Computational Studies of Polyetherimides: Beyond All-Atom Molecular Dynamics Simulations

Chengyuan Wen

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Physics

Shengfeng Cheng, Chair
Uwe C. Täuber
Jean Joseph Heremans
Guoliang Liu

December 13, 2019
Blacksburg, Virginia

Keywords: Molecular Dynamics, Monte Carlo, Coarse-Graining, Glass Transition Temperature, Polyetherimide

Computational Studies of Polyetherimides: Beyond All-Atom Molecular Dynamics Simulations

Chengyuan Wen

(ABSTRACT)

Polyetherimides are an important class of engineering thermoplastics used in a broad range of industries and applications because of their high heat resistance and stability, high strength and moduli, excellent electrical properties over a wide range of temperatures and frequencies, good processability, good adhesive properties, and chemical stability. All-atom molecular dynamics (MD) simulation is a useful tool to study polymers, but the accessible length and time scales are limited. In this thesis, we explore several computational methods that go beyond all-atom MD simulations to investigate polyetherimides. First, we have developed a transferable coarse-grained MD model of polyetherimides that captures their mechanical and thermal expansion properties. Our results show that in order to make the model transferable, it is critical to include an entropic correction term in the coarse-grained force field and require the coarse-grained model to capture the thermal expansion property of polyetherimides. Secondly, we have constructed a predictive model of the glass transition temperature (T_{g}) for polyimides by using machine-learning algorithms to analyze existing data on T_{g} reported in the literature. The predictive model is validated by comparing its predictions to experimental data not used in the training process of the model. We further demonstrate that the diffusion coefficients of small gas molecules can be quickly computed with all-atom MD simulations and used to determine T_{g}. Finally, we have developed a Monte Carlo (MC) program to model the polymerization process of branched polyetherimides and to compute their molecular

weight distribution for a wide range of systems, including fully reacted, partially reacted, stoichiometric, and nonstoichiometric ones. The MC results are compared to the predictions of the Flory-Stockmayer theory of branched polymers and an excellent agreement is found below the gel point of the system under consideration. Above the gel point, the Flory-Stockmayer theory starts to fail but the MC method can still be used to quickly determine the molecular weight distribution of branched polyetherimides under very general conditions.

Computational Studies of Polyetherimides: Beyond All-Atom Molecular Dynamics Simulations

Chengyuan Wen

(GENERAL AUDIENCE ABSTRACT)

Polyetherimides are an important category of engineering plastics with wide applications in many fields because of their superior mechanical, thermal, chemical, and electrical properties. All-atom molecular dynamics simulations serve as a useful tool to study the properties of polyetherimides *in silico*. However, such simulations are computationally expensive and therefore limited to small system sizes and short time scales. To overcome these issues, we employed various computational techniques in this thesis to model polyetherimides. First, we have developed a coarse-grained model of polyetherimides where atoms are grouped into beads. We show that molecular dynamics simulations on the basis of the coarse-grained model can be used to provide a reasonable description of the mechanical and thermal expansion properties of polyetherimides. Secondly, we have constructed a predictive model of the glass transition temperature, which is the temperature at which a material enters a glassy state when cooled rapidly, of polyimides using machine-learning algorithms. This model is capable of estimating the glass transition temperature of polyimides within an accuracy of $\pm$ 15 K even for those not synthesized yet. We further show that the diffusion coefficients of gas molecules, in addition to the polymer density, can be computed accurately with all-atom molecular dynamics simulations and used to determine the glass transition temperature of polyimides. Finally, we have developed a Monte Carlo scheme to efficiently model the polymerization and compute the chain-length distribution of branched polyetherimides under

very general conditions. The results from Monte Carlo simulations are compared to the predictions of the Flory-Stockmayer theory of branched polymers. The range of applicability of the theory is revealed. Overall, we have demonstrated several computational techniques that can be used to efficiently model polyetherimides, potentially other polymers as well, beyond the widely-used all-atom molecular dynamics simulations.

I dedicate this thesis to my beloved parents.

Acknowledgments

I would like to express my gratitude to my academic advisor, Prof. Shengfeng Cheng. This thesis would not exist without his guidance and mentorship. He introduced me to many different topics and gave me freedom to collaborate with other groups and researchers, which provided good opportunities for me to improve my communication skills and collaboration capabilities. When I ran into obstacles, his advice always pointed me to the right direction of moving forward. He also helped me improve my technical writing skills. Apart from research, he provided generous help to me in my life. I still remember the scene that he came to court to help me and my roommate argue with a rental company that tried to put incorrect charges on us. When I heard the news that my aunt was mortally ill, he immediately wrote an explanation letter to help me obtain a temporary leave and helped me cover my teaching-assistant duty while I was outside the US. I feel really lucky to have chosen him as my thesis advisor and I am proud to be his student.

I would also like to express my gratitude to the members of my Ph.D. committee, including Prof. Uwe C. Täuber, Prof. Jean Joseph Heremans, Prof. Guoliang Liu, and Prof. Eric Sharpe (former member). Their advice and support are always greatly appreciated. Prof. Sharpe was also my temporary advisor when I came to Virginia Tech. He helped me modify my essay once a week. Prof. Täuber gave constructive feedback and suggestions after all my talks, which helped me improve my presentation skills. He also read through this

thesis carefully and provided numerous detailed comments, which greatly helped enhance the quality of the thesis.

I thank Dr. Yanfei Tang who graduated from our group in 2018. I discussed with him about my research countless times when he was at Virginia Tech. I thank Binghan Liu who has helped me finish the work on T_g calculation using machine-learning methods. I also thank Britannia Vondrasek, whom I collaborate with on the project of developing functional polysulfones for water purification in the past several years. Some of our discussions on that project are valuable to my thesis work.

I also thank my roommates, including Xiangwen Wang, Weigang Liu, Wenchao Yang, and Wei Zhao. Their presence makes life much easier.

I thank all staff members of the Physics Department, in particular Betty Wilkins, Katrina Loan, Roger Link and Travis Heath for their help and assistance throughout my graduate life here.

I am grateful to the SABIC project team on designing new functional polyetherimides, including Prof. Tim Long, Prof. Shengfeng Cheng, Prof. Guoliang Liu, Dr. Roy Odle, Ke Cao and other team members. They contributed to ideas and insights that led to this thesis.

Finally, I would like to express my gratitude to my parents and my wife for their endless love and unconditional support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Engineers are seeking high-performance polymeric materials all the time. Materials with high strength, chemical resistance, and thermal/thermo-oxidative/thermomechanical stability are always of great interest as they can find wide applications in many real-world situations. Polyimide, first synthesized in 1908,[1] is one of the polymers that possess these excellent properties because of the presence of aromatic imide groups. However, polyimide is not readily melt-processable because of its infusible and insoluble nature due to their planar hetero-aromatic structure, making its production considerably expensive.[2] To overcome this disadvantage, synthetic chemists have synthesized polyetherimide by adding ether linkages to polyimide, making the polymer chains more flexible.[2] This modification causes the polymer to melt more readily but still retain the good properties brought about by the aromatic imide groups.

1.1 Polyetherimides

The first polyetherimide was developed in the 1970s by J. G. Wirth [3] and was introduced into the market with a commercial name "Ultem" by General Electric in 1982. The repeat unit of Ultem is shown in Fig. 1.1. In each repeat unit, there are two ether linkages, two imide groups, and a few aromatic rings. Since there are 68 atoms in one repeat unit, studying Ultem with all-atom models presents a big challenge. In the polymer industry, the name polyetherimide is usually exchangeable with Ultem. However, in this thesis we use the term polyetherimide to refer to a class of polymers containing ether and imide groups, which can be regarded as a sub-family of polyimide.

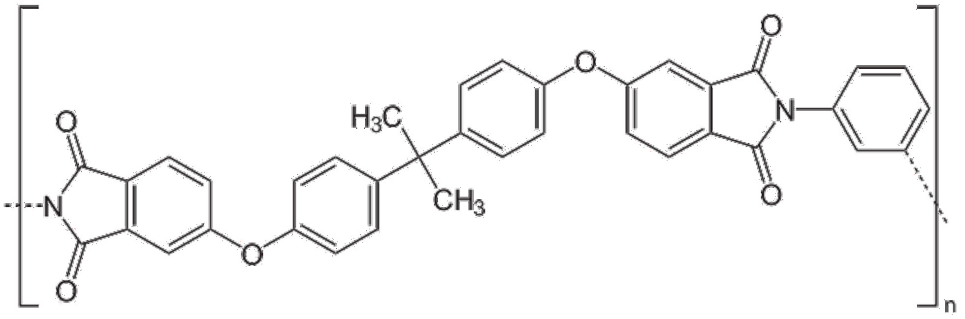

Figure 1.1: The repeat unit of Ultem

Polyetherimide maintains many good properties of polyimide, such as high strength to weight ratio, good thermal/thermo-oxidative stability, good heat resistance, strong chemical resistance, etc. It is widely applied in the automotive industry, aerospace engineering, electrical components, and medical devices.[4–6] Polyetherimide is also more amenable to processing. For example, Ultem can be processed at 340 °C. However, it is still a highly energy-intensive process and not very cost-effective. Therefore, synthetic chemists and materials scientists/engineers are still seeking new polyetherimides that can be processed at lower temperatures but at the same time maintain the desired properties of Ultem. For example, Cao *et al.*

synthesised UPy-terminated polyetherimides that are thermally stable and solution processable, and have mechanical properties still comparable to Ultem.[7] Hsiao *et al.* synthesised six different polyetherimides using triphenylamine-dietheramine monomers.[8] Tawade *et al.* synthesised three new polyetherimides with multiple ether linkages.[9]

Experimentally, synthesizing a new polyetherimide through an Edisonian trial-and-error approach is still time-consuming and expensive. A synthetic chemist may have to identify a correct pathway of synthesis via numerous tests. Even if the synthesis is successful, it is not guaranteed that the final product has the wanted properties. In the field of materials, there is a strong desire to use theory and computational modeling to speed up the process of material discovery, which is the principle behind the "Material Genome Initiative".[10] In this thesis, we explore several computational techniques to study polyetherimides, including all-atom molecular dynamics simulations, coarse-grained molecular dynamics simulations, Monte Carlo methods, and machine-learning algorithms, with the ultimate goal of assisting the *in silico* design and test of new functional polyetherimides. Such an effort is expected to reduce the cost and time of developing new chemical formulae that lead to polyetherimides with properties comparable to or even superior to Ultem but at the same time more suited for industrial processing and manufacturing.

1.2 Molecular Dynamics Simulations

Molecular dynamics (MD) simulation is a computational method to simulate a classical many-body system by solving the equations of motion of particles. To some extent, MD simulation fills the gap between experiment and theory, especially for a complex many-body system that is hard to treat analytically.[11] MD simulation is extremely useful in terms of mapping out atomistic and molecular scale details behind a mechanical/thermal/chemical

process and exploring extreme conditions, such as high temperature, high pressure, large voltage, vacuum, and gravity-free environments, which may be difficult to set up and control experimentally. These features make MD simulation an indispensable tool in the field of computational material science.[12, 13] In biology, MD simulation also plays a crucial role in elucidating and establishing structure-functional relationships for biological molecules, macromolecules, and complexes.[14]

In the early days, computational power was the main limiting factor of MD simulations. Only small systems could be simulated with MD methods. For example, Alder was one of the early developers and practitioners of MD simulations from the 1950s to 1970s.[15] The size of the systems that he studied was limited to about 1000 particles.[16–23] Interactions between particles were modeled with highly simplified intermolecular potentials. Since the late 1980s, more realistic force fields were developed and MD simulations were frequently used to model proteins.[24–29] Such simulations were usually limited to one or a few biomacromolecules with the total number of atoms in a system up to about 10,000. The evolution time of the system was typically at the order of about 1 ns or shorter. Around the same time, MD simulations were applied to study problems in polymer physics with model systems [30] and to compute physical properties of polymeric materials.[31–33] In the early studies of polymers, the bead-spring model with Lennard-Jones potentials was frequently used and results were usually qualitative rather than quantitative.[30, 32–34] As computation power was improved and more atomic/molecular force fields were developed with the aids of quantum-mechanics calculations and experimental data, quantitative studies of polymeric materials with all-atom MD simulations started to emerge and flourish.[35–37] However, even with computational power available nowadays, the length and time scales accessible to all-atom MD simulations are still limited to about 100 nm and 100 ns. Many processes such as fracture, rheology, and polymer solvation are still challenging problems for all-atom MD methods.[38, 39]

In this thesis, I apply three different molecular modeling approaches to study polyetherimides with the goal of overcoming the limitations of all-atom MD simulations. The first is to construct a coarse-grained model of polyetherimides that allows MD simulations on larger length scales and longer time scales. The second is to use machine-learning algorithms to train a predictive model of the glass transition temperature of polyimides. The predicted results of this model are then compared to experimental measurements and data from all-atom MD simulations. The last is to develop a Monte Carlo simulation method and software to study the polymerization process of branched polyetherimides. Below I introduce these approaches separately in a series.

1.3 Coarse-Grained Molecular Dynamics Simulations

All-atom MD simulations can yield details of atomistic and molecular motion and quantitative results of the properties of a system. However, the number of degrees of freedom is usually huge in an atomistic model. In order to simulate a larger system at longer time scales, a model with some details smoothed out is usually needed. In MD simulations, this technique is termed coarse-graining, where atoms are grouped and mapped to larger, coarse-grained beads. The underlying idea is that in the coarse-graining process, fast and unimportant degrees of freedom can be neglected and removed but other more relevant degrees of freedom are still kept. For example, when an atomistic model is used to model a polymer, there are usually a lot of hydrogen atoms connected to other much heavier atoms such as carbon, oxygen, and nitrogen. The bond between a hydrogen atom and a heavier atom typically has a high frequency but may not be a crucial factor controlling the physicochemical process of interest. If all hydrogen atoms are included in a simulation, then a small time step has to be used to resolve the dynamics of the bonds between hydrogen atoms and other heavier

atoms. In the so-called united-atom model, which is an example of a coarse-grained model, hydrogen atoms are not included explicitly but are combined with the heavy atoms that they are bonded to. By doing so, not only the number of atoms is reduced but also a larger time step can be used as the fast vibration modes related to the bonds between hydrogen and heavier atoms are removed. As a result, a larger system at a longer time scale can be treated after coarse-graining.

A clarification is needed here with regard to the exact meaning of "coarse-graining" and a coarse-grained model. To some extent, all physical models are coarse-grained models and all physical theories are coarse-grained theories. In an atomistic model of a material, the basic building blocks are atoms but they consist of electrons and nuclei, which are not treated explicitly in the model. A nucleus is in turn composed of neutrons and protons that are made of quarks. From the perspective of elementary particles, an atomistic model is therefore a coarse-grained model. In this thesis, a coarse-grained model is used in reference to a (more-detailed) atomistic model and coarse-graining refers to a process in which atoms are grouped into more coarse-grained beads and the interactions between those beads are parameterized on the basis of the atomistic model and possibly, experimental results.

Coarse-grained models have been used in molecular modeling for several decades. Levitt first used a simplified representation model to simulate the folding process of a protein.[40] Kremer and Grest used a bead-spring model to study the dynamics of polymer melts and illustrated the reptation process.[41] There are typically two approaches to develop a coarse-grained model, bottom-up and top-down. In a top-down approach, a coarse-grained model is constructed to reproduce some physical properties of a system and does not depend on a higher-resolution model (e.g., an atomistic model). In this approach, the interaction potentials are usually written down first in some functional forms and then the parameters are tuned to fit the computed results to experimental data. For example, Ouldridge employed a

top-down approach to construct a coarse-grained model of DNA.[42] A bead in such a grained model always represents a large group of atoms but its exact meaning can be obscure.

In recent years, bottom-up coarse-graining approaches have started to gain attention as they can be formulated rigorously on the basis of thermodynamics and statistical mechanics and may lead to a more systematic, hierarchical way of constructing coarse-grained models. These approaches usually start with a more-detailed model such as an atomistic model. A bottom-up approach can be structure-based such as in the iterative Boltzmann inversion,[43] force-based such as in the multi-scale coarse-graining approach and the force-matching method,[44–46] or pairwise-interaction-based such as in the conditional reversible work method.[47] In the following, these methods are briefly introduced. In Chapter 2 of this thesis, a bottom-up approach on the basis of the latter two methods will be used to develop a coarse-grained MD model of polyetherimides.

