Dynamics of biological and synthetic polymers through large-scale parallel computations

I. D. Dissanayake[a] and P. Dimitrakopoulos[a] * †

[a]Department of Chemical and Biomolecular Engineering,
University of Maryland, College Park, Maryland 20742, USA

We present the efforts of our research group to derive an optimized parallel algorithm so that we are able to efficiently perform large-scale numerical studies on the dynamics of biological and synthetic polymers. To describe the polymer macromolecules we employ Brownian dynamics simulations based on a semiflexible bead-rod chain model. Our algorithm has been parallelized by employing Message Passing Interface (MPI) on both shared- and distributed-memory multiprocessor computers with excellent parallel efficiency. By utilizing this methodology, we routinely employ up to 320 of the fastest processors to study polymer chains with more than $40,000$ beads. Our optimized parallel algorithm facilitates the study of polymer dynamics over a broad range of time scales and polymer lengths.

1. INTRODUCTION

The dynamics of polymer solutions is a problem of great technological and scientific interest since these systems are encountered in a broad range of industrial, natural and physiological processes. Common examples include biopolymers such as DNA, actin filaments, microtubules and rod-like viruses, as well as synthetic polymers such as polyacrylamides, Kevlar and polyesters. These polymers show a wide range of stiffness which results in some unique properties of their solutions and networks. For example, networks of actin filaments can provide biological cells with mechanical stability while occupying a significantly smaller volume fraction of the cytosol than would be required for a flexible network. In the case of the stiff synthetic polymers, stiffness is responsible for the macroscopic alignment of the chains in the system which imparts unique mechanical properties for these materials. Thus, the study of these systems is motivated by both the biological relevance and engineering applications.

The numerical study of polymer dynamics is computationally expensive due to the large molecule lengths and the requirement of monitoring the polymer properties over extended time periods. To overcome this obstacle, we have developed an optimized parallel algorithm and thus we are able to efficiently perform large-scale numerical studies on

---

*Corresponding author, email: `dimitrak@eng.umd.edu`

†This work was supported in part by the M. Martin Research Fund, the National Center for Supercomputing Applications (NCSA) in Illinois, and by an Academic Equipment Grant from Sun Microsystems Inc.

polymer dynamics. Our interest lies on understanding the physical properties of important biopolymers such as DNA, actin filaments and microtubules as well as of synthetic polymers. Our studies are further motivated by the recent development of experimental techniques which study individual biological molecules as well as by the recent development of micro-devices involving stretched tethered biopolymers (e.g. [1–3]).

2. MATHEMATICAL FORMULATION

To describe a semiflexible polymer chain, we employ the Kratky-Porod wormlike chain model [4,5] based on a Brownian dynamics method developed in Ref.[6]. This method considers a (flexible) bead-rod model with fixed bond lengths and ignores hydrodynamic interactions among beads as well as excluded-volume effects. The polymer chain is modeled as $N_B = (N+1)$ identical beads connected by N massless links of fixed length b (which is used as the length unit). The position of bead i is denoted as $\boldsymbol{X}_i$, while the link vectors are given by $\boldsymbol{d}_i = \boldsymbol{X}_{i+1} - \boldsymbol{X}_i$.

To account for polymer stiffness, we add a bending energy proportional to the square of the local curvature. For a continuous chain the bending energy is given by

$$\phi^{bend} = \frac{\mathcal{E}b}{2} \int_0^L (\frac{\partial \hat{\boldsymbol{d}}}{\partial s})^2 \mathrm{d}s = \frac{\mathcal{E}b}{2} \int_0^L (\frac{\partial^2 \boldsymbol{X}}{\partial s^2})^2 \mathrm{d}s, \tag{1}$$

where L is the (constant) contour length of the chain and $\hat{\boldsymbol{d}}$ the local unit tangent. The bending energy $\mathcal{E}$ is related to the persistence length L_p via $\mathcal{E}/k_B T \equiv L_p/b$, where k_B is the Boltzmann constant. The bending energy of the discrete model is given by

$$\phi^{bend} = \frac{\mathcal{E}b}{2} \sum_{i=2}^{N} (\frac{\boldsymbol{X}_{i+1} - 2\boldsymbol{X}_i + \boldsymbol{X}_{i-1}}{b^2})^2 b = \mathcal{E} \sum_{i=1}^{N-1} (1 - \frac{\boldsymbol{d}_i \cdot \boldsymbol{d}_{i+1}}{b^2}) \tag{2}$$

and thus it depends on the angle θ_i between two successive links since $\boldsymbol{d}_i \cdot \boldsymbol{d}_{i+1} = b^2 \cos\theta_i$. For a fixed b, the properties of the polymer chain are specified by the number of links N and the dimensionless bending energy $E = \mathcal{E}/k_B T$.

