GPU Based Molecular Dynamics Simulations of Polymer Rings in Concentrated Solution: Structure and Scaling

Daniel Reith,1,* Leonid Mirny,2,** and Peter Virnau1,***

1 Institut für Physik, Johannes Gutenberg-Universität Mainz, Staudingerweg 7, 55128 Mainz, Germany
2 Harvard-MIT Division of Health Sciences and Technology, and Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA

(Received May 8, 2011)

We report on equilibrium properties of a concentrated solution of non-concatenated ring polymers by Molecular dynamics simulations using HooMD-blue, a fast implementation on graphics processor units (GPUs). We are able to identify the intermediate scaling regime for the radius of gyration \( R_g \propto N^{0.4} \) as well as indication for a crossover to \( R_g \propto N^{1/3} \) for rings with chain length \( N \) in our fully flexible off-lattice polymer model. This crossover takes place between a ring size of 2500 and 7500 monomers for monomer density \( \rho = 0.5 \). Our results are in agreement with recent studies for lattice and stiff off-lattice models and show once again that this scaling is not model dependent at all. Furthermore this study demonstrates that GPUs are suited to tackle large scale problems in computational polymer science. To this end we present performance measurements relating the computing performance of a single NVidias GTX580 consumer graphic card to conventional state of the art supercomputers.

§1. Introduction

The dynamics in a polymer melt is driven by the reptation mechanism.1) Thereby a polymer chain moves along its contour in a snake-like motion because every other motion is hindered by entanglements imposed by neighboring chains. The lack of this mechanism in ring polymers changes the static and dynamic behavior dramatically.2)–22) Furthermore, it is impossible to change self-entanglements (knots) of a ring polymer without breaking bonds. The typical restriction to unknotted loops in this context corresponds to a constriction of phase space to a specific topological subclass. Knots also provide a direct measure of entanglements in single polymers,23)–25) akin to interchain concepts like the primitive path in melts26) and can serve as a starting point for the understanding of DNA in viral capsids,27), 28) proteins29)–34) and synthetic polymers.23), 35) Again, the presence or absence of such entanglements influences both static and dynamic properties.

While in a melt of open chains the radius of gyration in principle scales like a random walk (\( R_g \propto N^{1/2} \)) in the asymptotic limit, there are three different scaling regimes8), 9), 13), 18), 19), 21), 36) in a melt or concentrated solution of non-concatenated, unknotted rings. For very small polymer rings, \( R_g \propto N^{1/2} \), which is followed by an

---

* E-mail: reithd@uni-mainz.de
** E-mail: leonid@mit.edu
*** E-mail: virnau@uni-mainz.de
intermediate region with $R_g \propto N^{2/5}$ and finally a crossover to $R_g \propto N^{1/3}$. Computer
simulations of lattice models first showed evidence of this transition around 10 years
ago\textsuperscript{13} and recently a first off-lattice simulation of a semiflexible polymer model
confirmed these scaling laws.\textsuperscript{21,22} In this manuscript we investigate a fully flexible
off-lattice model. While the intermediate 2/5 region is consistent with a Flory-like
free energy minimization in which each topological constraint is penalized by $k_B T$, the
asymptotic 1/3 scaling is consistent with the fractal globule concept introduced
by Grosberg et al. in the late 1980s.\textsuperscript{37} Grosberg also concluded that DNA has to
be organized in exactly this structure in order to remain functional.\textsuperscript{38}

In the interphase (between two cell divisions) chromosomes are localized in so-
called chromosome territories in the eukaryotic cells,\textsuperscript{39–41} which can be visualized
with experimental techniques like Fluorescence In Situ Hybridization (FISH).\textsuperscript{42}
New experimental techniques like HiC\textsuperscript{43} allow for the visualization of the three-
dimensional architecture of whole genomes by constructing spatial proximity maps
with a resolution down to 1 megabase. With this spatial proximity maps it is possible
to calculate the probabilities for two loci being in contact, which are separated
by a distance $s$ along the genome. Several studies try to explain the development
of these chromosome territories in the context of polymer physics.\textsuperscript{44–49} While linear
chromosomes may simply not have enough time to fully equilibrate as argued in
Ref. 45), an analogy can be drawn to melts of rings which segregate due to topological
constraints.