1.3.1 Iterative Boltzmann Inversion

The iterative Boltzmann inversion (IBI) method is an approach to calculate the potentials of mean force (PMFs) between coarse-grained beads using the structural information of a system computed in all-atom MD simulations. A PMF can be iteratively calculated using the following equation,

$$V_{i+1}(r) = V_i(r) - k_\mathrm{B}T \ln \frac{g_i(r)}{g_t(r)} \ , \tag{1.1}$$

where r is the distance between a pair of coarse-grained beads, V_i is the PMF of this pair after i iterative steps, $g_i(r)$ is the pair correlation function computed for this pair in a coarse-grained MD simulation with V_i, and $g_t(r)$ is the target pair correlation function determined in an all-atom MD simulation. For a pair of beads of type A and B, respectively, the physical

meaning of its pair correlation function, $g(r)$, is that $\rho g(r) \times 4\pi r^2 \mathrm{d}r$ is the average number of A beads in a spherical shell from r to $r + \mathrm{d}r$ around a B bead, or vice versa. Here ρ is the average number density of A beads (or B beads) in the system.

To implement the IBI method, an initial guess of the PMF is needed, which is usually taken as $V_0 = -k_\mathrm{B}T \ln g_t(r)$. After many iterations, the PMF is expected to converge and then $g_i(r) \simeq g_t(r)$, indicating that the pair correlation function (i.e., structure) is reproduced with the resulting coarse-grained model. The convergence behavior of this method was explained by Soper without a rigorous proof.[43] Recently Hanke also analyzed the convergence property of the IBI method.[48] Villa *et al.* discussed several cases where an IBI implementation may have difficulty to yield a converged PMF and they introduced entropic corrections to improve the convergence.[49]

1.3.2 Multiscale Coarse-Graining

The multiscale coarse-graining approach was developed by Izvekov and Voth on the principle of force-matching.[44] The coarse-grained force field is obtained by minimizing the following quantity

$$\chi^2 = \frac{1}{3NL} \sum_{l=1}^{L} \sum_{i=1}^{N} |\mathbf{F}_{il}^{ref} - \mathbf{F}_{il}^{p}(g_1, g_2, ..., g_M)|^2 \ , \tag{1.2}$$

where $\mathbf{F}_{il}^{ref}$ is the reference force on the i-th coarse-grained bead in the l-th configuration computed in an all-atom MD simulation, $\mathbf{F}_{il}^{p}(g_1, g_2, ..., g_M)$ is the coarse-grained force on the same bead with a set of parameters, g_1, g_2, ..., and g_M, N is the number of coarse-grained beads, and L is the number of configurations examined. The final coarse-grained force field is therefore $\mathbf{F}_{il}^{p}(g_1, g_2, ..., g_M)$ that minimizes χ^2. In 2008, Noid *et al.* derived the expression of $\mathbf{F}_{il}^{ref}$ on the basis of thermodynamics and statistical mechanics.[45, 50] This theory is the

foundation of our coarse-graining approach and will be discussed in detail in Sec. 2.2.

1.3.3 Conditional Reversible Work

A coarse-grained model developed with the two previous approaches is usually system and state dependent. When a new system or a different state of a system is considered, the coarse-grained model has to be completely re-parameterized. That is, the coarse-grained model is not expandable and transferable. The conditional reversible work (CRW) method has the potential of addressing these issues.[47] In CRW, the interactions between coarse-grained beads are assumed to be pairwise. In order to compute the PMF between a pair of coarse-grained beads, two all-atom MD simulations are run where constraints are introduced to fix the centers of mass of the two groups that the pair of beads represent. In one simulation, the nonbonded interactions between the two groups (i.e., the nonbonded interactions between atoms from one group and those from the other group) are turned off. In the other simulation, these nonbonded interactions are turned on. In each simulation, the forces on the two groups are computed and the differences of these forces in the two simulations are taken as the pairwise, coarse-grained force field between the two corresponding beads.

Hess *et al.* developed the LINCS algorithm to constrain the distance between atoms treated as mass points.[51, 52] Van der Vegt and collaborators applied the LINCS algorithm to their calculations of PMFs via CRW and introduced an entropic correction term to improve the results. [48, 53, 54] However, there is no rigorous derivation of this correction and in some of their coarse-graining efforts using CRW, the correction term is not used.[47, 55]

In Chapter 2 of this thesis, we show an entropic correction term is indeed needed for groups of atoms (e.g., a segment of a polymer chain) that have shapes very different from a sphere. The PMF between a pair of such groups computed directly or via CRW is generally too soft,

resulting in a much higher polymer density in the coarse-grained model. We show how a modified entropic correction term can be introduced to enable the coarse-grained model to capture the density of polyetherimides over a wide range of temperatures. As a result, the coarse-grained model can reasonably capture the mechanical properties of polyetherimides and is transferable.

1.4 Predicting Glass Transition Temperature of Polymers via Machine-Learning Algorithms

The glass transition temperature (T_g) is an important material parameter. It has a strong influence on thermal and mechanical properties, dictates the processing temperature, and determines the conditions of application (e.g., working temperature) of a material. Unsurprisingly, many simulation techniques including all-atom MD simulations have been applied to compute T_g of a polymer.[32, 56–65] However, the difference between a simulation result from the corresponding experimental value can be large.[66] The discrepancy can be ascribed to several reasons, including the accuracy of the adopted force field,[67] the large cooling rate used in simulations (which is intrinsically related to the accessible time-scale issue of all-atom MD simulations),[59, 68] the much shorter chains in simulations as compared to experimental systems,[58] etc. Furthermore, all-atom MD simulations can still be time-consuming as the physical property such as density that is used to determine T_g, has to be quantified at a series of temperatures spanning T_g.

For each class of polymers, there are often a lot of data already reported in the literature on T_g. It is therefore desirable if a structure-property relationship can be established for T_g using the available data. Machine-learning methods provide tools to achieve this goal as they

are designed to identify patterns and correlations from data. In fact, there exist many works on predicting $T_{\mathbf{g}}$ using machine-learning algorithms for various classes of polymers.[69–76] However, such a model does not exist for polyetherimides and more generally, polyimides. Furthermore, the existing models still have issues such as small training datasets, accuracy of predictions, stability of predictive models, and identification of polymer features that are strongly correlated to $T_{\mathbf{g}}$.

In Chapter 3 of this thesis, we first show that the diffusion coefficients of small gas molecules can be readily computed in all-atom simulations and used to determine T_{g}. This method may be even more efficient and accurate than the traditional method based on the density-temperature curve. Then we develop a predictive model of T_{g} by applying machine-learning algorithms to analyze the data of T_{g} for polyimides reported in the literature. We show that various strategies and practices can be adopted in the machine-learning process to improve the accuracy and stability of the predictive model.

1.5 Monte Carlo Simulations of Polymerization

Polymerization is the process in which monomers, oligomers, and polymer chains react to form longer chains. New covalent bonds are formed in this process, which makes it a challenging task to model a polymerization process via MD simulations. Most classical force fields for MD simulations are not reactive and have a difficulty of describing bond formation and breaking. Quantum mechanical calculations can treat bonding between atoms but are computationally extremely demanding. This is the reason that the hybrid quantum mechanics and molecular mechanics (QM/MM) methods are frequently used to study chemical reactions.[77–80] However, such methods are still limited to reactions between small molecules. In order to allow the simulations of reactions in a large system, van Duin *et*

al. developed a reactive force field for MD simulations on the basis of the bond-order concept.[81, 82] Yet, MD simulations using such a reactive force field are always 10 to 50 times slower than similar simulations with classical nonreactive force fields.[83] Other computational techniques such as Monte Carlo (MC) methods may be more suitable for studies of reaction and polymerization processes.

In general, there are two types of polymerizations. One is addition (chain-growth) polymerization, in which a chain grows longer by addition of monomers one by one.[84] Another is condensation (step-growth) polymerization where monomers, oligomer, and shorter chains can react with each other to form longer chains, usually with the release of small molecules when new bonds are formed. Up to present most MC studies of polymer formation are on addition polymerization.[85–92] Only a few MC simulations are reported on condensation polymerization.[93, 94] For example, Johnson and O'Driscoll applied the Gillespie algorithm to study condensation polymerization without considering the reaction time.[93] The system they studied is the $AA + BB + XY$ type where AA, BB and XY are monomers and A and B groups can only react with X or Y groups. The authors were interested in the distribution of A and B groups in the resulting polymer chains. Hillegers *et al.* used MC methods to model condensation polymerization but did not use the Gillespie algorithm.[94] In their approach, at each MC step two monomers were picked to react but the reaction may be accepted or rejected, which was decided according to a Metropolis-type criterion. This approach is generally not efficient compared to the approaches based on the Gillespie algorithm.

In Chapter 4 we develop a MC scheme based on the Gillespie algorithm to study the formation of branched polyetherimides, which is an example of condensation polymerization with branching. We show that the MC method can be used to quickly determine the molecular weight distribution of branched polymers under general conditions and the results agree with

the predictions of the Flory-Stockmayer theory below the gel point of a polymerizing system. Above the gel point, the theory fails while the MC scheme still yields correct results. The software we developed on the basis of this MC scheme is currently regularly used by SABIC engineers to predict the molecular weight distribution of branched polyetherimides prior to their laboratory efforts.

Chapter 2

Coarse-Grained Molecular Dynamics Modeling of Polyetherimides

This chapter is based on a paper to be submitted:

Chengyuan Wen, Roy Odle, and Shengfeng Cheng, "Coarse-Grained Molecular Dynamics Modeling of Polyetherimides".

I designed and built all the models for molecular dynamics simulations. I performed all the data analyses and prepared figures. Dr. Odle from SABIC contributed some ideas and provided experimental insights. All authors contributed to the writing of this paper. My contributions to this paper were under Dr. Cheng's supervision.

2.1 Introduction

Polyetherimides are an important class of engineering thermoplastics used in many different industries and applications because of their high heat resistance and stability, high strength

and moduli, excellent electrical properties over a wide range of temperatures and frequencies, good processability, good adhesive properties, and chemical stability.[2]. For example, Ultem is a family of polyetherimide products that has superior heat, solvent, and flame resistance and has been widely used in the automotive industry, medical and chemical instrumentation, and aerospace engineering.[95] There is a strong interest in discovering and synthesizing new polyetherimides that maintain the existing good properties but possess other desired properties and functions, such as better rheological properties when molded.[9, 96, 97] A laboratory trial and error, Edisonian approach is of course possible but becomes expensive and time-consuming when there is a need to scan a wide range of potential chemical formulae. To expedite materials discovery, computational approaches including molecular dynamics (MD) simulations have evolved into indispensable tools to study polyetherimides and other polymers.[98–101]

In the past several decades, all-atom MD simulations have been applied to compute the physical properties of polymers and elucidate the underlying mechanisms of many physicochemical processes.[68, 102–104] However, they are limited to relatively small systems and short time scales.[66] To access larger size and longer time scales, other computational techniques must be adopted. One approach is to employ a coarse-grained description of a polymer system to get rid of fast degrees of freedom that play a less important role in the physical properties and processes of interest. The typical practice is to group atoms into coarse-grained beads and to parameterize the interactions between the coarse-grained beads on the basis of all-atom simulations or available experimental data on material properties. Several coarse-graining methods, including iterative Boltzmann inversion,[43] force matching,[44, 105] and top-down methods,[106–108] have been developed but the resulting coarse-grained models are usually state-dependent and not transferable.[50]

In this chapter, we aim to develop a coarse-grained model for branched polyetherimides that

is transferable and expandable. Being transferable means that the model can be used at different state points, for example, at different temperatures. The expandability requirement indicates that when a new functional group (e.g., a side group, an ionic group, etc.) is added to polyetherimides for which a coarse-grained model is already available, we just need to expand the existing coarse-grained model by adding new beads corresponding to the newly added groups. It this is the case, then the task becomes parameterizing the interactions between these new beads and the existing beads. We demonstrate in this chapter that by combining a chemistry-informed grouping method, potential of mean force calculations, and entropic correction of a coarse-grained force field, we move closer to the ultimate goal of constructing a transferable and expandable coarse-grained model of polyetherimides.

This chapter is organized as follows. The general theory of coarse-graining, first systematically discussed by Noid and coworkers, and how it can be practically approximated and implemented, are introduced in Sec. 2.2. Then we apply this theory to develop a coarse-grained model of polyetherimides, as detailed in Sec. 2.3. In Sec. 2.4, the coarse-grained model is applied to compute mechanical, structural, and rheological properties of polyetherimides and the results are compared to those from all-atom MD simulations and experiments. Finally, conclusions are included in Sec. 2.5.

2.2 General Theory of Coarse-graining

In this chapter, the term "coarse-graining" refers to the process of deriving a coarse-grained description of a polymeric system from a fine-grained, atomistic model of the same system. In other words, atoms in the detailed description are grouped into coarse-grained beads (i.e., pseudo-atoms) with some fine details smoothed over. The coarse-graining procedure thus involves two key steps: the mapping from groups of atoms to coarse-grained beads and

the determination of interactions between the coarse-grained beads. We are motivated by the consideration that the final coarse-grained model being developed for the prototypical polyetherimide should be transferable, reusable, and expandable when a new formula of polyetherimide is the target of modeling. That is, when a new functional group (e.g., a new side group, or a new end-capping group) is added to the prototypical polyetherimide, we just need to add a new coarse-grained bead for the new functional group to the existing coarse-grained model and parameterize the interactions between this new bead and all existing coarse-grained beads, without completely reconstructing the coarse-grained model.

We will delay the discussion on the coarse-graining mapping to Sec.2.3.1 where our coarse-grained model of a branched polyetherimide is introduced. In this section, we introduce the general theory underlying the coarse-graining approach adopted here to determine the coarse-grained force field (i.e., the interaction potentials between the coarse-grained beads), which was originally proposed by Noid and coworkers.[45] We then discuss the implementation of this theory through potential of mean force (PMF) calculations for a pair of atomic groups. Technical issues in the implementation, including many-body effects on the coarse-grained potentials and how to fix the center of mass of a group of atoms in the PMF calculations, are discussed in detail. Finally, we present a test case with a pair of benzene molecules to validate the coarse-graining approach.

2.2.1 Theory of Noid and Coworkers

We first introduce the theory of coarse-graining formulated by Noid and coworkers.[45] Consider an atomistic system with n atoms. We map this system to a system of N coarse-grained beads through a mapping function $\boldsymbol{M}_{\boldsymbol{R}_I}(\boldsymbol{r}^n)$ for coordinates and $\boldsymbol{M}_{\boldsymbol{P}_I}(\boldsymbol{p}^n)$ for momenta, where $\boldsymbol{R}_I$ and $\boldsymbol{P}_I$ are coordinates and momenta of the coarse-grained beads while $\boldsymbol{r}_i$ and $\boldsymbol{p}_i$

are atomistic coordinates and momenta. Usually, linear mappings are used such as

$$\boldsymbol{M}_{\boldsymbol{R}_I}(\boldsymbol{r}^n) = \sum_{i=1}^{n} c_{Ii}\boldsymbol{r}_i \ , \tag{2.1}$$

and

$$\boldsymbol{M}_{\boldsymbol{P}_I}(\boldsymbol{p}^n) = M_I \sum_{i=1}^{n} c_{Ii}\boldsymbol{p}_i/m_i \ , \tag{2.2}$$

where I runs from 1 to N, the notation $\boldsymbol{r}^n$ ($\boldsymbol{p}^n$) represents the collection of atomistic coordinates (momenta), c_{Ii} are coefficients in the linear mapping, m_i is the mass of the i-th atom, and M_I is the mass of the I-th coarse-grained bead.