Assuming that the bead inertia is negligible, the sum of all forces acting on each bead i must vanish, which leads to the following Langevin equation

$$\zeta \frac{\mathrm{d}\boldsymbol{X}_i}{\mathrm{d}t} = \boldsymbol{F}_i^{bend} + \boldsymbol{F}_i^{rand} + \boldsymbol{F}_i^{ten} + \boldsymbol{F}_i^{cor}, \tag{3}$$

where ζ is the friction coefficient and $\boldsymbol{F}_i^{rand}$ the Brownian force due to the constant bombardments of the solvent molecules. The force $\boldsymbol{F}_i^{ten} = T_i \boldsymbol{d}_i - T_{i-1} \boldsymbol{d}_{i-1}$, where T_i is a constraining tension along the direction of each link $\boldsymbol{d}_i$, ensures the link inextensibility. $\boldsymbol{F}_i^{cor}$ is a corrective potential force added so that the equilibrium probability distribution of the chain configurations is Boltzmann [6]. The bending force $\boldsymbol{F}_i^{bend}$ is derived from the discrete form of the bending energy, Eq.(2),

$$\boldsymbol{F}_i^{bend} = -\frac{\partial \phi^{bend}}{\partial \boldsymbol{X}_i} = \frac{\mathcal{E}}{b^2} \sum_{j=i-2}^{i} \left[\delta_{j,i-2}\, \boldsymbol{d}_{i-2} + \delta_{j,i-1}\, (\boldsymbol{d}_i - \boldsymbol{d}_{i-1}) - \delta_{j,i}\, \boldsymbol{d}_{i+1} \right]. \tag{4}$$

(In the equation above as well as in all the equations in this paper, a term exists only if its index can be defined within its permitted bounds.) The resulting system based on Eq.(3) may be solved in $O(N)$ operations facilitating the study of long and/or stiff chains. The polymer properties are determined as the ensemble averages of the corresponding instantaneous values by employing 10^4 to 10^5 independent initial configurations.

Note that although we model the polymer molecule as a bead-rod semiflexible chain, the mathematical formulation presented above is more general since it can also represent other polymer models including bead-spring models with Hooke, Fraenkel or FENE-type springs [7]. Our preference for the bead-rod semiflexible model results from the fact that this model allows a continuous crossover from a freely-jointed flexible chain to a rigid rod as the bending energy increases while it preserves the link lengths and thus the contour length of the entire chain. We emphasize that the tensions required to keep the link lengths constant play a significant role in the polymer dynamics both near and far from equilibrium as recent studies have revealed [6,8–10].

3. RELEVANT TIME SCALES

The Brownian forces give rise to a series of time scales associated with the diffusive motion of the chain's increasing length scales, from that of one bead $\tau_{rand} = \zeta b^2/k_B T$, up to the time scale for the entire chain length which is $\tau_f = N^2\,\tau_{rand}$ for flexible chains and $\tau_s = N^3\,\tau_{rand}$ for stiff ones. Similarly, the bending forces give rise to a series of time scales associated with the bending vibrations of portions of the polymer chain with increasing length. The smallest bending time scale is associated with the relaxation of the angle between two successive links given by $\tau_{bend} = \zeta b^2/\mathcal{E} = \tau_{rand}/E \ll \tau_{rand}$, while the largest time scale is associated with the entire polymer chain and given by $\tau_\perp = \zeta L^4/\mathcal{E} b^2 = (N^4/E)\,\tau_{rand}$. If we consider that the length N is large for polymer molecules, it is clear that all polymers, and especially the stiff chains, are associated with very small time scales while their behaviors cover extended time periods.

Therefore the numerical study of polymer chains is computationally expensive; even with the present powerful single-processor computers it may take months or even years to simulate a long polymer chain over an extended time period. To overcome this obstacle, we have developed an efficient parallel algorithm which utilizes the great computational power of current multiprocessor supercomputers, especially the relatively inexpensive PC clusters, as discussed in the next section.

4. PARALLEL ALGORITHM FOR POLYMER DYNAMICS

Due to the computationally expensive nature of Brownian dynamics simulations of polymer molecules, it is imperative that the corresponding algorithm is fully optimized. To optimize our code, we employ highly-tuned LAPACK routines for the solution of the tridiagonal systems resulting from the equation of motion, Eq.(3). For the calculation of the chain's normal modes we use highly-tuned FFT algorithms provided by system libraries. In addition, our FORTRAN code has been optimized using cache optimization techniques.