In the following we report on simulations of fully flexible coarse-grained ring
polymers in the semi-dilute regime. We see indication for a crossover from the
intermediate $\nu = 2/5$ to the asymptotic $\nu = 1/3$ scaling in the Flory exponent and
analyse structures in both regimes. All calculations were notably done on low-budget
graphics processor units (GPUs) the performance of which we compare to state of
the art supercomputers in the second part of the paper.

\section{Model and simulation techniques}

Our model system consists of simple flexible Lennard-Jones + FENE homopolymers\textsuperscript{23,50,51} with cut and shifted Lennard-Jones interactions between all monomers

\begin{equation}
V_{LJ} = \begin{cases} 
4\epsilon \left[ (\frac{\xi}{\sigma})^{12} - (\frac{\xi}{\sigma})^6 + \frac{127}{16384} \right], & \text{if } r < r_{\text{cut}} \\
0, & \text{otherwise}
\end{cases}, \quad (2.1)
\end{equation}

and FENE interactions between adjacent beads:

\begin{equation}
V_{\text{FENE}} = -33.75\epsilon \ln \left[ 1 - \left( \frac{r}{1.5\sigma} \right)^2 \right]. \quad (2.2)
\end{equation}
Table I. Parameters of the simulated systems. All boxes are of cubic size $L \times L \times L$. We used periodic boundary conditions in all three directions.

<table>
<thead>
<tr>
<th>type</th>
<th>chain length</th>
<th>number of chains</th>
<th>$N_c$</th>
<th>number of monomers $N_{total}$</th>
<th>side length $L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>rings</td>
<td>250</td>
<td>4000</td>
<td>100000</td>
<td>58.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>2000</td>
<td>100000</td>
<td>58.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>108</td>
<td>108000</td>
<td>60.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2500</td>
<td>100</td>
<td>250000</td>
<td>79.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5000</td>
<td>100</td>
<td>500000</td>
<td>100.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7500</td>
<td>100</td>
<td>750000</td>
<td>114.472</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10000</td>
<td>100</td>
<td>1000000</td>
<td>125.992</td>
<td></td>
</tr>
<tr>
<td>open chains</td>
<td>1000</td>
<td>1000</td>
<td>1000000</td>
<td>125.992</td>
<td></td>
</tr>
</tbody>
</table>

Unless noted, $r_{cut} = 2^{1/2}$. The single chain simulations took place at $T = 4 \epsilon/k_B$, a temperature significantly above the $\theta$-temperature $T_\theta = 3.2 \epsilon/k_B$ (for a chain of infinite length), to ensure sufficient interpenetration. Note that at these high temperatures, the attractive part of the Lennard-Jones potential could in principle be omitted to further speed up the computations. The molecular dynamics simulations were performed using the GPU accelerated code HooMD blue revision 3574 in the NVT ensemble with a timestep of $dt = 0.002$ for the production and $dt = 0.004$ for the equilibration runs. We used the included DPD-thermostat with $r_{cut,DPD} = 2^{1/2}$ and $\gamma = 0.5$. Note that all calculations were done using the GPU’s single-precision arithmetic. All simulations started from globular single loop configurations, which were placed in the simulation box. For $N = 5000, 7500$ and $10000$ we prepared additional starting configurations from the final configurations for $N = 2500$ and $N = 5000$ to check for equilibrium. (Volume and loops were scaled up from $N = 2500$ and $N = 5000$ and additional particles were placed in between neighboring particles.) The potentials of Eqs. (2.1) and (2.2) were originally designed to prevent bond crossings. However, the probability of crossing a bond is small but non-zero and does depend on time step and details of the thermostat. To address this concern, we performed a primitive path analysis of the final configuration for $N = 5000$, which ran for more than three months. Unfortunately, we did discover a single link between two rings. A knot analysis, however, did not reveal any knots. Hence we do not believe that this single entanglement (which probably occurred in the initial relaxation phase) significantly alters our results. Further details on the specific systems can be found in Table I. A crucial parameter of the model is the entanglement length. This length can at least for open chains be estimated according to Ref. 26) via an empirical relation between packing and entanglement length