The Hamiltonian of the atomistic system is

$$h(\boldsymbol{r}^n, \boldsymbol{p}^n) = \sum_{i=1}^{n} \frac{1}{2m_i}\boldsymbol{p_i^2} + u(\boldsymbol{r}^n) \ , \tag{2.3}$$

where $u(\boldsymbol{r}^n)$ is the potential energy of the atomistic system. The coarse-grained Hamiltonian is

$$H(\boldsymbol{R}^N, \boldsymbol{P}^N) = \sum_{I=1}^{N} \frac{1}{2M_I}\boldsymbol{P}_I^2 + U(\boldsymbol{R}^N) \ , \tag{2.4}$$

where $\boldsymbol{R}^N$ ($\boldsymbol{P}^N$) represents the collection of coarse-grained coordinates (momenta) and $U(\boldsymbol{R}^N)$ is the potential energy of the coarse-grained system.

For the atomistic system held at a constant temperature and volume, its (equilibrium) physical properties are determined by the canonical probability distribution

$$p_{rp}(\boldsymbol{r}^n, \boldsymbol{p}^n) = p_r(\boldsymbol{r}^n)p_p(\boldsymbol{p}^n) \ , \tag{2.5}$$

where the configurational part is controlled by the Boltzmann factor via

$$p_r(\boldsymbol{r}^n) \propto \exp\left(-\frac{u(\boldsymbol{r}^n)}{k_\mathrm{B}T}\right) , \tag{2.6}$$

and the momentum part is

$$p_p(\boldsymbol{p}^n) \propto \exp\left(-\sum_{i=1}^{n}\frac{\boldsymbol{p}_i^2}{2m_i k_\mathrm{B}T}\right) . \tag{2.7}$$

Here k_B is the Boltzmann constant and T is temperature.

Similarly, for the coarse-grained system the canonical probability distribution function in the phase space reads

$$P_{RP}(\boldsymbol{R}^N, \boldsymbol{P}^N) = P_R(\boldsymbol{R}^N)P_P(\boldsymbol{P}^N) , \tag{2.8}$$

where the configurational part can be written as

$$P_R(\boldsymbol{R}^N) \propto \exp\left(-\frac{U(\boldsymbol{R}^N)}{k_\mathrm{B}T}\right) , \tag{2.9}$$

and the momentum part is

$$P_P(\boldsymbol{P}^N) \propto \exp\left(-\sum_{I=1}^{N}\frac{\boldsymbol{P}_I^2}{2M_I k_\mathrm{B}T}\right) . \tag{2.10}$$

The CG mapping yields a probability distribution function in terms of CG variables from

the atomistic probability distribution functions, which reads

$$\begin{aligned} p_{RP}(\boldsymbol{R}^N, \boldsymbol{P}^N) = & \int d\boldsymbol{r}^n \int d\boldsymbol{p}^n p_{rp}(\boldsymbol{r}^n, \boldsymbol{p}^n) \\ & \times \delta\left(\boldsymbol{M}_{\boldsymbol{R}}^N(\boldsymbol{r}^n) - \boldsymbol{R}^N\right) \delta\left(\boldsymbol{M}_{\boldsymbol{P}}^N(\boldsymbol{p}^n) - \boldsymbol{P}^N\right) \\ = & p_R(\boldsymbol{R}^N) p_P(\boldsymbol{P}^N) \ , \end{aligned} \tag{2.11}$$

where

$$p_R(\boldsymbol{R}^N) = \int dr^n p_r(\boldsymbol{r}^n) \delta\left(\boldsymbol{M}_R^N(r^n) - \boldsymbol{R}^N\right) \ , \tag{2.12}$$

$$p_P(\boldsymbol{P}^N) = \int dp^n p_p(\boldsymbol{p}^n) \delta\left(\boldsymbol{M}_P^N(p^n) - \boldsymbol{P}^N\right) \ , \tag{2.13}$$

and

$$\delta\left(\boldsymbol{M}_{\boldsymbol{R}}^N(\boldsymbol{r}^n) - \boldsymbol{R}^N\right) \equiv \prod_{I=1}^{N} \delta\left(\boldsymbol{M}_{\boldsymbol{R}I}(\boldsymbol{r}^n) - \boldsymbol{R}_I\right) \ , \tag{2.14}$$

$$\delta\left(\boldsymbol{M}_{\boldsymbol{P}}^N(\boldsymbol{p}^n) - \boldsymbol{P}^N\right) \equiv \prod_{I=1}^{N} \delta\left(\boldsymbol{M}_{\boldsymbol{P}I}(\boldsymbol{p}^n) - \boldsymbol{P}_I\right) \ . \tag{2.15}$$

A perfect coarse-grained model, which yields identical equilibrium properties as the atomistic model, thus requires

$$P_R(\boldsymbol{R}^N) = p_R(\boldsymbol{R}^N) \ , \tag{2.16}$$

$$P_P(\boldsymbol{P}^N) = p_P(\boldsymbol{P}^N) \ . \tag{2.17}$$

These two identities indicate that

$$\exp\left(-\frac{U(\boldsymbol{R}^N)}{k_{\mathrm{B}}T}\right) \propto \int dr^n \exp\left(-\frac{u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right) \delta\left(\boldsymbol{M}_R^N(r^n) - \boldsymbol{R}^N\right) \ , \tag{2.18}$$

and

$$\exp\left(-\sum_{I=1}^{N} \frac{\boldsymbol{P}_I^2}{2M_I k_{\mathrm{B}}T}\right) \propto \int dp^n \exp\left(-\sum_{i=1}^{n} \frac{\boldsymbol{p}_i^2}{2m_i k_{\mathrm{B}}T}\right) \delta\left(\boldsymbol{M}_P^N(p^n) - \boldsymbol{P}^N\right) \ . \tag{2.19}$$

Equation (2.18) can be used to determine the CG potential via

$$U(\boldsymbol{R}^N) = -k_{\mathrm{B}}T \ln z(\boldsymbol{R}^N) + \text{const} \ , \tag{2.20}$$

where

$$z(\boldsymbol{R}^N) \equiv \int d\boldsymbol{r}^n \exp\left(-\frac{u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right) \delta\left(\boldsymbol{M}_{\boldsymbol{R}}^N(\boldsymbol{r}^n) - \boldsymbol{R}^N\right) \ . \tag{2.21}$$

The coarse-grained force field, $\boldsymbol{F}_I(\boldsymbol{R}^N)$, is then given by

$$\begin{aligned} \boldsymbol{F}_I(\boldsymbol{R}^N) =& -\frac{\partial U(\boldsymbol{R}^N)}{\partial \boldsymbol{R}_I} \\ =& \frac{k_{\mathrm{B}}T}{z(\boldsymbol{R}^N)} \frac{\partial z(\boldsymbol{R}^N)}{\partial \boldsymbol{R}_I} \\ =& \frac{k_{\mathrm{B}}T}{z(\boldsymbol{R}^N)} \int d\boldsymbol{r}^n \exp\left(-\frac{u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right) \\ & \times \prod_{J \neq I} \delta\left(\boldsymbol{M}_{\boldsymbol{R}J}(\boldsymbol{r}^n) - \boldsymbol{R}_J\right) \frac{\partial}{\partial \boldsymbol{R}_I} \delta\left(\sum_{i \in \mathcal{I}_I} c_{Ii} \boldsymbol{r}_i - \boldsymbol{R}_I\right) \ , \end{aligned} \tag{2.22}$$

where the set $\mathcal{I}_I$ consists of the indices of a group of atoms that are mapped into the I-th coarse-grained bead. For the partial derivative, the following identity holds,

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{R}_I}\delta\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right) =& \frac{\partial\delta\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right)}{\partial\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right)}\frac{\partial\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right)}{\partial \boldsymbol{R}_I} \\ =& -\frac{\partial\delta\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right)}{\partial\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right)}\frac{\partial\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right)}{c_{Ik}\partial \boldsymbol{r}_k} \\ =& -\frac{\partial}{c_{Ik}\partial \boldsymbol{r}_k}\delta\Big(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\Big)\ , \end{aligned} \tag{2.23}$$

where $k \in \mathcal{I}_I$. After this transformation, we can do the integral on the right side of Eq.(2.22) by parts and the partial derivative $\partial/\partial \boldsymbol{r}_k$ is then acted upon $\exp\left(-\frac{u(\boldsymbol{r}^n)}{k_\mathrm{B}T}\right)\prod_{J\neq I}\delta\big(\boldsymbol{M}_{\boldsymbol{R}J}(\boldsymbol{r}^n) - \boldsymbol{R}_J\big)$. This partial derivative can be greatly simplified if we focus on atoms that are only mapped to the I-th coarse-grained bead but not other beads. That is, the index k is limited to $k \in \mathcal{I}_I$ but $k \notin \mathcal{I}_J$ for $J \neq I$. For such indices, we can introduce a nonzero factor d_{Ik} such that

$$\sum_{k\in\mathcal{S}_I} d_{Ik} = 1 \quad \text{for all } I\ , \tag{2.24}$$

where $\mathcal{S}_I$ is the set of indices for atoms that only belong to the I-th coarse-grained bead in the coarse-graining mapping. The combination of Eq. (2.23) and Eq. (2.24) yields the following identity,

$$\frac{\partial}{\partial \boldsymbol{R}_I}\delta\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right) = \sum_{k\in\mathcal{S}_I}\frac{d_{Ik}}{c_{Ik}}\frac{\partial}{\partial \boldsymbol{r}_k}\delta\left(\sum_{i\in\mathcal{I}_I} c_{Ii}\boldsymbol{r}_i - \boldsymbol{R}_I\right)\ . \tag{2.25}$$

Using Eq. (2.25), we can rewrite Eq. (2.22) as

$$\begin{aligned}
F_I(\boldsymbol{R}^N) =& -\frac{k_{\mathrm{B}}T}{z(\boldsymbol{R}^N)}\int d\boldsymbol{r}^n \exp\left(\frac{-u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right) \\
&\times \prod_{J\neq I}\delta\left(\boldsymbol{M}_{\boldsymbol{R}J}(\boldsymbol{r}^n)-\boldsymbol{R}_J\right)\sum_{k\in\mathcal{S}_I}\frac{d_{Ik}}{c_{Ik}}\frac{\partial}{\partial \boldsymbol{r}_k}\delta\left(\sum_{i\in\mathcal{I}_I}c_{Ii}\boldsymbol{r}_i-\boldsymbol{R}_I\right) \\
=& -\sum_{k\in\mathcal{S}_I}\frac{d_{Ik}}{c_{Ik}}\frac{k_BT}{z(\boldsymbol{R}^N)}\int d\boldsymbol{r}_{\overline{k}}^{n-1}\prod_{J\neq I}\delta(\boldsymbol{M}_{\boldsymbol{R}J}(\boldsymbol{r}^n)-\boldsymbol{R}_J)\int d\boldsymbol{r}_k \\
&\exp\left(\frac{-u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right)\frac{\partial}{\partial \boldsymbol{r}_k}\delta\left(\sum_{i\in\mathcal{I}_I}c_{Ii}\boldsymbol{r}_i-\boldsymbol{R}_I\right)\ ,
\end{aligned} \tag{2.26}$$

where $d\boldsymbol{r}_{\overline{k}}^{n-1}\equiv\prod_{j=1,j\neq k}^{n}d\boldsymbol{r}_j$. An integration by parts yields

$$\begin{aligned}
F_I(\boldsymbol{R}^N) =& \sum_{k\in\mathcal{S}_I}\frac{d_{Ik}}{c_{Ik}}\frac{k_BT}{z(\boldsymbol{R}^N)}\int d\boldsymbol{r}_{\overline{k}}^{n-1}\prod_{J\neq I}\delta\left(\boldsymbol{M}_{\boldsymbol{R}J}(\boldsymbol{r}^n)-\boldsymbol{R}_J\right)\int d\boldsymbol{r}_k \\
&\delta\left(\sum_{i\in\mathcal{I}_I}c_{Ii}\boldsymbol{r}_i-\boldsymbol{R}_I\right)\frac{\partial}{\partial \boldsymbol{r}_k}\exp\left(\frac{-u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right) \\
=& -\sum_{k\in\mathcal{S}_I}\frac{d_{Ik}}{c_{Ik}}\frac{1}{z(\boldsymbol{R}^N)}\int d\boldsymbol{r}^n\prod_{J}\delta\left(\boldsymbol{M}_{\boldsymbol{R}J}(\boldsymbol{r}^n)-\boldsymbol{R}_J\right) \\
&\exp\left(\frac{-u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right)\frac{\partial u(\boldsymbol{r}^n)}{\partial \boldsymbol{r}_k} \\
=&\frac{1}{z(\boldsymbol{R}^N)}\int d\boldsymbol{r}^n\exp\left(\frac{-u(\boldsymbol{r}^n)}{k_{\mathrm{B}}T}\right)\delta\left(\boldsymbol{M}_{\boldsymbol{R}^N}(\boldsymbol{r}^n)-\boldsymbol{R}^N\right) \\
&\sum_{k\in\mathcal{S}_I}\left(-\frac{d_{Ik}}{c_{Ik}}\right)\frac{\partial u(\boldsymbol{r}^n)}{\partial \boldsymbol{r}_k} \\
=&\langle\sum_{k\in\mathcal{S}_I}\left(-\frac{d_{Ik}}{c_{Ik}}\right)\frac{\partial u(\boldsymbol{r}^n)}{\partial \boldsymbol{r}_k}\rangle_{\boldsymbol{R}^N} \\
=&\langle\sum_{k\in\mathcal{S}_I}\frac{d_{Ik}}{c_{Ik}}\boldsymbol{f}_k(\boldsymbol{r}^n)\rangle_{\boldsymbol{R}^N} \\
=&\langle\boldsymbol{\mathcal{F}}_I(\boldsymbol{r}^n)\rangle_{\boldsymbol{R}^N}\ .
\end{aligned} \tag{2.27}$$

This equation is the theoretical foundation of computing the coarse-grained force field.

Here the ensemble average is performed under the coarse-graining mapping summarized in Eq. (2.1). In the mapping adopted here, all atoms are separated into nonoverlapping groups, each of which is mapped to one and only one coarse-grained bead. Furthermore, the coarse-grained bead is placed at the center of mass of the group of atoms that it represents. In this case, $c_{Ik} = \frac{m_k}{\sum_{k \in \mathcal{S}_I} m_k}$ and if $k \in \mathcal{S}_I$, then $k \notin \mathcal{S}_J$ for all $J \neq I$. A natural choice is $d_{Ik} = c_{Ik}$, which indicates that the force on the I-th coarse-grained bead is the ensemble average of the forces on atoms with indices $\in \mathcal{S}_I$ from all other atoms in the system, subjected to the mapping from atoms to beads. Noid *et al.* further proved that this scheme also guarantees that the momentum part of the phase-space probability distribution function in the coarse-grained model matches that in the atomistic model.[45]

To summarize, the derivation in this section shows that the coarse-grained model and the atomistic model have consistent probability distribution functions of states in the phase space as long as

- One group of atoms is mapped to one coarse-grained bead.
- No atom is shared by more than one group.
- A coarse-grained bead is placed at the center of mass of the group of atoms that the bead represents.
- The force on a coarse-grained bead is the ensemble average of the forces exerted on all atoms in the group, which the bead represents, by all other atoms in the atomistic system.

The coarse-graining method on the basis of these constraints is called force matching coarse-graining.[109]

2.2.2 Pairwise Nonbonded Coarse-grained Force Field

The general theory of coarse-graining presented in the previous section is valid in any system. However, the resulting coarse-grained force field is a many-body force field, which is very difficult to compute in general. Furthermore, the coarse-grained model is state-dependent as the coarse-grained force field is computed for a given thermodynamic state, as shown in Eq. (2.27). Therefore, the coarse-grained model developed using the general theory is actually not transferable, though it is proved to be self-consistent and rigorous from the perspective of thermodynamics.[44]

One approximated coarse-graining scheme was proposed by Wang *et al.*[46] Their effective-force coarse-graining (EF-CG) method results in pairwise coarse-grained force field. In EF-CG, an atomic system is divided into groups of atoms, with each group mapped to a coarse-grained bead. The interaction between a pair of beads is computed from the average force between the corresponding atomic groups, under the presence of all other atoms in the atomistic system. The advantage of this scheme is that the many-body nature of the coarse-grained force field is naturally captured. However, the implementation of EF-CG requires MD simulations of the entire atomistic system and nontrivial constraints on the pair of atomic groups under consideration. As a result, the resulting coarse-grained model is not transferable and a new model has to be developed when a new polyetherimide is studied.