We note that further optimization efforts have been made. Many physical problems involve polymers at equilibrium. For example, a DNA molecule, before undergoing a shearing flow, may be at equilibrium. To determine the equilibrium configuration, one has to allow the chain to equilibrate first over many time-consuming Brownian dynamics steps. In this case it is more efficient to start directly from an equilibrium chain configuration. To produce these equilibrium configurations we apply the following procedure based on the chain bending energy. In particular, for completely flexible chains (i.e. $E = 0$), an equilibrium configuration is derived as a "random-walk" polymer configuration generated by choosing successive bead positions from random vectors distributed over the surface of a sphere. For chains with non-zero bending energy, we employ a Monte Carlo-Metropolis algorithm to calculate an equilibrium chain configuration based on the chain bending energy. The Monte-Carlo algorithm for calculating equilibrium configurations is much faster than the Brownian dynamics algorithm; thus, significant savings are achieved. We emphasize that for any chain's bending energy, our ability to efficiently produce equilibrium configurations is a consequence of the corrective potential force $\boldsymbol{F}^{cor}$ we described in section 2. (Thus, bead-rod models which do not include this correction need to be equilibrated first by employing Brownian dynamics time steps.)

To study the dynamic evolution of polymer macromolecules, we numerically determine the ensemble average value of properties of interest including the chain configuration, the polymer stress and the solution birefringence. This is achieved by starting from (different) initial polymer configurations and following a different sequence of Brownian forces to simulate different chain realizations. Our optimized algorithm can successfully be employed on single-processor computers for relatively short chains and time periods; for larger polymer molecules we employ multiprocessor supercomputers and suitable parallelization.

To describe this, observe that the ensemble average calculation is by nature parallel while we can use as many processors as the number of initial configurations which is of order 10^4–10^5. In practice, a moderate number of processors (of order 10–300) is adequate for most cases. For the parallelization we use Message Passing Interface (MPI) [11,12]. The great advantage of MPI is its portability to different platforms. With this parallelization, we have the ability to employ both shared-memory multiprocessor computers (e.g. IBM pSeries 690) and distributed-memory supercomputers (e.g. Linux clusters). On all machines, the efficiency is almost 100% even for a high number of processors as shown in Table 1. This is due to the limited time of communication among the processors compared with the CPU time consumed for calculations on each processor. Therefore, our numerical code is ideal for all types of the state-of-the-art supercomputers.

By utilizing this optimized parallel algorithm, we routinely employ up to 320 of the fastest processors to study polymer chains with more than $40,000$ beads and thus we are able to identify the polymer behavior of very long chains at very small times and for extended time periods. For example as we discussed in Refs.[13,10] for the problem of the relaxation of initially straight flexible and stiff chains, we were able to determine the polymer evolution over 17 time decades for flexible chains and over 28 time decades for stiff ones. In addition, the length scales we study (up to at least $N = 40,000$ beads) correspond to DNA molecules up to 2.6 mm long and to synthetic molecules of polystyrene with molecular weight of $O(10^7)$!

Therefore, by employing this optimized parallel algorithm we are able to identify all physical behaviors even if they appear at very small times and span extended time periods. For example, we have studied the dynamics of chain stiffening due to temperature change

Table 1
Efficiency versus the number of processors N_p for a typical problem by employing our parallel Brownian Dynamics algorithm on the shared-memory IBM pSeries 690 and the distributed-memory Linux Xeon Supercluster provided by the National Center for Supercomputing Applications (NCSA) in Illinois. Note that efficiency denotes the ratio of the wall time for the serial execution T_s to that for the parallel execution T_p multiplied by the number of processors, i.e. efficiency $\equiv T_s/(T_p\, N_p)$.

NCSA IBM pSeries 690	
N_p	Efficiency (%)
1	100.00
2	99.81
5	99.75
10	99.64
16	99.24

NCSA Xeon Supercluster	
N_p	Efficiency (%)
1	100.00
10	99.39
20	99.09
25	98.65
50	97.16
160	97.02

or salt addition in the polymer solution [14–17]. An industrial application of this process involves the manufacturing of better trapping gel networks, micro-electronic devices from plastics and novel drug delivery systems [18]. In Figure 1 we present the shape evolution of an initially coil-like chain during stiffening towards its final rod-like shape. The associated transition from coil to helix and from helix to rod is obvious from this figure. These transitions are also shown in Figure 2 where we plot the time evolution of the polymer length and its widths. Observe that the time of the maximum chain's width corresponds to the helix creation while the time of the maximum chain's length corresponds to the rod-like final configuration.