$$N_e = p^3 \times \frac{4}{5 \times 0.00226}.$$  \hspace{1cm} (2.3)

The packing length is defined as $p = NR_e^2/\rho$, where $R_e^2$ is the squared end-to-end vector, $N$ the number of monomers per chain and $\rho$ the overall monomer density. Putting in the numbers for the open chain simulation we obtain a value of roughly $N_e \approx 400$. Therefore, the biggest rings of size $N = 10000$ considered correspond approximately to $25N_e$ of the corresponding linear polymer at the same density.
§3. Results and discussion

One of the central static properties of concentrated solutions or melts of polymers is the radius of gyration of the polymer which measures the size of a single chain. Figure 1 shows the squared radius of gyration \( \langle R_g^2 \rangle \) as a function of time to determine if chains are equilibrated. At each time the average is calculated over all rings in the simulation box. While the systems appear to be equilibrated for ring sizes up to 5000 monomers, the plateau is not reached yet for ring sizes of 7500 and 10000 monomers. Our control runs from our “blown-up” configurations (to the right of Fig. 1) yield the same results for \( N = 5000 \) and similar results for the not yet fully equilibrated sizes \( N = 7500 \) and \( N = 10000 \). The scaling of the radius of gyration \( \langle R_g \rangle \) is expected to behave like \( \langle R_g \rangle \propto N^{0.4} \) for the intermediate and like

Fig. 1. Squared radius of gyration \( \langle R_g^2 \rangle \) over time \( t \) for different ring sizes. \( N = 7500 \) and \( N = 10000 \) are not fully equilibrated, yet, and only the last few configurations were used for analysis. The three time series on the right-hand side of the graph are runs starting from alternative starting configurations as described in the main text.

Fig. 2. Squared radius of gyration \( \langle R_g^2 \rangle \) as a function of chain length \( N \) in double logarithmic representation. The two straight lines \( \propto N^{0.8} \) and \( \propto N^{2/3} \) are guides to the eyes only to illustrate the intermediate and the asymptotic scaling regime. The additional squares at \( N = 5000 \), 7500 and 10000 denote the averages from runs with alternative initial configurations. Note that \( N = 7500 \) and \( N = 10000 \) are not fully equilibrated, yet.
\( \langle R_g \rangle \propto N^{1/3} \) for the asymptotic regime. We observe the intermediate scaling regime between ring sizes of \( N = 500 \) and \( N = 2500 \) monomers (Fig. 2). A fit yields an exponent \( \nu = 0.403 \) using ring sizes \( N = 500, \ 1000, \ 2500 \). At ring size \( N = 5000 \) monomers the scaling starts to deviate from the \( \nu = 0.4 \) scaling. Ring sizes of \( N = 7500 \) and \( N = 10000 \) are not fully equilibrated, yet, but indicate a crossover to \( \nu = 1/3 \). Following Vettorel et al.\(^{19}\) the onset of the asymptotic regime can be estimated by an argument using the entanglement length \( N_e \) of linear chains at the same density. Topological constraints start to play an important role in melts or concentrated solutions of open chains on a length scale larger than the entanglement length \( N_e \). On such a length scale ‘tube defects’ become entropically unfavorable if the strands leaking out of the chain’s tube are longer than \( L_e = b \sqrt{N_e} \), where \( b \) is the bond length. Vettorel et al. concluded that large rings favor collapsed states, because rings cannot reptate out of this situation. Relating the core size \( r_c \) of such a collapsed states to \( L_e \) and assuming \( r_c > L_e \) they expect to observe collapsed states with ring sizes

\[
N > N_e = \frac{4\pi}{3} \rho L_e^3 = \frac{4\pi}{3} \rho b^3 (N_e)^{3/2}.
\]