In this chapter, we aim to develop a coarse-grained model for polyetherimides that is transferable and expandable. In particular, the former feature indicates that the model can be used to simulate polyetherimides under different temperatures and the latter means that when a new polyetherimide containing one or several new functional groups is dealt with, the same number of new coarse-grained beads will be added to the coarse-grained model. The only parameterization that is needed to update the coarse-grained model is to compute

the interactions between the new beads and all existing beads, i.e., the coarse-grained force field is expanded with the newly added beads. We term this method a library-like approach, which is obviously appealing as it avoids the need to construct a new coarse-grained model every time when the polymer of interest is updated. However, the theoretical foundation of such an approach is not well understood at the moment.[50] In this chapter we provide evidences demonstrating that such an approach may be viable.

The simplest system is one that can be divided into two groups of atoms and can be coarse-grained into two beads. The force on one bead is then the sum of the forces on all atoms in the group this bead represents by all atoms in the other group. If the center of mass of each group is chosen as the location of the corresponding coarse-grained bead, then the coarse-grained model with a pairwise interaction is fully consistent with the general theory of coarse-graining discussed by Noid *et al.*. This is easy to understand as in such a simplified system, there is no many-body effect since there are only two coarse-grained beads anyway. In a more general case, there are of course more than two beads in the coarse-grained model. Then parameterizing the coarse-grained force field in a pairwise fashion assumes the force field is additive and automatically neglects its intrinsic many-body nature as required by the thermodynamic consistency between the atomistic and coarse-grained models. In this chapter we show that an entropic correction term can be added to the pairwise coarse-grained force field derived with PMF calculations to not only make the model transferable but also effectively compensate for the error introduced by neglecting many-body effects,

2.2.3 Fixing the Center of Mass of a Group of Atoms

When computing the force between two groups of atoms, one technical key is to fix the center of mass of each group, which will then allow the separation between the two centers of mass

to be used as a coarse-grained coordinate. Fritz *et al.* used the LINCS constraint algorithm to fix the center of mass of a group of atoms and compute the coarse-grained force field between atomic groups.[110] Here we adopt a different approach implemented in LAMMPS. In the starting configuration, the velocity of the center of mass of each group of atoms is set to zero (i.e., the total momentum of the group is set to zero). When a group of atoms interact with atoms from other groups, all atoms first move according to the Newtonian equation of motion. After one time step in MD simulations, the atomic coordinates are updated and the new location of the center of mass of each group is computed. The displacement vector, $\vec{d}_c$, of a group's center of mass from its original location to the new one is determined as well as its velocity, $\vec{v}_c$. Then all atoms in that group are displaced by $-\vec{d}_c$ such that the center of mass of the group is shifted back to its original location. At the same time, $\vec{v}_c$ is subtracted from the velocity of each atom in the group to ensure that the velocity of the group's center of mass is restored to zero. To justify this "recentering" approach, in Appendix A we prove that it is equivalent to the method where a constraint force is added to each atom in a group such that the group's center of mass does not move (i.e., the total force including the constraint forces on the group is zero) and the group does not exhibit any artificial rotation (i.e., the net torque from the constraint forces on the group is zero but there can be torques from interactions with other groups). In this chapter, we employ the "recentering" approach to constrain the center of mass of a group of atoms.

2.2.4 Potential of Mean Force Calculations: Test

To demonstrate the method of using PMF calculations to approximate the ensemble average in Eq. (2.27) that connects the coarse-grained force field to the atomistic force field, we utilize a model system that consists of one benzene molecule and one oxygen atom, as shown in Fig. 2.1(a). A configuration of this model system can be approximately characterized

by three parameters, the magnitude of the vector $\vec{r}$ pointing from the benzene's center of mass to the oxygen and two angles that describe the orientation of $\vec{r}$ relative to the benzene molecule. A Cartesian coordinate system can be set up using the benzene's center of mass as its origin and three orthonormal vectors, $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$. Among these, $\vec{v}_1$ and $\vec{v}_2$ define the plane in which the benzene molecule lies in and $\vec{v}_3$ is normal to this plane. At a finite temperature, the six carbon atoms and six hydrogen atoms in the benzene molecule actually have a three-dimensional conformation. However, we can always define a plane that captures the planar nature of the benzene molecule. For example, a plane that minimizes the sum of the square distances or has zero average distance of the carbon atoms from this plane can be used. The location of the oxygen atom in this coordinate system is therefore given by the vector $\vec{r}$, i.e., its length $r \equiv |\vec{r}|$ as well as two angles: the polar angle ω and the azimuthal angle ϕ, as shown in Fig. 2.1(a).

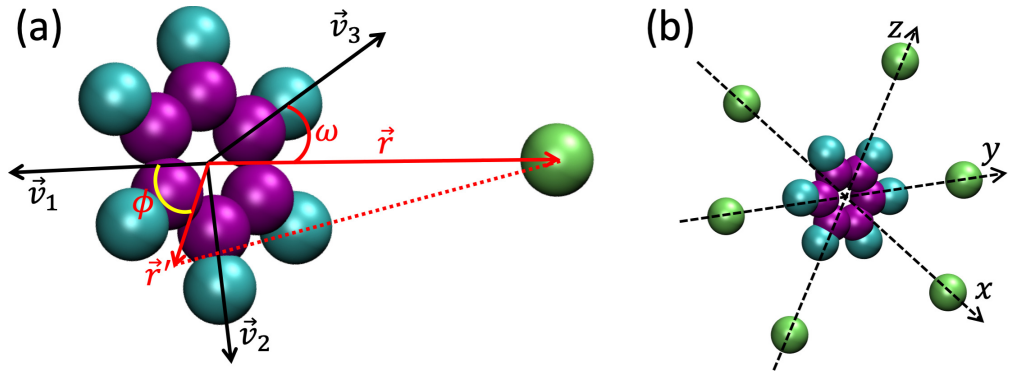

Figure 2.1: (a) A snapshot of a model system consisting of one benzene molecule and one oxygen atom. (b) A snapshot of a model system consisting of one benzene molecule at center and six surrounding oxygen atoms.

In this chapter, all atomistic systems were built using MAPs.[111] Bond increment charges were used to set atomic charges [112] and PCFF force field was adopted for all-atom MD simulations.[36] The equation of motion was integrated with the velocity-Verlet algorithm with a time step set as 1 fs. The cutoff of both Van der Waals and Coulomb interactions was set as 12 Å. The particle-particle particle-mesh (PPPM) method was used to calculate

the long-range part of Coulomb interactions. When needed, a Langevin thermostat and Berendsen barostat was used to control temperature and pressure, respectively.

The "recentering" method of fixing a center of mass discussed previously is applied to the benzene-oxygen system and used to fix r at any chosen value that is physically allowed. At a given r, we use all-atom MD simulations to sample various configurations parameterized by ω and ϕ and compute the probability density, $P(\omega, \phi)$ of a given configuration. Considering the rotational symmetry of a benzene molecule around $\vec{v}_3$, we investigate the probability density integrated over ϕ, i.e., $P(\omega) \equiv \frac{1}{2\pi} \int_0^{2\pi} P(\omega, \phi) d\phi$.

We compute $P(\omega)$ in various ways. First, it is directly estimated using the trajectory (i.e., a series of configurations) generated by a MD simulation at a constant temperature T. Namely, $P(\omega)d\omega = \frac{\#[\omega, \omega+d\omega]}{\text{total number of states}}$, where $\#[\omega, \omega + d\omega]$ is the number of states with the polar angle in $[\omega, \omega + d\omega]$. Secondly, the energy of each configuration, $\epsilon(\omega, \phi)$, from the MD simulation is used to compute $P(\omega)$ through the canonical distribution, $P(\omega) = \frac{\sum_\phi e^{-\beta\epsilon(\phi,\omega)} \sin\omega}{\sum_\omega \sum_\phi e^{-\beta\epsilon(\phi,\omega)} \sin\omega}$, where $\beta = 1/(k_{\mathrm{B}}T)$ and k_{B} is the Boltzmann constant. Thirdly, we compute $\epsilon(\phi, \omega)$ for a series of static configurations of the benzene-oxygen system. Noting the symmetry of the system, the polar angle ω is varied from 0 to $\pi/2$ in steps of $\pi/90$. For each ω, the azimuthal angle ϕ is varied from 0 to π in increments of $\pi/90$. For each configuration, the benzene molecule is in its ground-state conformation (i.e., a planar hexagonal configuration) and the interaction energy between benzene and oxygen, $\epsilon(\omega, \phi)$, is computed. The canonical distribution is then used to determine $P(\omega)$. In the final way, $P(\omega)$ is determined with the trajectory generated from a MD simulation of a model system that consists of one benzene molecule at center and six surrounding oxygen atoms, as shown in Fig. 2.1(b). This system is designed to speed up the sampling of various configurations, in particular, rare configurations that may have significant contributions to an ensemble average when r is small. In this case, extremely long MD simulations are needed if a single pair of benzene and oxygen is used. Later in

this chapter we show that this strategy of using six pairs simultaneously is practically useful when PMF calculations are performed for groups of atoms of which the conformations deviate significantly from a sphere.

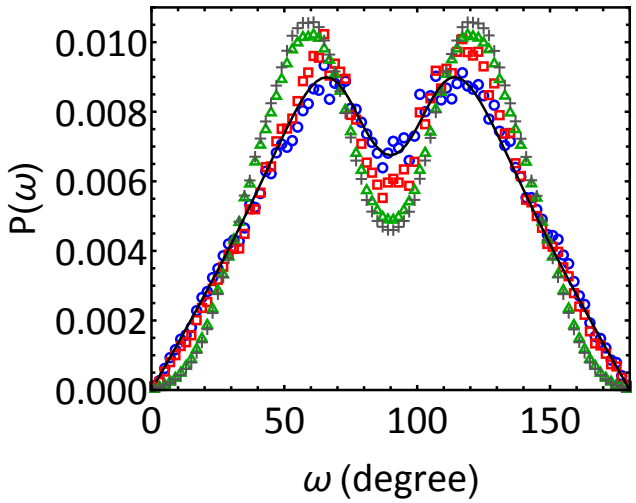

Figure 2.2: Comparison of the probability density, $P(\omega)$, calculated in various ways for benzene-oxygen systems. The data are based on MD trajectory (◯) and energy (□) of one benzene-oxygen pair at $T = 300$ K, MD trajectory (△) and energy (+) of the one benzene/six oxygen system at $T = 300$ K, and static configuration energy (solid line) of one benzene-oxygen pair (effectively at $T = 0$ K).

The results of $P(\omega)$ calculated in the various ways discussed previously are shown in Fig. 2.2. For both one benzene-oxygen pair and one benzene/six oxygen system, the results based on the MD trajectory (circles in Fig. 2.2) and the corresponding energy (squares in Fig. 2.2) are close but differ slightly around $\omega = 60°$, 90°, and 120°. This difference is likely due to the limited number of configurations sampled in the MD simulation as it can be noted that the difference between the two (triangles and pluses in Fig. 2.2) is much smaller for the one benzene/six oxygen system in Fig. 2.1(b), which can sample more configurations in the same number of MD time steps. The results computed using the energies of uniformly-scanned static configurations (solid line in Fig. 2.2), on the other hand, show a very good agreement with those based on the MD trajectory of one benzene-oxygen pair. However, this agreement may just be a coincidence as the energy of a static configuration is essentially the energy of

the system in that configuration at 0 K and it is unclear why it seems to produce reasonable results when that energy is used to compute the Boltzmann factor at 300 K. The results of the system with six benzene-oxygen pairs are in reasonable agreement with those from one pair, though the former seem to slightly overestimate (oversample) configurations for $30^\circ \lesssim \omega \lesssim 60^\circ$ and $120^\circ \lesssim \omega \lesssim 150^\circ$, while underestimate (undersample) configurations in the remaining ranges of ω. Although the discrepancy is noted, we will utilize the six-pair setup in Fig. 2.1(b) to speed up the calculation of pairwise nonbonded interactions between a pair of coarse-grained beads when developing a coarse-grained model of polyetherimides. The error introduced by this choice will be balanced out when entropic corrections are included in the coarse-grained force field, as discussed later.

2.3 Development of a Coarse-Grained Model of Polyetherimides

We employ the methodology discussed in the previous section to develop a coarse-grained model of branched polyetherimide chains and use this model to compute their mechanical, structural, and rheological properties. The protocol of developing such a model is outlined in Fig. 2.3, which mainly includes three steps: the grouping of atoms into coarse-grained beads, the parameterization of bonded (i.e., bond, angle, dihedral, etc.) interactions, and the parameterization of nonbonded interactions between the coarse-grained beads. In the following sections we describe each step in detail.

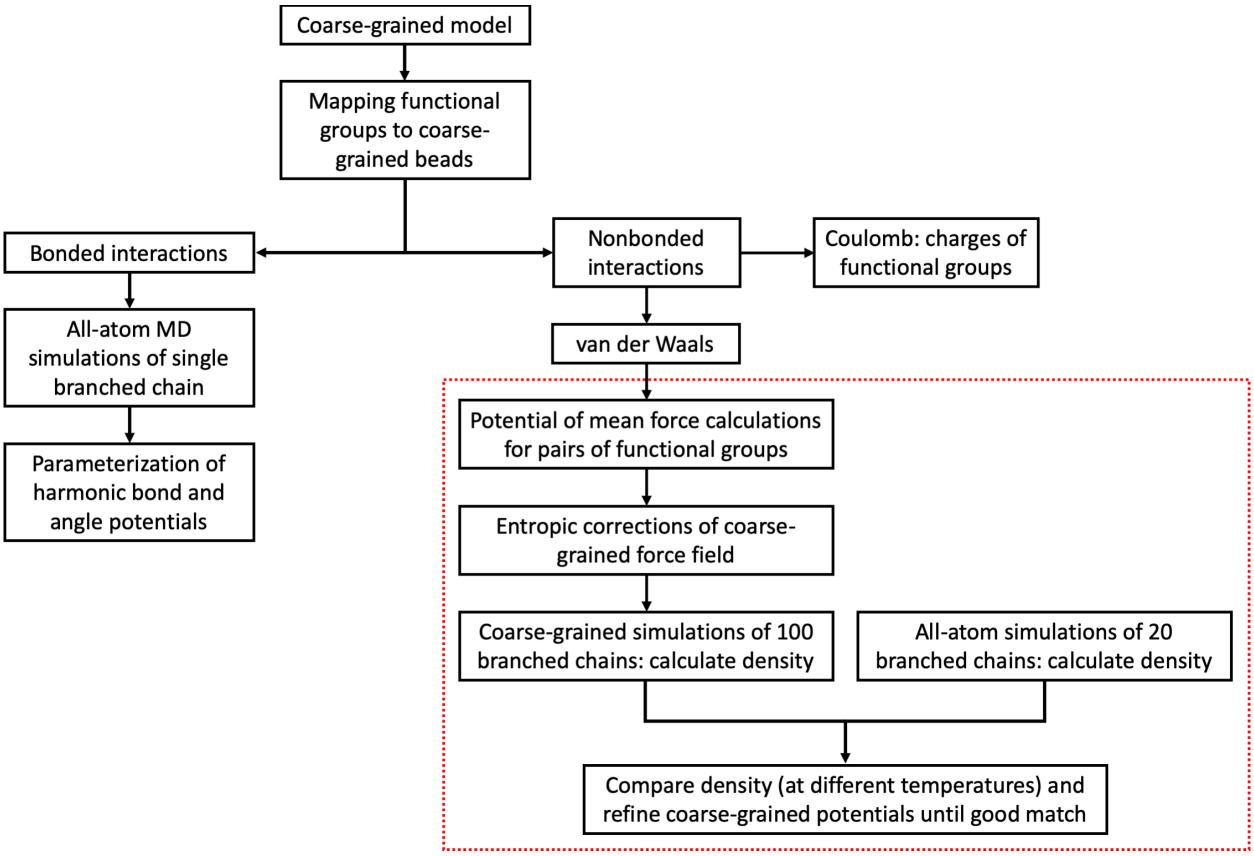

Figure 2.3: The coarse-graining flowchart.