5. CONCLUSIONS

The dynamics of polymer solutions is a problem of great technological and scientific interest but the numerical study of these systems is computationally expensive due to the large molecule lengths and the requirement of monitoring the polymer properties over extended time periods. To overcome this obstacle, we have developed an optimized parallel Brownian dynamics algorithm for the dynamics of biological and synthetic polymers via Message Passing Interface (MPI) based on ensemble average calculations. By employing both shared- and distributed-memory multiprocessor computers, we have the ability to determine and understand the polymer properties over a wide range of time scales and polymer lengths which has never been achieved before.

REFERENCES

1. T. T. Perkins, S. R. Quake, D. E. Smith and S. Chu, Science 264 (1994) 822.
2. T. T. Perkins, D. E. Smith and S. Chu, in Flexible Polymer Chains in Elongational Flow, T. Q. Nguyen and H.-H. Kausch (eds.), Springer, 1999.

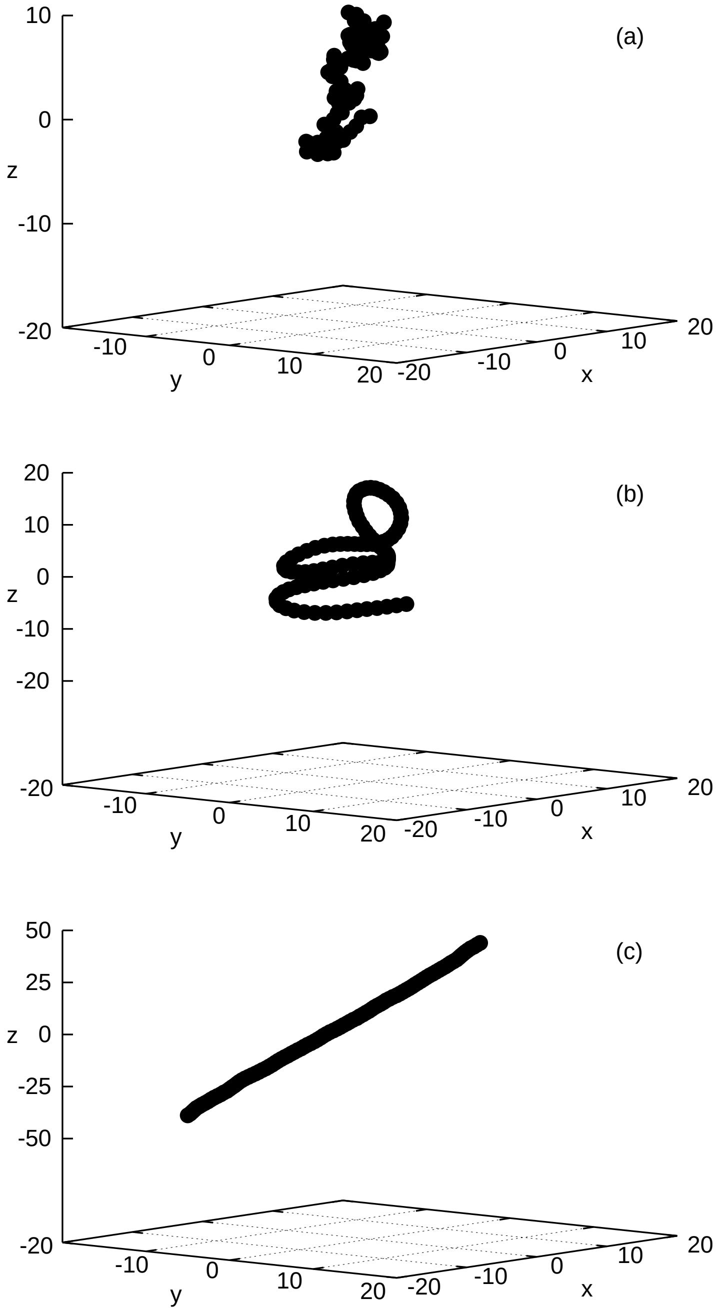

Figure 1. Evolution of a polymer molecule during backbone stiffening. The polymer length is $N = 100$ while the chain's final stiffness is $E/N = 10$. Three distinct configurations are presented: (a) the initial coil-like shape, (b) the intermediate helical shape, and (c) the final rod-like configuration.

