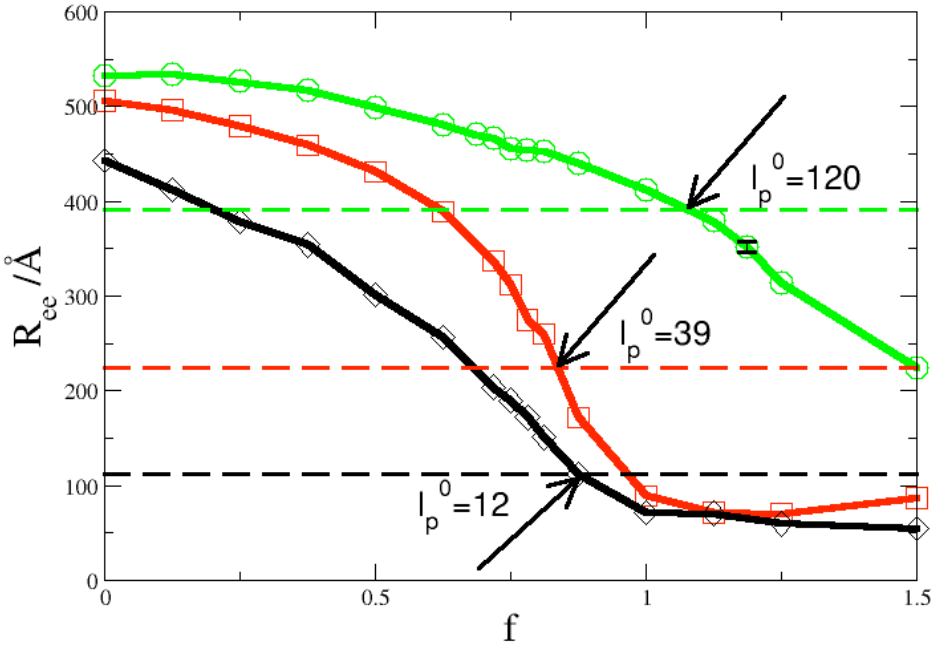


All CPUs do not have a flat histogram - the sum has

Case study: Polyelectrolytes with intrinsic stiffness

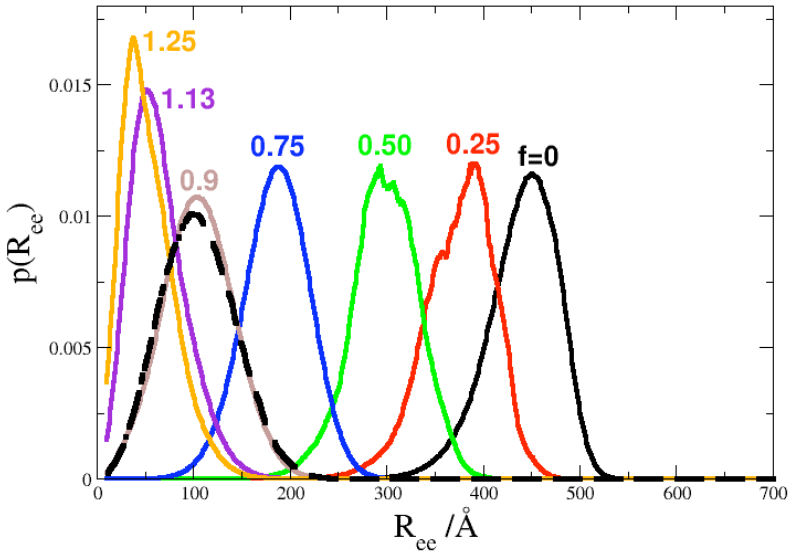
Polyelectrolyte, N=128, monovalent counterions + added tetravalent salt

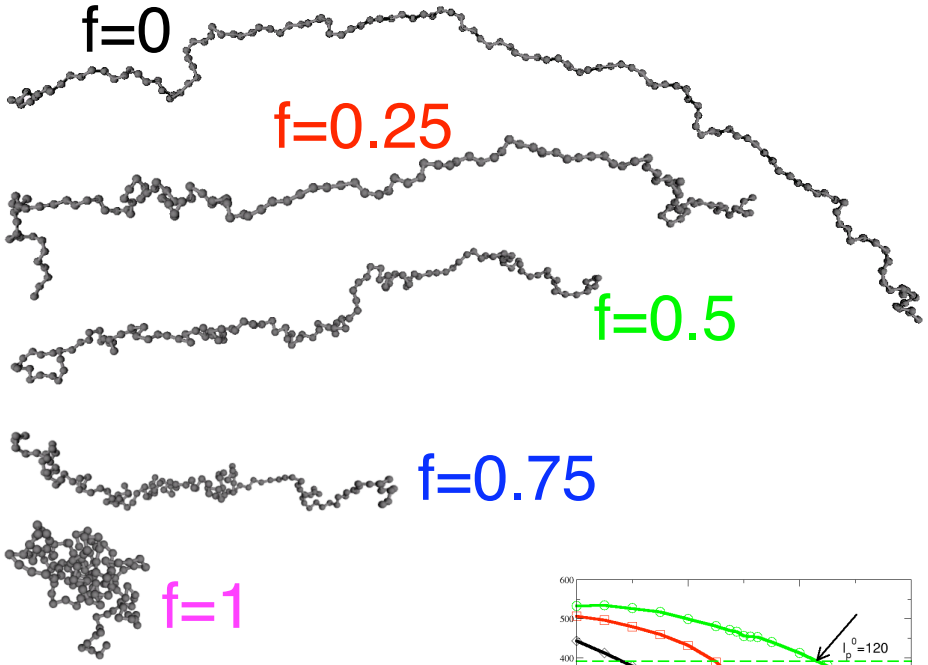
Scales to 64 processors on Power 5 (VPAC) and Itanium2 (APAC)