2.3.1 Mapping Groups of Atoms into Coarse-Grained Beads

The branched polyetherimides dealt with in this chapter are polymerized from two backbone monomers [4,4'-bisphenol A dianhydride (BPADA) and m-phenylenediamine (MPD)], a chain terminator [phthalic anhydride (PA)], and a tri-functional branching agent [tris[4-(4-aminophenoxy)phenyl] ethane (TAPE)]. One short branch, terminated on the end with PA and connected to TAPE on its other end, is shown in Fig. 2.4. After several attempts, we settle with using 5 types of coarse-grained beads for such a system: type-A beads for phthalimide groups, type-B beads for oxygen atoms, type-C beads for the bisphenol groups, type-D beads for the aromatic rings on the phenylenediamine groups, and type-E bead for

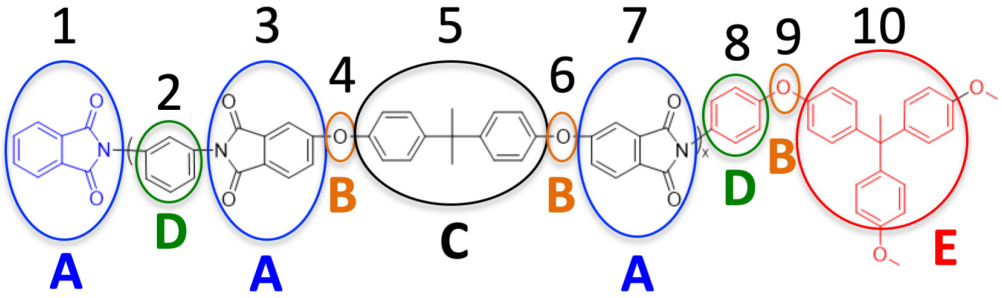

Figure 2.4: Mapping atomic groups into coarse-grained beads. For clarity, only one short branch connected to TAPE and terminated with PA is shown.

the core part of TAPE. The groupings are shown schematically in Fig. 2.4. In total there are 10 atomic groups in this short branch. Groups 1 is chemically almost the same as groups 3 and 7 except that there is one extra hydrogen atom in group 1 as it is at the end of the branch. The mass and charge of group 1 are therefore slightly different from those of groups 3 and 7. However, we map all three groups to type-A beads (with different masses and charges) to simplify the non-Coulombic part of the coarse-grained force field. The error of this treatment is partially compensated for by the entropic correction of the coarse-grained force field introduced later. Groups 2 and 8 are chemically almost identical except for the location of one hydrogen atom. As a result, they have the same mass but their charges are somewhat different. For simplicity, we map these groups into type-D beads with different charges. The charges and masses of all atomic groups are summarized in Table 2.1.

Table 2.1: Charges and masses of atomic groups defined in Fig. 2.4.

Coarse-grained Bead	Atomic Group Index	Charge (e)	Mass (10^{-25} g)
type-A	1	0.05	2426.46
	3, 7	0.0765	2409.72
type-B	4, 6, 9	-0.053	265.67
type-C	5	0.053	3226.05
type-D	2	-0.1	1263.64
	8	-0.0235	
type-E	10	0.0795	4240.02

We adopt the chemistry-informed mapping scheme in Fig. 2.4 because it leads to unimodal

distributions for all bonds and angles between the coarse-grained beads and therefore a Gaussian approximation can be used to derive the stiffness constants of bonds and angles, as discussed below. Other mapping schemes, including those more coarsened, usually cause the angle potential to have two or more local minima, which is not desirable for the parameterization of the bond and angle interactions between the coarse-grained beads.

2.3.2 Parameterization of Coarse-Grained Bonded Interactions

The second step of coarse-graining is to parameterize the bond and angle interactions in the coarse-grained chain using results from all-atom MD simulations. For two neighboring groups of atoms, each mapped to a coarse-grained bead, we compute the distance between their centers of mass and examine the probability distribution of the distances. For three consecutive groups of atoms, we compute the angle formed by the corresponding centers of mass and examine the probability distribution of the angles. All probability distributions based on the grouping scheme in Fig. 2.4 are well approximated by Gaussian distributions, indicating that the bonds and angles can be described by harmonic potentials as in

$$U(x) = \frac{1}{2}k_x(x - x_0)^2 \ , \tag{2.28}$$

and the corresponding probability density is

$$p(x) = \sqrt{\frac{k_x}{2\pi k_{\mathrm{B}}T}}e^{-\frac{U(x)}{k_{\mathrm{B}}T}} \ , \tag{2.29}$$

where x is the length r for a bond and the angle θ for an angle, x_0 the equilibrium bond length or angle, k_x the corresponding stiffness, and $U(x)$ the harmonic bond or angle potential.

Figure 2.5 shows the probability distributions of the 2-3-4 angle and the 6-7-8 angle (see

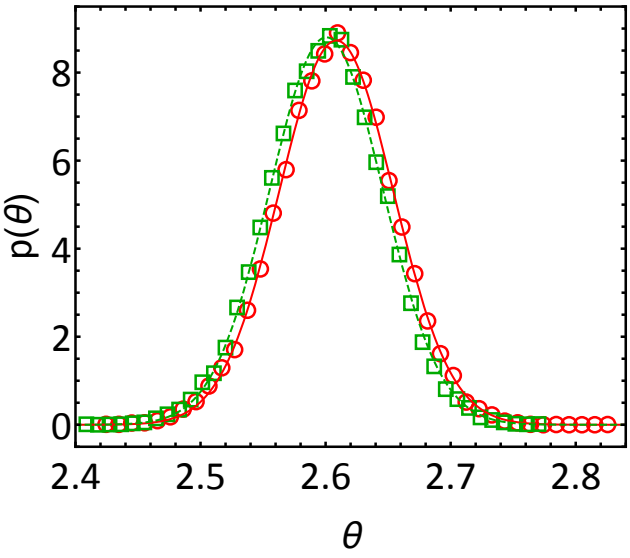

Figure 2.5: The probability distribution of the 2-3-4 angle at $T = 450$ K: all-atom MD data (circles) and a Gaussian fit (solid line) with k_θ=428 kcal/mol/rad^2 and θ_0=2.61 rad; of the 6-7-8 angle at the same T: all-atom MD data (squares) and a Gaussian fit (dashed line) with k_θ=435 kcal/mol/rad^2 and θ_0=2.60 rad.

Fig. 2.4 for the definition of bead types and atomic group indices), which are both D-A-B type angles and are well fit by a Gaussian distribution. The fitting allows us to extract a spring constant and an equilibrium angle for the harmonic angle potential. The spring constants and equilibrium values for the same type angle are very close, as shown in Fig. 2.5. Similar results are obtained for all bonds and angles. The average spring constants, equilibrium bond lengths, and equilibrium angles for all bonds and angles are summarized in Table 2.2 and Table 2.3. The results validate the usage of harmonic potentials for the bond and angle interactions in the coarse-grained force field. For the same type bonds and angles appearing more than once even in one branched chain, they all have similar stiffnesses and equilibrium values. This finding justifies the simplification of mapping 10 atomic groups into 5 types of coarse-grained beads. We further confirm that the parameters of the coarse-grained bond and angle potentials are insensitive to temperature for the range of temperatures of interest here.

Table 2.2: Coarse-grained bond parameters.

Bond Type	Bonds	k_r (kcal/mol/Å^2)	r_0 (Å)
A-D	1-2	267	4.86
	2-3	255	4.87
	7-8	286	4.91
A-B	3-4	381	4.09
	6-7	400	4.07
B-C	4-5	88	4.99
	5-6	83	5.04
B-D	8-9	505	2.80
B-E	9-10	80	5.58

Table 2.3: Coarse-grained angle parameters.

Angle Type	Angles	k_θ (kcal/mol/rad^2)	θ_0 (rad)
A-D-A	1-2-3	254	2.10
D-A-B	2-3-4	427	2.62
	8-7-6	425	2.61
A-B-C	3-4-5	74	2.17
	7-6-5	44	2.07
B-C-B	4-5-6	124	1.70
A-D-B	7-8-9	108	2.23
D-B-E	8-9-10	507	3.11
B-E-B	9-10-9	79	2.20

2.3.3 Parameterization of Coarse-Grained Nonbonded Interactions

2.3.3.1 Potential of Mean Force Calculations

The next step of constructing the coarse-grained model is to parameterize the nonbonded interaction between a pair of beads that belong to the same chain but are separated by at least three bonds or that belong to different chains. In each PMF calculation, the centers of mass of the two groups of atoms are fixed but the atoms in each group are still mobile, causing each group to rotate and vibrate around its center of mass. A Langevin thermostat is used to keep the system at a target temperature. By sufficiently sampling the relative configurations

and orientations of the two groups, we calculate the average force between them as a function of the separation between their centers of mass. The results show that the average force is along the vector connecting the two centers of mass. Therefore the separation between two centers of mass can be used as a coarse-graining coordinate. Integrating the average force over separation, we obtain the coarse-grained potential for each pair of coarse-grained beads.

Since atoms carry charges in an all-atom model, we split the coarse-grained potential into two parts: the Coulomb component and the van der Waals component. The Coulomb interaction between two atomic groups is included in the coarse-grained force field using the total charge of a group as the charge of the corresponding coarse-grained bead. The charges of all atomic groups for branched polyetherimides can be found in Table 2.1. If one or two groups from the pair being parameterized are charged, the Coulomb force is subtracted from the mean force between the two groups, resulting in a van der Waals component that is still along the vector connecting the centers of mass. The integration of this component over separation is called the nonbonded, van der Waals PMF. Its mathematical expression is

$$U(r) = \int_{r_m}^{r} \langle f_c \rangle_s ds, \tag{2.30}$$

where r_m is a large separation at which $U(r_m) \simeq 0$, and $\langle f_c \rangle_s$ is the force between the two groups at separation s with the Coulomb force between them subtracted. Below we reserve PMF specifically for the van der Waals component of the nonbonded interactions between two atomic groups.

As an example, Fig. 2.6 shows the results of PMF for a pair of benzene molecules. Two sets of results are included. One is computed from all-atom MD simulations with a single pair of atomic groups, similar to the setup shown in Fig. 2.1(a). The data show the typical feature of intermolecular interactions, i.e., the force is attractive at large and repulsive at

short separations, with a well-defined minimum at an equilibrium separation around 5 to 6 Å.

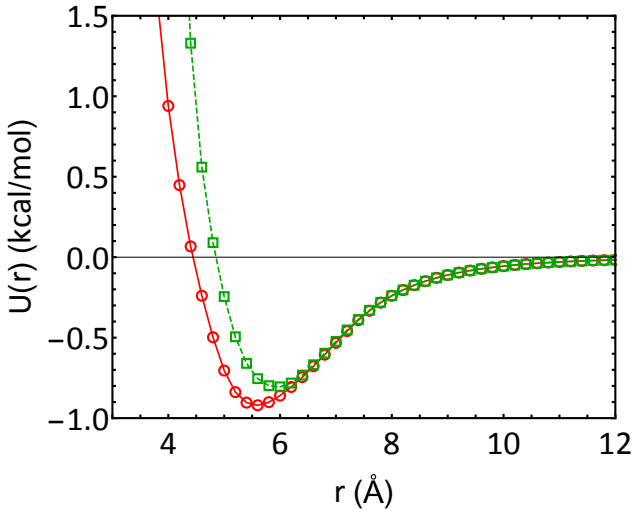

Figure 2.6: The van der Waals PMF, $U(r)$, as a function of separation r for a benzene dimer (i.e., a pair of benzene molecules). The results are from all-atom MD simulations with a single pair (◯) and a system with one group at center and six surrounding groups (□, see Fig. 2.1(b) for the setup). The lines are guides to the eye.

Using the setup in Fig. 2.1(a), we have parameterized the PMFs for all 15 pairs of atomic groups for polyetherimides. For atomic groups carrying different charges but mapped to the same type of coarse-grained beads (with the corresponding charges), the results show that after the subtraction of Coulomb interactions, the van der Waals PMFs are similar. This is the reason that only 15 PMFs are included in the coarse-grained force field. However, a large discrepancy is observed between the coarse-grained model and the all-atom model. Under standard conditions for temperature and pressure, the mass density of polyetherimides is 1.209 g/cm^3 but the density from the coarse-grained model, without any modification as discussed below, is much higher at 1.622 g/cm^3. This difference reflects a generic "softness" issue of a coarse-grained model constructed via a bottom-up approach on the basis of PMF calculations,[113] i.e., the coarse-grained potentials are usually much softer than the atomistic ones, making the coarse-grained system to have a much larger density than the atomistic

system.

We can illustrate the "softness" issue more clearly with benzene. Under standard conditions, the mass density of benzene at 300 K is 0.795 g/cm^3 from all-atom MD simulations, which is only slightly lower than the experimental value of 0.876 g/cm^3. During coarse-graining, each benzene is grouped into a bead and the van der Waals PMF between beads serves as the coarse-grained force field. If this PMF is computed with a setup similar to Fig. 2.1(a) where only a single pair of benzene molecules is utilized, we find that the mass density from the coarse-grained model is much higher at 1.579 g/cm^3.

The "softness" issue is related to the insufficient sampling of relative configurations in a PMF calculation. For a benzene dimer at large separations, their conformations are not strongly correlated and an all-atom MD simulation can sufficiently sample all possible configurations. However, when they get close, the two benzene molecules prefer to be in the T-shaped or parallel-displaced configurations, which are energetically favored.[114] In the PMF calculation of benzene-benzene interactions with a benzene dimer, the contributions of these configurations dominate, where the two benzene molecules are close to each other. However, in an atomistic model of benzene system, the local packing of two benzene molecules is affected by other surrounding molecules and cannot all assume the lowest-energy configurations. As a result, the average separation between adjacent benzene molecules is larger than the separation at which the PMF calculated with a single pair has a minimum. Furthermore, in a polymeric material containing aromatic rings, the rings are connected to other atomic groups. The T-shaped or parallel-displaced configurations are still favored by the aromatic rings but are subjected to the constraints set by the presence of other groups. As a result, a pair of aromatic rings cannot be as close as in the situation where only two rings are present. In this sense, the "softness" issue is the outcome of using pairwise nonbonded interactions between coarse-grained beads to approximate the many-body interactions among atomistic

groups.

To overcome the "softness" problem, we resort to the setup illustrated in Fig. 2.1(b). To compute the PMF between a pair of atomic groups, we place the center of mass of one group at the origin of a Cartesian coordinate system (xyz), replicate the other group six times, and place the six groups around the central group on x, $-x$, y, $-y$, z, and $-z$ axes with their centers of mass at equal distance from the origin. In each snapshot, there are therefore six possible configurations between the central group and the surrounding groups, i.e., there are six pairs simultaneously but all in different configurations. In this approach, the sampling of unfavored and rare configurations of the pair is enhanced and the system is more suitable to capture the many-body nature of nonbonded interactions between atomic groups in an atomistic model. For the benzene-benzene pair, the PMF calculated with this six-pair geometry is included in Fig. 2.6, which is obviously more repulsive at short distances as compared to the PMF calculated with only one pair. The location of minimum of the potential from the six-pair setup has also shifted to a larger value. As a result, the density of benzene from the coarse-grained model with nonbonded PMFs calculated with the six-pair geometry is reduced to 0.914 g/cm^3, much closer to the result of the atomistic model.

If more simultaneous pairs are used in a PMF calculation, we expect the resulting coarse-grained nonbonded potentials to be even more repulsive at short separations and the corresponding polymer density from the coarse-grained model to be reduced further. We have tested a setup in which twelve groups of atoms, all replicates of the same group, are placed around one central group. The placement is similar to the arrangement of twelve nearest neighbors around one atom in a face-centered-cubic crystal. Indeed, the resulting PMF is more repulsive at small distances. For benzene, the density from the corresponding PMF is reduced to 0.692 g/cm^3, which is smaller than the density from the atomistic model. Later on, we will show that the twelve-pair setup makes it harder to introduce a uniform entropic-

correction term to the coarse-grained force field of polyetherimides. Therefore, we settle with the six-pair setup for PMF calculations. It should also be pointed out that with atomic groups that have shapes resembling spheres, the results from the one-pair and six-pair setups are very close. In this regard, benzene molecules, which are almost planar, are an ideal model system illustrating the difficulty of developing coarse-grained models for molecular and polymeric systems.