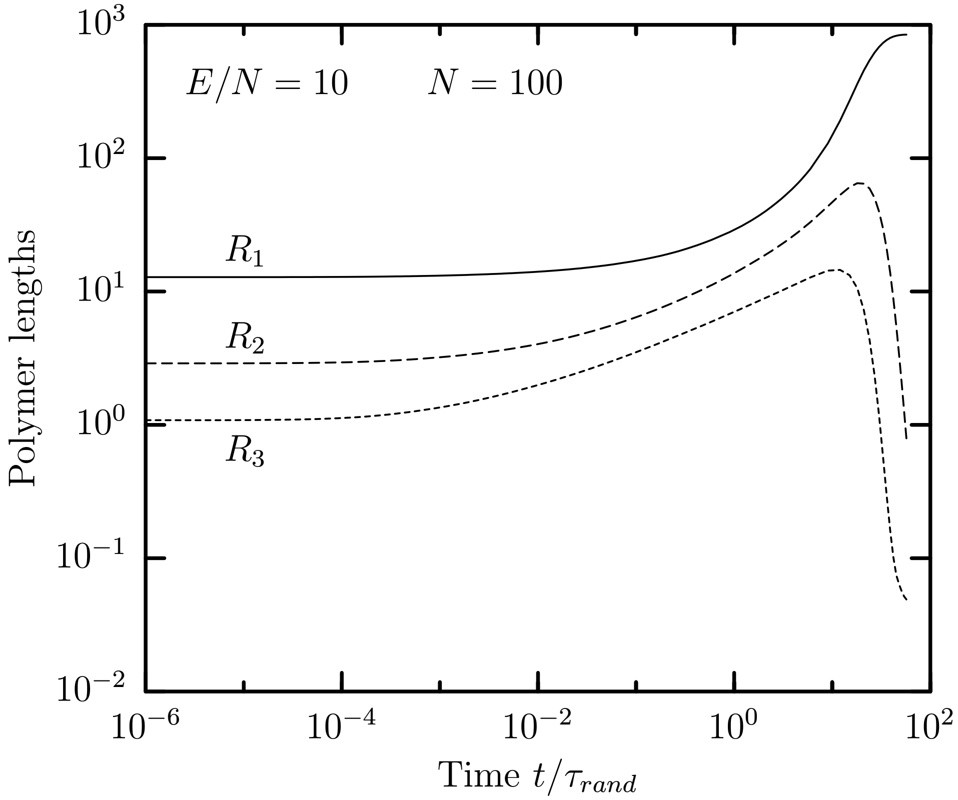

Figure 2. Time evolution of the lengths of a polymer molecule during backbone stiffening. The polymer's contour length is $N = 100$ while the chain's final stiffness is $E/N = 10$. Note that R_1 is the polymer length while R_2 and R_3 are the chain's widths.

3. G. J. L. Wuite, S. B. Smith, M. Young, D. Keller and C. Bustamante, Nature 404 (2000) 103.
4. H. Yamakawa, Helical Wormlike Chains in Polymer Solutions, Springer, 1997.
5. M. Doi and S. F. Edwards, The Theory of Polymer Dynamics, Clarendon, 1986.
6. P. S. Grassia and E. J. Hinch, J. Fluid Mech. 308 (1996) 255.
7. R. B. Bird, C. F. Curtiss, R. C. Armstrong and O. Hassager, Dynamics of Polymeric Liquids: Volume 2, Kinetic Theory, Wiley, 1987.
8. R. Everaers, F. Jülicher, A. Ajdari and A. C. Maggs, Phys. Rev. Lett. **82** (1999) 3717
9. V. Shankar, M. Pasquali and D. C. Morse, J. Rheol. **46** (2002) 1111
10. P. Dimitrakopoulos, Phys. Rev. Lett. 93 (2004) 217801.
11. Message Passing Interface Forum (eds.), Message Passing Interface, 2003 (available at `http://www.mpi-forum.org` and `http://www-unix.mcs.anl.gov/mpi`).
12. I. Foster, Designing and Building Parallel Programs, Addison-Wesley, 1995.
13. P. Dimitrakopoulos, J. Chem. Phys. 119 (2003) 8189.
14. Y. Y. Gotlib and L. I. Klushin, Polymer 32 (1991) 3408
15. C. G. Baumann, S. B. Smith, V. A. Bloomfield and C. Bustamante, Proc. Natl. Acad. Sci. 94 (1997) 6185
16. K. Terao, Y. Terao, A. Teramoto, N. Nakamura, M. Fujiki and T. Sato, Macromolecules 34 (2001) 4519.
17. J. R. Wenner, M. C. Williams, I. Rouzina and V. A. Bloomfield, Biophysical J. 82 (2002) 3160.
18. C. K. Ober, Science 288 (2000) 448.