(3.1)

This would yield \( N_c \approx 15000 \) for our model, assuming that rings squeeze each other out completely and that at the center of mass of each ring there is no overlap with the surrounding rings at all. Halverson et al.\(^{21}\) did not see this complete squeeze out in their simulations and we observe a similar behavior for the case of a fully flexible model at a lower density. Furthermore they estimated the self-density \( \rho_S \approx 0.23 \) (for their largest system) at a melt density of \( \rho = 0.85 \) for their data. In other words in their semi-flexible model at melt density roughly 75 percent of the monomers in the vicinity of the center of mass have to be attributed to other chains. Figure 3 shows that we see a similar behavior for our fully flexible model. In our case the self-density converges at roughly \( \rho = 0.22 \) at the center of mass (55 percent of monomers in the vicinity of the center of mass belong to neighboring chains). Halverson et al.\(^{21}\) again proposed to include this observation in the arguments of Vettorel et al. and modify Eq. (3.1) such that the overall density \( \rho \) is replaced by the self-density \( \rho_S \). Hence, we end up with a new estimate of the critical ring size of \( \tilde{N}_c \approx 6700 \).
In Fig. 4 we report on internal properties of the rings. The contact probability of two monomers over the separation of these two monomers along the chain is plotted in double logarithmic representation (Fig. 4 (left)) and defined as follows: Two monomers are in contact if their distance is smaller than $d_{\text{min}} = 2 \cdot \sqrt{2}$. The data for open chains of length $N = 1000$ is also included for comparison in the plot. The open chain contact probability $P(s) \propto s^{-3/2}$ scales nicely with an exponent of $-3/2$ over two orders of magnitude over the range from 10 to almost 1000 as predicted by theory and simulations.\cite{56,57} The contact probability of two monomers on a ring, however, differs already at length scales larger than 10 monomers. Halverson et al.\cite{21} found evidence, that the structure of non-concatenated rings in a melt is consistent with the model of the fractal globule\cite{37} with fractal dimension $d_f = 3$. For such structures the contact probability scales with exponent $-1$ as recently demonstrated in Refs. 43 and 57). Theoretically, the fractal globule is only expected for ring sizes $N \gg N_c$.\cite{37} Unfortunately we are not able to clearly observe the scaling with exponent $\gamma = -1$ in our data, probably because the rings are still too small compared to the linear entanglement length of $N_e \approx 400$. For our largest system ($N = 10000$, which is not yet fully equilibrated) $\gamma = -1$ scaling starts to emerge for $N > 1000$. Right next to the contact probability, Fig. 4 (right) shows the radius of gyration of sub-segments $s$ as an additional measure for the internal structure of a ring. Again, data for open chains of length $N = 1000$ is plotted for comparison. The sub-segment radius of gyration $R_g(s)$ appears to grow at a smaller rate for rings than for open chains. For larger sub-segments we are not able to observe a clear scaling law in our data. If the rings have a similar structure as a fractal globule, the sub-segment radius of gyration $\langle R_g(s) \rangle$ should scale with $\nu = 1/3$ because a fractal globule is
a self-similar structure. As for the contact probability, a ring size of roughly $25N_e$ is probably not sufficient to reach the $N \gg N_e$ limit and to observe the expected scaling.