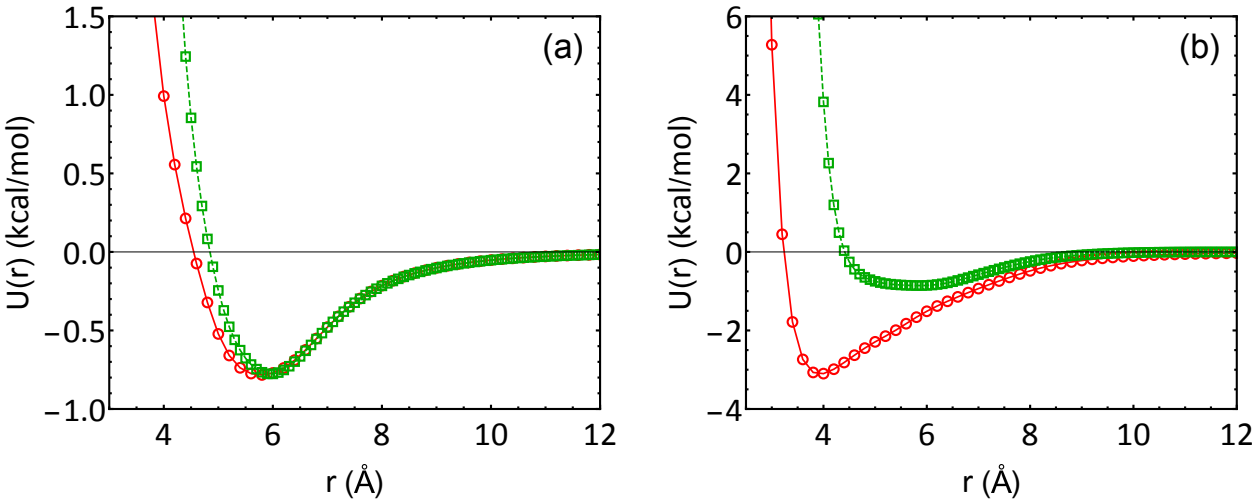

Figure 2.7: The van der Waals PMF, $U(r)$, as a function of separation r for (a) D-D pair and (b) A-D pair. The types of coarse-grained beads are defined in Fig. 2.4. The results are from all-atom MD simulations with a single pair (◯) and a system with one group at center and six surrounding groups (□). The lines are guides to the eye.

For the 5 types of coarse-grained beads defined in Fig. 2.4 for polyetherimides, We have performed the PMF calculations for all the 15 pairs. Two examples, for the D-D and A-D pair, are shown in Fig. 2.7. As expected, the results for the D-D pair are quite similar to those for the benzene dimer since the D bead represents an aromatic ring. The PMF from the six-pair setup is slightly more repulsive than that from the one-pair setup at short separations. However, for the A-D pair, the correction introduced by the six-pair setup is quite significant. The location of the PMF minimum shifts from about 4 Å to about 6 Å, which indicates that the PMF from the six-pair set up is much more repulsive than that

from the one-pair set up when the two corresponding atomic groups approach each other. The PMFs for other pairs also show shifts comparable to those shown in Figs. 2.7(a) and (b).

2.3.3.2 Entropic Corrections of Coarse-Grained Force Field

The six-pair setup discussed above has improved the PMF calculations and caused the resulting coarse-grained force field to better capture the many-body nature of nonbonded interactions between atomic groups in an atomistic model. However, the coarse-grained force field is still too attractive, leading to a polymer density higher than that from the atomistic model. An entropic correction can be introduced to the nonbonded, van der Waals PMF,[53, 110]

$$\begin{aligned} U(r) &= \int_{r_m}^{r} \left[\langle f_c \rangle_s + \frac{2k_{\mathrm{B}}T}{s} \right] ds \\ &= \int_{r_m}^{r} \langle f_c \rangle_s ds + 2k_{\mathrm{B}}T \ln\left(r/r_m\right) \ , \end{aligned} \tag{2.31}$$

where the term $2k_{\mathrm{B}}T\ln\left(r/r_m\right)$ corrects for the entropic volume contribution. This contribution can be written as $2k_{\mathrm{B}}\ln r$ as the volume sampled by the two atomic groups, which have a fixed center-to-center separation, r, and can rotate around each other, scales as r^2.[53, 110] Correspondingly, an entropic correction term

$$\delta\langle f_c \rangle_r = k_{\mathrm{B}}T\frac{2}{r} \tag{2.32}$$

should be added to the force between the two groups from the directly computed PMF. Note that the factor 2 in the above expression only reflects the fact that the sampled volume scales with r^2. In practice there is a prefactor in the expression of the sampled volume and

therefore the entropic correction term in the force can be more generally written as

$$\delta\langle f_c\rangle_r = k_{\mathrm{B}}T\frac{\alpha}{r}\ , \tag{2.33}$$

where α is treated as a fitting parameter that can be tuned to render the coarse-grained force field to better match the atomistic one. In this chapter, a coarse-grained force field including an entropic correction term as shown in Eq. (2.33) is termed a CG_α model.

To find the optimized value of α, we built an atomistic system consisting of 64 branched polyetherimide chains using MAPs.[111] The system was first heated to 800K and relaxed at that temperature for 5 ns. Then the system was cooled down to 300K within 5 ns. During these steps, the pressure of the system was kept to be one atmosphere. The configuration of the system during the cooling process at a given temperature (between 300 K and 800 K) was taken as a starting state for computing the density of branched polyetherimides. At each temperature, the system was first relaxed for 3 ns and its density was calculated in the subsequent 2 ns using an NPT ensemble. For the coarse-grained system, the same protocol was followed but the system consisted of 1000 branched polyetherimide chains and the time step of the MD simulations was set to 2 fs.

We find that a single value of α is sufficient to cause the density of polyetherimides from the CG_α model to match that from all-atom MD simulations at a given T. For example, $\alpha = 1.97$ for $T = 300$ K, as shown in Fig. 2.8. However, when the temperature is raised, the CG_α model predicts a more compressible polymer system as compared to the atomistic model. For example, at $T = 600$ K, the CG_α model gives a density 1.031 g/cm^3 for polyetherimides, while the density from the atomistic model at this temperature is higher at 1.150 g/cm^3.

Later we will discuss the results on the mechanical properties of polyetherimides from the CG_α model, which are significantly different from those from the all-atom model as well

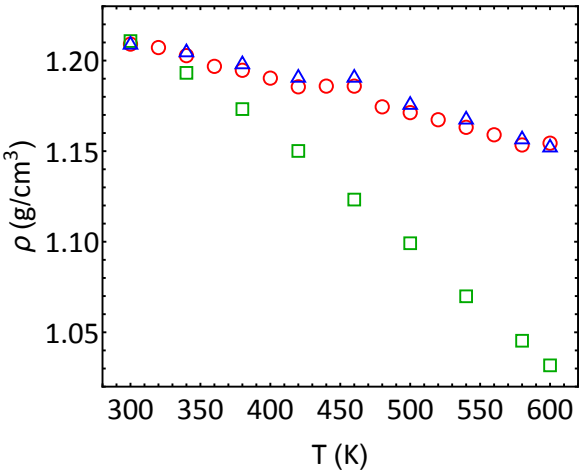

Figure 2.8: Density (ρ) of polyetherimides as a function of temperature (T) from all-atom MD simulations (◯), the CG_α model (□), and the CG_α^n model (△).

as experimental values. We find out that to improve the coarse-grained model, we need to modify the entropic correction term in the force as

$$\delta\langle f_c\rangle_r = k_B T \frac{\alpha}{r}\left(\frac{r_0}{r}\right)^n \ , \tag{2.34}$$

where n is an additional fitting parameter and r_0 is the separation between a pair of atomic groups at which their mutual force from the PMF calculation reaches minimum (i.e., where the force is most attractive). The rationale behind this modification is that the entropic correction should be more significant at small separations between a pair of atomic groups, where all-atom MD simulations are limited in terms of efficiently and sufficiently sampling all possible configurations of the pair. The form of the entropic correction term in Eq. (2.34) is negligible if $r \gg r_0$ and its importance grows in a power law for $r < r_0$. The coarse-grained force field with Eq. (2.33) as the entropic correction term is designated as the CG_α^n model.

Since the CG_α^n model has two fitting parameters, α and n, we can tune the model to have a density match with the atomistic model at two temperatures. As shown in Fig. 2.8, with

$\alpha = 0.395$ and $n = 15$, the density of polyetherimides from the CG_α^n model matches that from the all-atom model at both $T = 300$ K and 600 K. Furthermore, without further tuning the densities at other intermediate temperatures also match between the CG_α^n model and the all-atom model. That is, the CG_α^n model captures the thermal expansion property of polyetherimides from the atomistic model. This turns out to be a key point of making the coarse-grained model to better capture the mechanical properties of polyetherimides, as discussed below.

It should be clarified that the six-pair setup for PMF calculations and the entropic correction in either Eq. (2.33) or Eq. (2.34) are both needed as complementary steps to improve a coarse-grained model. If the six-pair setup was not used, then we would need to have a separate α or (α, n) combination for each pair of atomic groups (i.e., for each PMF) in order to achieve a density match (at one temperature for the CG_α model or a range of temperatures for the CG_α^n model). The number of fitting parameters would be too many for an optimization process to quickly converge. With the six-pair setup, just one α [or one (α, n) combination] is needed for all PMFs in the CG_α model (or the CG_α^n model)

2.4 Applications of the Coarse-Grained Model of Poly-etherimides

In this section we apply the coarse-grained model developed in this chapter to study the mechanical, structural, and rheological properties of branched polyetherimides. We show that the CG_α^n model reasonably captures these properties of polyetherimides.

2.4.1 Mechanical Moduli

We compute the mechanical moduli of branched polyetherimides with both atomistic and coarse-grained models. The setup of such simulations is shown in Fig. 2.9. A system of polyetherimide chains (64 atomistic chains or 1000 coarse-grained chains) is first equilibrated in a NPT ensemble at 300 K and 1 atmosphere. Periodic boundary conditions are used in all directions. After equilibration, a NVT ensemble is used and the stress tensor of the equilibrated system is computed as a reference. The simulation box size is 62.2 Å × 62.2 Å × 62.2 Å for the atomistic system and 155.7 Å × 155.7 Å × 155.7 Å for the coarse-grained system. Then either a tensile or shear strain is applied to deform the simulation box, as shown in Fig. 2.9. After deformed, the system is relaxed to remove transient effects and the stress tensor under the given strain is computed. The change of the stress tensor is analyzed as a function of the applied strain, which yields the mechanical moduli as well as Poisson's ratio of the materials.

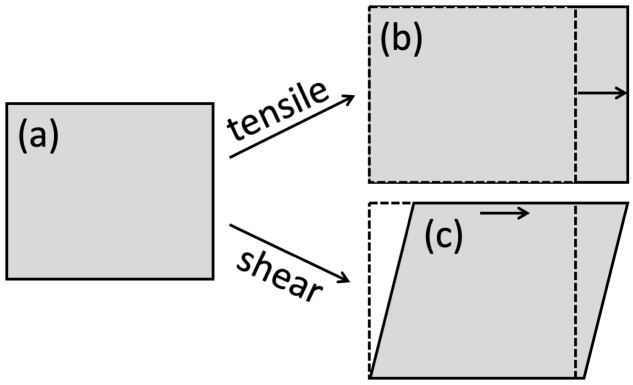

Figure 2.9: Setup of simulations used to compute the mechanical moduli of polyetherimides. The polymer domain is elongated in one direction in a tensile deformation [(a)→(b)] and sheared [(a)→(c)].

For an isotropic material, its mechanical moduli are determined by only two independent parameters, λ and μ, where λ is Lamé's first parameter and μ is Lamé's second parameter

or the shear modulus of the material. The mechanical moduli are then given by

$$\begin{aligned} K &= \lambda + \frac{2}{3}\mu \\ G &= \mu \\ E &= \mu\left(\frac{3\lambda + 2\mu}{\lambda + \mu}\right) \\ \nu &= \frac{\lambda}{2(\lambda + \mu)} \ , \end{aligned} \tag{2.35}$$

where K is the bulk modulus, G is the shear modulus, E is the Young modulus, and ν is Poisson's ratio.

Lamé's parameters can be determined by computing the stiffness tensor,

$$C = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix} . \tag{2.36}$$

The matrix element C_{ij} can be computed from $C_{ij} = \sigma_i/\epsilon_j$, where ϵ_j is a component of the strain tensor and σ_i is a component of the corresponding stress tensor with both tensors expressed in a vector form. The subindices, i and j, run from 1 to 6, denote the xx, yy, zz, yz, xz, and xy components, respectively, of the strain or stress tensor expressed in a 3×3 matrix form in a three-dimensional Cartesian system. In particular, $\sigma_1 = \sigma_{xx}$, $\sigma_2 = \sigma_{yy}$, $\sigma_3 = \sigma_{zz}$, $\sigma_4 = \sigma_{yz}$, $\sigma_5 = \sigma_{xz}$, and $\sigma_6 = \sigma_{xy}$ for the stress tensor. For the strain tensor, $\epsilon_1 = \epsilon_{xx}$, $\epsilon_2 = \epsilon_{yy}$, $\epsilon_3 = \epsilon_{zz}$, $\epsilon_4 = 2\epsilon_{yz}$, $\epsilon_5 = 2\epsilon_{xz}$, and $\epsilon_6 = 2\epsilon_{xy}$. Therefore, ϵ_1 is a tensile strain along the x-axis while ϵ_4 is twice the shear strain applied along the y-axis on a surface

perpendicular to the z-axis.

We first use tensile deformations to compute the top-left block of the stiffness tensor. Then Lamé's parameters are computed via

$$\lambda = \frac{1}{6}(C_{12} + C_{13} + C_{21} + C_{23} + C_{31} + C_{32}) \tag{2.37}$$

$$\mu = \frac{1}{6}(C_{11} + C_{22} + C_{33} - 3\lambda). \tag{2.38}$$

Shear deformations are also simulated to determine μ and the results are consistent with those from tensile deformations.

Table 2.4: Mechanical moduli and Poisson's ratio of polyetherimides at 300 K. The unit of K, G, E, and λ is GPa. The row of ν and the column for the parameter w are dimensionless.

	Atomistic	CG_α	CG_α^n	Experimental	w
K	4.185	0.727	2.110	4.297–4.942	0.50
G	0.770	0.120	0.469	1.059-1.070	0.61
E	2.178	0.342	1.311	2.965	0.60
λ	3.671	0.647	1.797	3.584-4.236	0.49
ν	0.413	0.422	0.397	0.385-0.400	

Table 2.5: Mechanical moduli and Poisson's ratio of polyetherimides at 400 K. The unit of K, G, E, and λ is GPa. The row of ν and the column for the parameter w are dimensionless.

	Atomistic	CG_α	CG_α^n	w
K	3.754	0.345	1.724	0.46
G	0.653	0.0167	0.310	0.47
E	1.852	0.0491	0.879	0.47
λ	3.319	0.334	1.517	0.46
ν	0.418	0.476	0.415	

The results on the mechanical moduli of branched polyetherimides are computed with the atomistic, CG_α, and CG_α^n models and are summarized in Table 2.4 for $T = 300$ K and in Table 2.5 for $T = 400$ K. The experimental values at $T = 300$ K are also included in Table 2.4. All models yield very good results on Poisson's ratio that match with the experimental value.

Regarding the mechanical moduli, the data further show that the results from the atomistic model are close to the experimental ones and the CG_α^n model is significantly improved from the CG_α model in terms of matching the atomistic model. It is also interesting to notice that if we define w as the ratio between the value of a mechanical modulus from the CG_α^n model and that from the atomistic model, then its value is around 0.5 for all mechanical moduli at either $T = 300$ K or 400 K. This comparison indicates that the CG_α^n model developed in this chapter is able to capture the mechanical properties of polyetherimides with an almost constant scaling factor about 0.5 and is transferable temperature-wise.