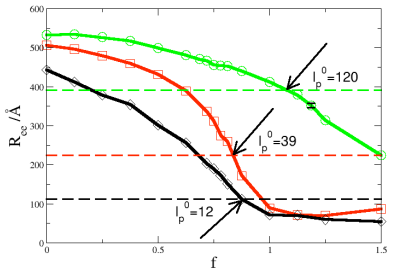


Distribution functions - Flexible PE

Polyelectrolyte, N=128, monovalent counterions + added tetravalent salt



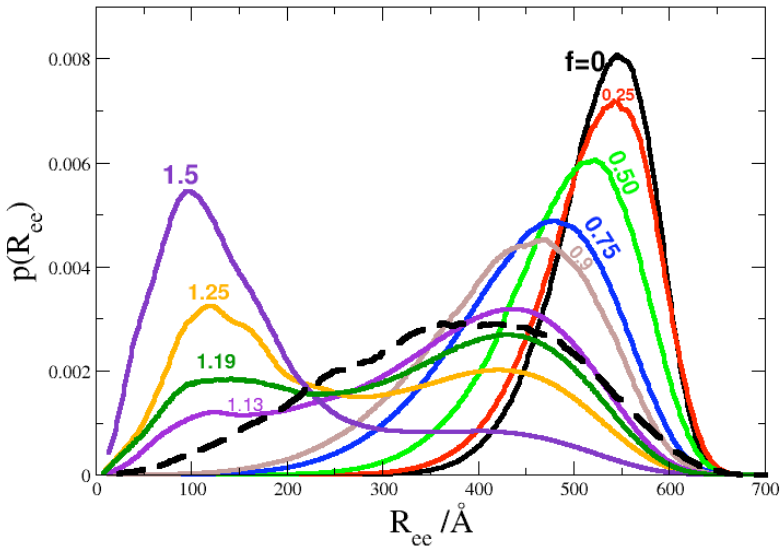


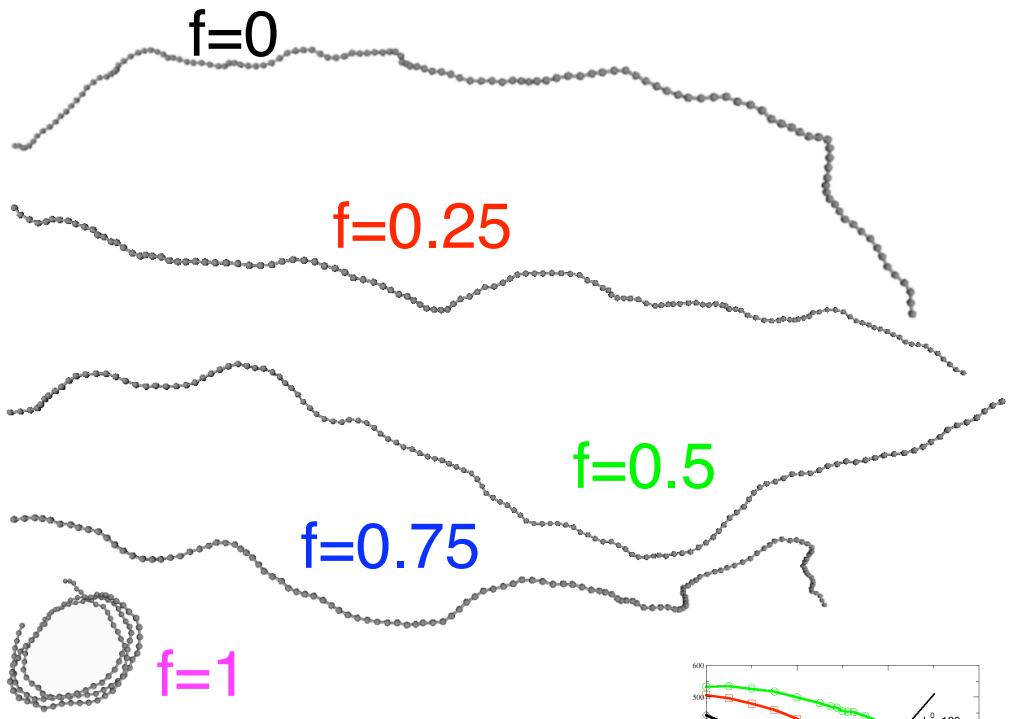


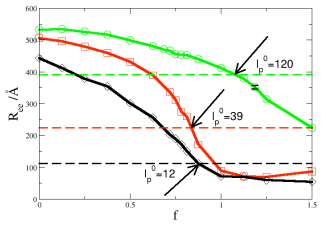
l_p^0=12 Å

Distribution functions - stiff PE

Polyelectrolyte, N=128, monovalent counterions + added tetravalent salt



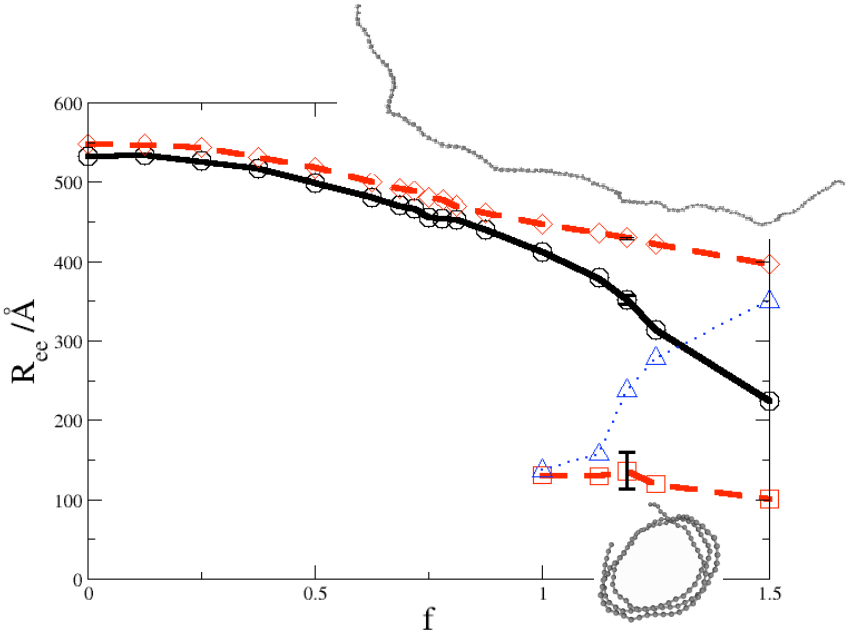




l_p^0=120 Å

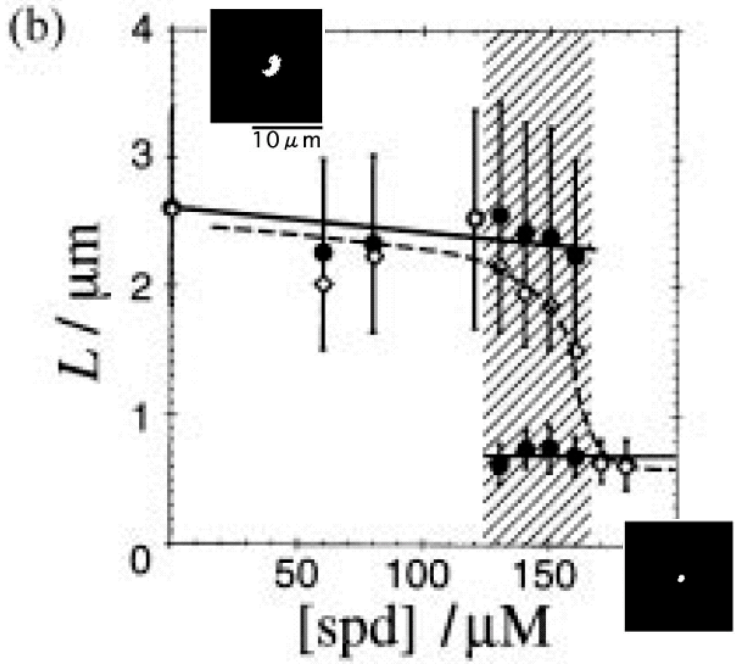
Summary - Stiff polyelectrolytes

Monte Carlo



Fluorescence microscopy of DNA

K. Yoshikawa et al, *Phys. Rev. Letts.,* 76, 3029, 1996



- Possible to simulate all or nothing phase transition of stiff polyelectrolytes, see double maxima for n_c=38 on previous page
- Simulation time for each point is 1 week on 24 processors on brecca (VPAC 2.8GHz Xeon with Myrinet interconnect) (5.6 CPU months)

Summary - Parallel flat histograms

- ✓ Gives the free energy directly
- ✓ Allows exploration of areas of phase space which are difficult to reach with conventional MC - complements importance sampling
- ✓ Parallelisation is easily implemented and shown to scale linearly to a large amount of CPUs on clusters
  - Time between updates is independent of N_{CPU}
  - Error is independent of N_{CPU}
  - Distribution is flat over all CPUs not every individual one
  - CPU time does not increase with N_{CPU} (for large systems)

Acknowledgments

People

Derek Chan – Big boss
Simon Petris – PhD student: double layers
Gareth Kennedy – MSc student: computational methods
Malek Ghantous – Honours student: polymer conformation

Grants & Organisations

Wenner-Gren Foundation – post doc grant
ARC
PFPC at The University of Melbourne
VPAC and APAC

Australia