§4. Performance

Modern Graphics Processing units (GPUs) are able to perform massively parallel scientific computations at low cost, but only few algorithms from the field of statistical physics have been proven to run efficiently on such devices. Besides calculations for the Ising-model\textsuperscript{58,59} and analysis of time series\textsuperscript{60} a lot of effort has focussed on porting Molecular dynamics simulations to GPUs.\textsuperscript{53,61,62} In the meantime traditional Molecular dynamics packages for parallel CPUs such as Gromacs,\textsuperscript{63} Lammps,\textsuperscript{64} Espresso\textsuperscript{65} and NAMD\textsuperscript{66,67} have also announced GPU support. In this section we would like to show that GPUs can be used to simulate systems, which originally relied on the usage of supercomputing clusters. As a test case we studied a concentrated solution of non-concatenated rings and their scaling behavior on traditional supercomputers and compare those results with single GPUs.

The system used for the benchmark is almost identical to the biggest system considered in Ref. 21), a Kremer-Grest (KG) based model.\textsuperscript{50} With this model, the authors demonstrated for the first time the discussed scaling relations for an off-lattice model. As it also includes a chain stiffness and as the densities are higher ($\rho = 0.85$), the largest systems ($N = 1600$) actually corresponds to 57 entanglement lengths according to Eq. (2.3) and is effectively larger than the biggest system used by us. Hence, the system is also better suited for identifying scaling relations. The nonbonded interaction is modeled with the help of a Weeks-Chandler-Anderson (WCA) potential, very similar to the one described in Eq. (2.1), but with a smaller cutoff $r_{\text{cut}} = \frac{6}{\sqrt{3}}$ and shifted in such a way, that $V_{WCA}(r_{\text{cut}}) = 0$ holds. The bonded interactions are described with the help of the FENE potential in Eq. (2.2). However the model considered in Ref. 21) is a semi-flexible and not a fully-flexible model as discussed in the previous section. To this end we introduce a chain stiffness by an additional angular potential

$$V_{\text{angular}} = \frac{1}{2}k(\theta - \theta_0)^2$$

(4.1)

with $k = 1.5$ and $\theta_0 = \pi$. Note that this is not exactly identical to the angular potential of Ref. 21) but rather its Taylor expansion up to second order. The benchmark system consists of 200 rings of size 1600 monomers with a total number of 320000 monomers in the system. They are packed into a cubic simulation box with side length $72.207\sigma$ with periodic boundary conditions in all three dimensions. This packing corresponds to a monomer density of $\rho = 0.85$. Figure 5 shows an equilibrated configuration of the system.

We compare HoomD-blue revision 3574\textsuperscript{53} running on different types of GPUs with Gromacs 4.x\textsuperscript{63} running on two state of the art parallel computing clusters at the JSC Jüllich: An Intel Xeon cluster (JUROPA) and a Blue Gene P supercomputer (JUGENE). To this end 100000 NVE integration steps with a timestep $dt = 0.01$
Table II. Details on the benchmarked platforms.

<table>
<thead>
<tr>
<th>name</th>
<th>GPU</th>
<th>CPU</th>
<th>MD code</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juropa</td>
<td>-</td>
<td>Intel <a href="mailto:X5570@2.93GHz">X5570@2.93GHz</a></td>
<td>Gromacs</td>
<td>4.5.3</td>
</tr>
<tr>
<td>JUGENE</td>
<td>-</td>
<td>PowerPC 450@850MHz</td>
<td>Gromacs</td>
<td>4.0.7</td>
</tr>
<tr>
<td>GTX580</td>
<td></td>
<td>HooMD-blue</td>
<td>r3547</td>
<td></td>
</tr>
<tr>
<td>Tesla M2070</td>
<td></td>
<td>HooMD-blue</td>
<td>r3547</td>
<td></td>
</tr>
</tbody>
</table>

were undertaken in each run to obtain a reasonable estimate for the speed of the integration. Note that both HooMD and Gromacs use single precision floating point operations. Of course, single precision is not particularly well-suited for NVE simulations. Therefore, these runs should only be considered for benchmark purposes.