The much improved performance, including temperature transferability, of the CG_α^n model is due to the fact that it captures the thermal expansion property of polyetherimides when compared to the atomistic model. This behavior can be explained with the Grüneisen law that uses a parameter, γ, to describe the effect of a changing temperature on the size of dynamics of a crystal lattice. One expression of γ is

$$\gamma = \frac{\alpha_V K}{C_V \rho} \ , \tag{2.39}$$

where α_V is the volume thermal expansion coefficient and C_V is the constant-volume heat capacity of the crystal. The physical implication of the Grüneisen law is that the thermal expansion behavior of a crystal is intrinsically connected to its mechanical properties. If we assume the Grüneisen law also applies to polyetherimides, then the law indicates that the ratio $\gamma C_V/K$ should be the same for the atomistic and the CG_α^n model as they yield matching polymer densities as well as volume thermal expansion coefficients. The fact that the value of K from the CG_α^n model is about 50% of that from the atomistic model thus indicates that γC_V should scale similarly between the two models.

2.4.2 Pair Correlation Functions

We compute the pair correlation functions, $g(r)$, of all 15 pairs of atomic groups with all-atom MD simulations and the corresponding functions of all pairs of coarse-grained beads with the CG_{α}^{n} model. The comparison is shown in Fig. 2.10. The locations of peaks in $g(r)$ generally match well but their heights differ significantly. The results of $g(r)$ from the atomistic model indicates that the arrangement of atomic groups in polyetherimides are rather structureless, particularly beyond the first peak of $g(r)$. On the other hand, the coarse-grained beads show more local ordering as evidenced by a strong first peak in $g(r)$ for almost all pairs. Beyond the first peak, the pair correlation functions of the coarse-grained beads match reasonably well with those of the corresponding atomic groups. The results in Fig. 2.10 are not surprising as $g(r)$ never enters the process when the coarse-grained model is constructed.

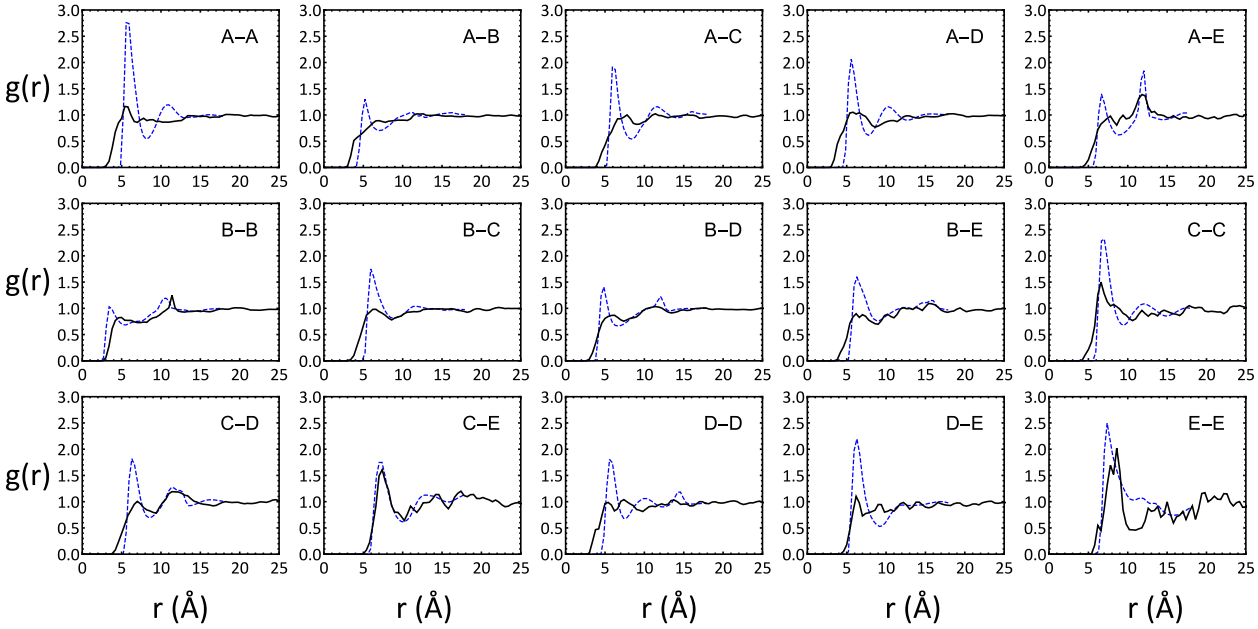

Figure 2.10: Comparison of pair correlation functions, $g(r)$, from the atomistic model (black solid line) and the CG_{α}^{n} model (blue dashed line) for all pairs of atomic groups or corresponding coarse-grained beads. Refer to Fig. 2.4 for the types of coarse-grained beads.

2.4.3 Rheology

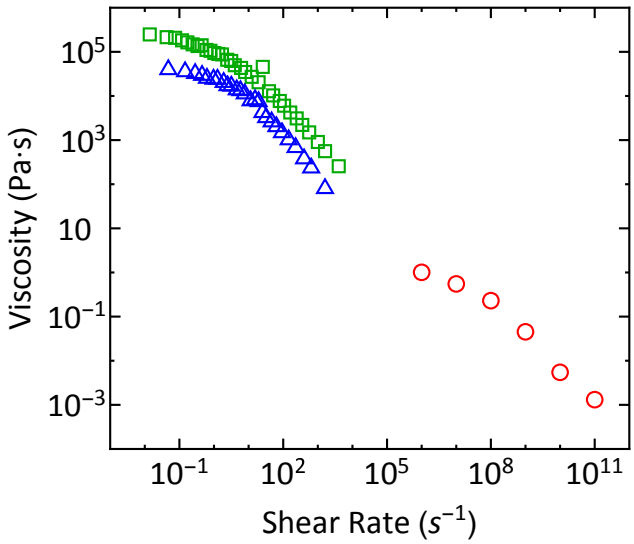

Figure 2.11: Shear viscosity as a function of shear rate from experiments (△ for branched polyetherimides with $M_n = 18.8$ kDa and □ for linear polyetherimides with $M_n = 20$ kDa) and MD simulations (◯ for branched polyetherimides with $M_n = 2.74$ kDa). For all systems, $T = 563$ K.

Finally, we applied the CG^n_α model discussed above to study the rheological properties of branched polyetherimides. The simple-shear setup shown in Fig. 2.9 was used to compute the viscosity of branched polyetherimides in a triclinic simulation box under a given shear rate. The system consisted of 8000 branched polyetherimide chains with $M_n = 2.74$ kDa initially in a cubic simulation box with side length 315.38 Å. The temperature was fixed at $T = 563$ K, at which branched polyetherimides formed a melt. This was also the temperature at which the experimental data on viscosity were obtained. At a fixed shear rate, the shear stress in the melt was computed in MD simulations with the CG^n_α force field and used to determine the viscosity. The MD data for shear rates ranging from 10^6 s^{-1} to 10^11 s^{-1} are shown in Fig. 2.11, together with experimental data for branched polyetherimides with $M_n = 18.8$ kDa and linear polyetherimides with $M_n = 20$ kDa. The shear thinning behavior is obvious from both MD and experimental results. Furthermore, the upper range of the

experimentally-probed shear rates is about 10^4 s^{-1} to 10^5 s^{-1}. The CG^n_α model enables us to approach the experimental range of shear rates and as well as the range of low shear rates where the Newtonian plateau occurs.

2.5 Conclusions

In this chapter, we have developed a coarse-grained model of polyetherimides on the basis of chemistry-informed grouping of atoms, parameterization of bond and angle interactions by fitting the distributions of bond lengths and angles to Gaussian functions, and parameterization of nonbonded interactions via pontential of mean force calculations. Our results show that a six-pair setup can be used to improve configuration sampling in a potential of mean force calculation and an entropic correction term can be introduced to make the coarse-grained model to capture the thermal expansion property of the polymer. As a result, the coarse-grained model is transferable temperature-wise and can capture the mechanical moduli of polyetherimides within a (temperature-independent) constant scaling factor, which is around 0.5 here. The coarse-grained model further enables us to approach the range of shear rates accessible to rheology experiments and probe rheological behavior such as shear thinning. The coarse-grained model only fairly captures the structural property of the polymer and future improvements are still needed in this respect.

Chapter 3

Determination of the Glass Transition Temperature of Polyimides from All-Atom Simulations and Machine-Learning Algorithms

This chapter is based on a paper to be submitted:

Chengyuan Wen, Binghan Liu, Roy Odle, and Shengfeng Cheng, "Determination of Glass Transition Temperature of Polyimides from All-Atom Molecular Dynamics Simulations and Machine Learning Algorithms".

I designed and built all the models for molecular dynamics simulations. I performed all the data analyses and prepared figures. Assisted by Mr. Liu, I developed the predictive model of glass transition temperature of polyimides using machine-learning algorithms. Dr. Odle from SABIC contributed some ideas and provided experimental insights. All authors

contributed to the writing of this paper. My contributions to this paper were under Dr. Cheng's supervision.

3.1 Introduction

When a polymer is rapidly cooled below a certain temperature, it undergoes a transition into a glassy phase where the polymer has an amorphous structure but exhibits rigidity on experimental time scales. The temperature at which this transition occurs is termed the glass transition temperature (T_g) and is one of the most important properties of a polymer that determine its performance and applications. For example, if a polymer has to stay as a hard solid in a certain application, its T_g should be much higher than the environmental temperature, T_e. On the other hand, if a rubber or a polymer melt is desired, then T_g needs to be lower than T_e. The difference of T_g and T_e also strongly affects other physical properties of the polymer, such as its density and the diffusion behavior of guest gas molecules in the polymer. In other words, many physical properties of a polymer exhibit changes, which can be significant, when T_e is varied to cross T_g. This observation is underlying a variety of methods that are used to determine T_g via measuring or computing these physical properties as a function of temperature. The glass transition temperature is therefore a critical parameter to be considered when the target is to design or identify a polymeric material that meets the requirement of a given application. In this chapter, we explore three approaches to determine T_g for various polyimides *in silico*, including computing density and gas diffusion coefficients with atomistic molecular dynamics (MD) simulations and predicting T_g with a model derived with machine-learning algorithms applied to the existing structure-property data on T_g of polyimides collected from the literature.

Polyimides are a category of engineering plastics that have wide applications in the auto-

motive and aerospace industries because of their relatively high T_g, high strength, and good heat resistance properties.[115, 116] Experimentally, T_g can be measured with differential scanning calorimetry and thermo-mechanical analysis techniques. However, these procedures usually require careful sample preparation and control of the measurement conditions. As a supplementary approach to expedite material characterization, MD simulations have been used to quantify T_g since the 1980s. Rigby *et al.* calculated T_g for Kremer-Grest chains consisting of Lennard-Jones beads as a model of polyethylene.[32] The temperature dependence of the density, self-diffusion coefficient, and the characteristic ratio $\langle r^2 \rangle / (nl^2)$ was used to identify the glass transition and T_g. Han *et al.* calculated T_g for five different polymers using the curve of specific volume against temperature.[56] Abu-Sharkh *et al.* used a similar method to compute T_g for poly(vinylchloride)s with the force field determined with an *ab initio* method.[57] In these studies, usually only one polymer chain was simulated for each system because of the limitation of computational power. Morita *et al.* simulated 100 coarse-grained polymer chains and introduced a method to compute T_g by examining the mean-square displacement of a polymer segment at different temperatures.[58] Buchholz *et al.* studied the cooling-rate-dependence of T_g with the Kremer-Grest model, where T_g was found from the curves of nonbonded energy or system volume versus temperature.[59] Following these initial efforts, other researchers started to compute T_g with atomistic MD simulations for various polymers,[60–65] typically by investigating the density change of a polymer when the temperature is varied. Lyulin *et al.* calculated the T_g of several polyimide polymers and pointed out that the results from MD simulations depend on cooling rate and vary if the atomistic model of a polymer considers partial charges or not.[68] Mohammadi *et al.* computed T_g of poly(methyl methacrylate) using the first peak of the pair correlation function, the mean square displacement of polymer segments, the self-diffusion coefficient, and the internal energy of the system. [117] They found that though all these different physical quantities exhibit an obvious transition at T_g, the values from MD simulations are

usually lower than the experimental value. All the reported work thus shows that MD simulations can be a useful tool to obtain T_g but the results can suffer from small system sizes, short chain lengths, and high cooling rates, all reflecting the limitations of MD methods. Furthermore, the results may depend on the particular force field being used in a study. [67, 118, 119]

Although MD methods can be used to calculate T_g for a polymer or even predict it for a polymer that is not yet synthesized, the calculations can still take a long time and may be limited by available computational resources. Therefore, a predictive model of T_g is highly desirable, which uses certain features of a polymer such as the chemical identity of the monomer, sequence structure of the chain, and others as inputs. Such a model can be applied to quickly yield T_g that can be tested later with MD calculations or experiments. This capability will allow quick screening of a series of polymers when a particular application is in consideration. Efforts of generating so-called quantitative structure property relationships (QSPRs) for the glass transition temperature of a polymer have been ongoing since the 1990s.[69, 70] Joyce *et al.* used neural network algorithms to train a model for T_g prediction with data on 360 monomers and the model can predict T_g for other 89 monomers with a root mean square error (RMSE) of about 35 K. [70] The large error may be caused by the fact that the 360 monomers picked by Joyce *et al.* were for a broad range of polymers and with a small dataset, neural network algorithms could easily lead to overfitting problems. Yu *et al.* applied the multiple linear stepwise regression method to establish the predictive model of T_g with a RMSE around 15.2 K.[71]

Chen *et al.* [72], Ning *et al.* [73], and Xu *et al.*[74] also developed predictive models of T_g with different accuracy for a variety of polymers. The number of data points in their training set ranges from 52 to 107. Pei *et al.* applied a support vector regression (SVR) optimized by an integrated particle swarm optimization to predict T_g.[75] They used 25 sample points

to train the model and 7 other data points to test the model. The RMSE of their model is around 4 K. However, the penalty factor C of 56700186162.908470 used in their model training procedure is not replicable. Chen *et al.* applied multiple linear regression analysis to establish a predictive model of T_g with 60 training data points.[76] The test set contained 20 data points and the error of T_g was found to be around 58 K.

Despite the existing efforts of constructing predictive models of T_g for a range of polymers, the stability of such models has not been proved or discussed. It is unclear if the models reported in the literature are robust and possess the same predictive power and accuracy if the training and test datasets are split in different ways. In this chapter, we discuss the instability issue of the commonly used regularization method termed least absolute shrinkage and selection operator (LASSO) and find that a bragging approach can be used to reduce the instability of the predictive model of T_g derived with LASSO. Furthermore, we compare T_g predicted by the model from machine learning algorithms to those computed with atomistic MD simulations for several polyimides that are yet synthesized. This comparison not only serves as a test of the predictive model but also validates the power of MD simulations of yielding T_g for polyimides with new formulae.

The chapter is organized as follows. In Sec. 3.2, the methods of determining T_g with atomistic molecular dynamics simulations are introduced and the results are analyzed and discussed. Then in Sec. 3.3, the procedure of building a predictive model of T_g for polyimides using machine learning algorithms is discussed in detail, including dataset preparation, the digitization of polymer structures, the conversion of polymer structures to proper SMILES notations, and the construction of the predictive model via machine learning. Finally, conclusions of this chapter are presented in Sec. 3.4.