Table II provides further details on the platforms. Figure 6 shows the speed of the NVE-integration over the number of used CPU cores. No performance loss through communication between the different CPU cores implies a perfect scaling and corresponds to a straight line with slope one in the double logarithmic scale of the plot. Gromacs measures the duration of the integration in hours per nanosecond (ns). In these units $dt = 0.1$ ps. HooMD-blue gauges its speed in timesteps per second. The HOOMD-blue units can be converted to the ones of GROMACS by multiplying their reciprocal value by $1000/(dt \times 60 \times 60)$. Hence the speed in units of ns/hour is plotted over the number of CPU cores in Fig 6. This plot shows that the specific systems scale rather nicely except for the biggest number of CPU cores taken into account on JUROPA as well as JUGENE. GPUs are massive parallel computing devices, but HOOMD-blue r3574 is not capable to run on more than one device in parallel. Therefore, the runtime of HOOMD-blue r3574 on different types of GPUs is denoted by different horizontal lines in the plot. The intersections of these horizontal lines with the scaling lines from JUROPA and JUGENE denote the number of CPU cores needed to obtain the same computing power from the respective cluster. In our
Polymer Rings in Concentrated Solutions

Fig. 6. Reciprocal time (speed) over the number of CPU cores on the JUROPA and JUGENE supercomputers used for integrating the benchmark system with Gromacs as described in the main text. The horizontal lines are the speed of integration with HOOMD-blue on different GPUs. A GTX580 roughly corresponds to 75 2.93 Ghz JUROPA cores and 850 BlueGene P cores on JUGENE.

test runs, a GTX 580 consumer card roughly corresponds to 75 2.93 Ghz JUROPA cores and 850 Blue Gene P cores on JUGENE. The corresponding speed-ups for the professional Tesla card are somewhat lower due to lower clock speeds and a lower number of compute cores. The exact speed up factors should, however, be taken with a grain of salt and rather serve as a guideline.

§5. Conclusion

The aim of this paper is twofold. First we demonstrate that GPUs can be used to tackle state of the art problems in the field of computational polymer physics. To this end we investigate the structure and scaling behavior of non-concatenated ring polymers in concentrated solution with a fully flexible off-lattice model by Molecular dynamics. Although our model and the parameters differ from models considered in the literature before, our results confirm once again that structural properties and scaling do not depend on a specific model. We find indication for a crossover to $\nu = 1/3$-scaling in the radius of gyration in the framework of a fully flexible off-lattice model. Furthermore we are able to give some additional evidence that the self-density of a ring at its center of mass does not converge towards the system density in a fully flexible model. The analysis of the contact probability and the sub-segment radius of gyration indicate, that even the longest chains considered in this study are not yet long enough to observe the scaling that is predicted for fractal globules.

In the second part of the paper we compare the speed of Molecular dynamics integration on a GPU using HooMD-blue to corresponding integrations on massively parallel computing clusters using Gromacs. We demonstrate that GPUs are an efficient and economic alternative to massive parallel clusters for Molecular dynamics integration of coarse-grained systems consisting of 100000 to 1000000 particles. Systems much bigger than 1 million particles cannot be simulated yet because of memory...
limitations, and an efficient implementation of Molecular dynamics parallelized over tens of GPUs is still missing. We found that a high end consumer graphic card (which at the time of writing costs less than 400 €) yields the same number of Molecular dynamics integration timesteps as 75 Xeon processor cores and roughly 850 Blue Gene P processor cores. These numbers should however only serve as a guideline.

Acknowledgements

We would like to thank A. Stasiak, K. Binder, H.-P. Hsu, A. Tröster, K. Koschke and M. Imakaev for fruitful discussions on the topic and J. Ahrens for technical support. We are also grateful to K. Kremer for sending preprints of Refs. 21) and 22). This work received financial support from the Deutsche Forschungsgemeinschaft (SFB 625-A17). Computing time on the JUROPA supercomputer at the NIC Jülich and the GPU cluster of the Center for Computational Sciences Mainz at the ZDV Mainz are also gratefully acknowledged.

References

7) S. Nam, J. Leisen, V. Breedveld and H. W. Beckham, Macromolecules 42 (2009), 3121.