3.2 Determining Glass Transition Temperature with All-Atom Molecular Dynamics Simulations

3.2.1 All-Atom Molecular Dynamics Simulation Methods

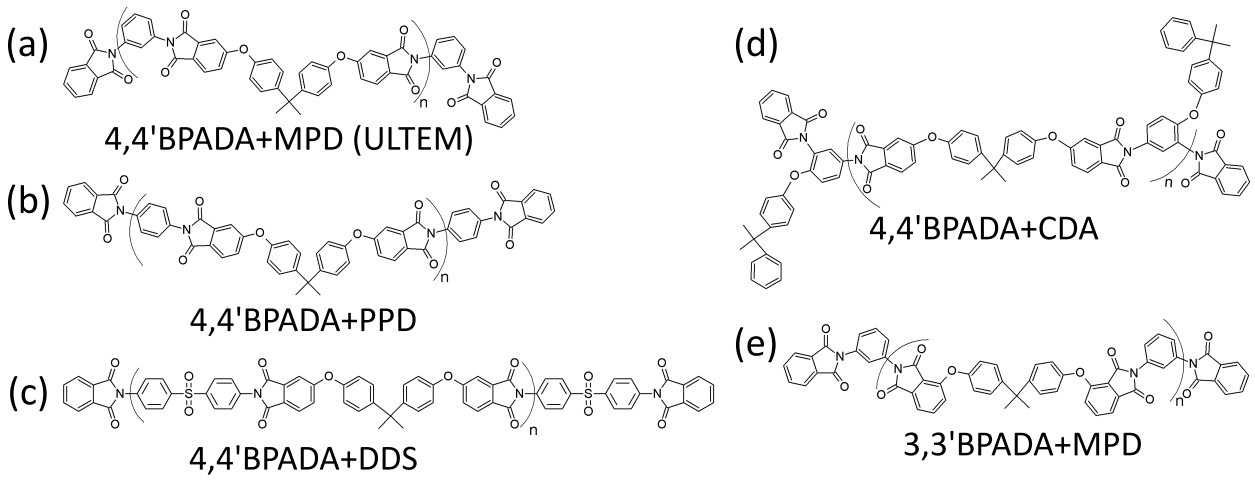

Figure 3.1: Structures of polyetherimides studied here.

Atomistic MD simulations were employed previously to model the mechanical, thermal, and dielectric properties of polyetherimides.[67, 120, 121] The polyimide chains in our study were built with MAPS builder.[111] In this section five polyetherimides are studied, including 4,4'BPADA+MPD (ULTEM), 4,4'BPADA+PPD, 3,3'BPADA+MPD, 4,4'BPADA+CDA, and 4,4'BPADA+DDS.[1] The chemical formulae of these structures are shown in Fig. 3.1. Each chain consists of four repeating units and phthalic anhydride (PA) groups are added to cap the chain. All MD simulations were performed using LAMMPS [122] with the PCFF force field.[123] The equations of motion were integrated with a velocity-Verlet algorithm with time step $\Delta t = 1$ fs. The Mulliken charge was included in the model and calculated

[1]BPADA: 4,4'-bisphenol A dianhydride; MPD: m-phenylenediamine; PPD: para-phenylenediamine; CDA 1,2-dihydroxybenzene dianhydride: DDS: diphenyl sulfone.

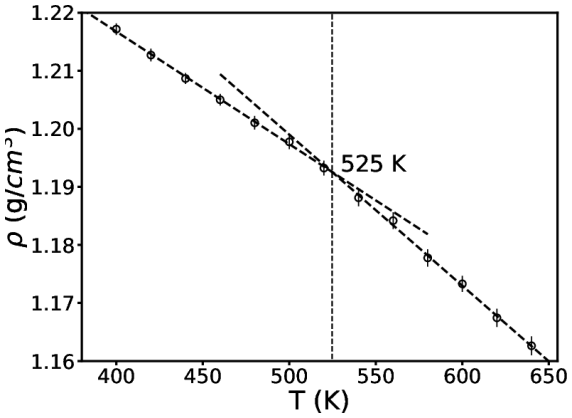

Figure 3.2: $\rho(T)$ vs. T for 4,4'BPADA+MPD (ULTEM), for which $T_g = 525$K.

using Gaussian09 software with the semi-empirical PM6 method as the basis set.[124] The cutoffs of nonbonded and Coulomb interactions are both set as 12 Å and the long-range part of Coulomb interactions was calculated using the particle-particle particle-mesh (PPPM) method. Each system contained 512 chains. A hydrostatic pressure of 1000 atm was used to compress the system at 300 K until it reached a density around 1.2 g/cm^3, close to the experimental value of ULTEM.[125] Then the system was heated up to 800 K under 1 atm and equilibrated at 800 K for 5 ns. After this step, the system was gradually cooled down to 300 K under 1 atm. In this process, many configurations were created for a series of temperatures between 300 K and 800 K. At a given temperature, the corresponding configuration was taken as a starting state and the system was equilibrated further for 2 ns. The density of the polymer was then computed in a NPT ensemble. The equilibrated system was also used for simulating the diffusion of gas molecules in the polymer. In these simulations, a NVT ensemble was adopted with the temperature controlled by a Nose-Hoover thermostat.

A commonly used protocol to determine T_g for a polymer is to calculate its density as a function of temperature, i.e., to obtain the $\rho(T)$ curve.[126–131] One example is shown in Fig. 3.2 for ULTEM (i.e., 4,4'BPADA+MPD). The value of T_g can be determined from the intersection of the two linear fits to the data on $\rho(T)$, one for the lower and the other for

the higher temperature region. At room temperature, the density of ULTEM from MD simulations is slightly lower than the experimental value, 1.27 g/cm^3. The data indicate that the variation of density with temperature is captured by the PCFF force field. The value of T_g computed from MD data on $\rho(T)$ is 498K for ULTEM, which is very close to the experimental value, 490K. The $\rho(T)$ curve is also used to determine T_g for polyetherimides with other chemical formulae and all results are listed in Table 3.1, where a good agreement is found between MD calculations and available experimental data.

In addition to density, there are other properties of a polymer that can be used to determine T_g, including volume, free volume, specific volume, radial distribution functions, mean-square displacements, nonbonded energy, dihedral torsion energy, etc.[126, 127] Many studies also showed that the diffusion behavior of gas molecules in a polymer matrix changes when the polymer undergoes a glass transition.[132, 133] This can be understood by examining the temperature dependence of the diffusion coefficient of a gas molecule in a polymer matrix, which has an Arrhenius form,

$$D = D_0 \exp\left(-\frac{E_A}{RT}\right) \ , \tag{3.1}$$

where D is the diffusion coefficient, D_0 is a constant with the same unit as D, E_A is the activation energy for diffusion, T is the absolute temperature and $R = 8.314$ J mol^{-1} is the gas constant. Note that E_A may be temperature dependent but in many cases the dependence is weak and negligible.[134] For a glassy polymer, the value of E_A changes when the polymer enters a glassy state from a melt state. Therefore, a plot of $D(T)$ vs. $1/T$ on a log-linear scale will show a straight line for $T > T_g$ and another straight line with a different slope for $T < T_g$. The intersections between these two lines can be used to determine T_g.

In MD simulations, the diffusion coefficient D of a gas molecule in a polymer can be computed from its mean-square displacement, $\langle r^2 \rangle$. In the diffusive regime, its time dependence can

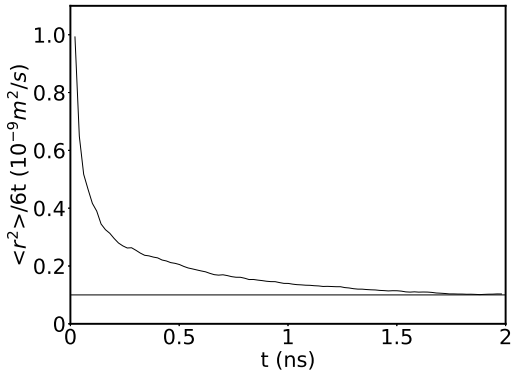

Figure 3.3: $\langle r^2 \rangle/(6t)$ vs t for argon diffusing in 4,4'BPADA+MPD at 400K, which yields $D = 1.0 \times 10^{-10}\mathrm{m}^2/\mathrm{s}$.

be expressed as

$$\langle r^2 \rangle = 6Dt + C, \tag{3.2}$$

where t is time and C is a constant. One example is shown in Fig. 3.3, where we plot $\langle r^2 \rangle/(6t)$ vs t. Clearly, $\lim_{t\to\infty} \langle r^2 \rangle/(6t) = D$. This calculation can be performed at various temperatures to generate the $D(T)$ curve. One example is included in Fig. 3.4 for 4,4'BPADA+MPD by plotting $\log D$ against $1/T$. The two regions in which $\log D$ depends $1/T$ linearly are clearly visible. The corresponding linear fits and their intersection are used to determine T_g of this polyetherimide to be about 504 K.

3.2.2 Molecular Dynamics Simulation Results and Discussion

Figure 3.4 shows $D(T)$ vs. $1/T$ for 4,4'BPADA+MPD, from which T_g can be determined as 504K. The value of T_g from $\rho(T)$ vs. T is around 525K. The two results are in a good agreement. Similar agreements are also found for polyetherimides with other chemical formulae. All results on T_g are summarized in Table. 3.1. It is noted that the results from MD simulations match well with available experimental values.

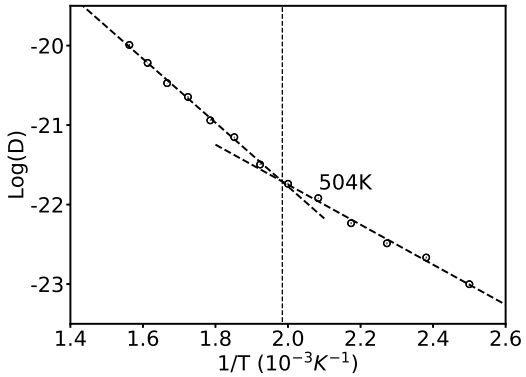

Figure 3.4: D vs. $1/T$ for helium at a concentration of 0.71 atoms/nm^3 diffusing in 4,4'BPADA+MPD, which yields $T_g = 504$ K.

Table 3.1: Summary of T_g for various polyetherimides from experiments and MD calculations of $\rho(T)$ and $D(T)$ respectively.

	4,4'BPADA + MPD	4,4'BPADA + PPD	4,4'BPADA + DDS	4,4'BPADA + CDA	3,3'BPADA + MPD
Experiment	490K	504	520 K	473 K	510
MD [$\rho(T)$]	525 K	539 K	542 K	516 K	524 K
MD [$D(T)$]	504 K	488 K	532 K	495 K	500 K

The method of determining T_g from $D(T)$ vs. $1/T$ has several advantages. First of all, the diffusion coefficient of a gas molecule can be computed quickly and accurately with MD simulations. The self-diffusion coefficient of a polymer can also be used to pinpoint T_g. However, a polymer typically diffuses much more slowly than gas molecules. As a result, much longer MD simulations are needed to compute the self-diffusion coefficient of a polymer to the same precision as in the diffusion coefficient of gas molecules. Secondly, to compute $D(T)$ we can use a NVT ensemble with temperature well controlled by a suitable thermostat (e.g., Nose-Hoover thermostat). However, to compute the $\rho(T)$ curve, a NPT ensemble is required. In MD simulations, it is practically very challenging to control pressure accurately, particularly if the pressure is as small as 1 atm.[135, 136] Our MD data show that when the target pressure is 1 atm, the actual pressure in the system can fluctuate significantly

from about -190 atm to about 190 atm. As a result, the density also fluctuates strongly. Computing $D(T)$ instead of $\rho(T)$ circumvents this issue and leads a much faster convergence of the data that can be used to determine T_g.

3.3 Predictive Model of Glass Transition Temperature Trained with Machine-Learning Algorithms

3.3.1 Machine-Learning Methods

Machine learning is considered a subset of artificial intelligence. A machine-learning algorithm is a mathematical model that can be trained by a set of sample data without requiring the system to be explicitly programmed to generate (pre-determined) outputs on the basis of various inputs. After the training process, such a mathematical model can be used to make a future decision or prediction given new data. There are three basic machine learning paradigms: supervised, unsupervised, and reinforcement. The process adopted here is a supervised learning method as the sample dataset used for training includes both inputs (e.g., polymer chemical identity and sequence) and desired outputs (e.g., glass transition temperature). The outcome of the learning process is an optimized objective function that connects the chemical information of a polymer, particularly its monomer type and sequence, to its measurable physical property.

Many efforts have been devoted to synthesize various polyimides, characterize their structures, and measure their properties including T_g.[137–141] By collecting the available data published in the literature and applying machine learning approaches to analyze the data, we can develop a predictive model for the glass transition temperature of polyimides. This

model can be used to probe polyimides that are yet to be synthesized. As a first step, we will compare the predictions of T_g from the mapping function derived via machine learning to those computed with atomistic MD simulations for a few selected polyimides not currently made in a lab. This comparison serves as a test of the machine-learning-generated predictive model. In the future, various polyimides with potential values in terms of their performance and application will be screened with the predictive model and then selected formulae will be synthesized in a lab to validate and improve the model.

3.3.1.1 Database and Feature Generation

We collected 225 data points on the glass transition temperature of polyimides from the literature, including 160 data points from Ref. [142] and 65 data points from Ref. [143]. Some sample data are shown in Table. 3.2 For each polymer, the chemical identity of the monomer is taken as the input. The skeleton notations of the polyimides were drawn and converted into an expression called Simplified Molecular-Input Line-Entry System (SMILES), which is a line-notation system using an ASCII string to represent the structure of a polymer. Then a feature-generating engine called E-dragon was utilized to read in the generated SMILES notations and to extract the available features for each polyimide.[144] In polymer informatics, features are also called descriptors, consisting of individual measurable properties of a molecule or a polymer.[145–147] The ensemble of descriptors represents the characteristics of the polymer/molecule being studied. As polyimides are made of dianhydrides and diamines, we calculated the features for a dianhydride and a diamine group separately. For each polyimide, E-dragon generates 1342 descriptors for its dianhydride group and the same set for its diamine group. Sample features include molecular weight, sum of atomic van der Waals volumes, and sum of atomic polarizabilities, etc. All polymer features are included and explained in Appendix C.

Table 3.2: Sample dataset of T_g of polyimides from experiments.

No.	Polyimide's Name	SMILES Notation	T_g (K)
1	4,4'TDPA+1,4,4APB	Nc1ccc(cc1)···(=O)OC(=O)c7c6	234
2	3,3'ODPA+M,M'DABP	Nc1cccc(c1)···(Oc4cccc3C(=O)OC(=O)c34)c5C6=O	234
3	4,4'ODPA+M,M'DABP	Nc1cccc(c1)···c5ccc6C(=O)OC(=O)c6c5	235
4	3,3'ODPA+M,M'DDS	Nc1cccc(c1)···(Oc4cccc3C(=O)OC(=O)c34)c5C6=O	241
5	3,4'TDPA+1,4,4APB	Nc1ccc(cc1)···(Sc4ccc5C(=O)OC(=O)c5c4)c6C7=O	242
6	3,4'ODPA+M,M'DABP	Nc1cccc(c1)···(Oc3ccc4C(=O)OC(=O)c4c3)c5C6=O	243
7	4,4'BPDA+M,M'ODA	Nc1cccc(c1)···c6C(=O)OC(=O)c6c5	243
8	4,4'BPDA+p,p'ODA	Nc1ccc(cc1)...OC(=O)c6c5	262
9	4,4'BTDA+m,m'MDA	Nc1cccc(c1)...c5ccc6C(=O)OC(=O)c6c5	272
10	4,4'BTDA+o,o'MDA	Nc1ccc(cc1)...c7cccc6C(=O)OC(=O)c67	283
...	...	...	...

3.3.1.2 Data Splitting into Training Set and Test Set

For polyimides, the values of T_g collected from the literature range from 273 K to 697 K. However, the distribution is not uniform in this range. The majority of the data is between 466 K and 583 K. The distribution of the 225 data points on T_g is shown in Fig. 3.5, with a peak around 530 K. The nonuniform nature of the distribution must be considered when the dataset is split into a training set and a test set as it is important for the training set to be representative of the entire dataset. This is particularly a concern if the number of available data is limited, as in the case here. To examine the influence of splitting the dataset on the performance of the resulting predictive model of T_g, we test two different ways of dividing the dataset into a training and a test set. To this end, we only use the 160 data points from Ref. [142] to train the model and reserve the 65 data points from Ref. [143] as a completely independent test of the predictive capability of the machine-learning-trained model. To ensure that the relatively small dataset can be split consistently, we first remove the data points of T_g at the tail of the probability distribution, i.e., those below 423 K or above 623K. The total number of the remove data points is 9, leaving 151 points in the dataset. In the first way, this dataset is randomly split into a training set containing 85% of

the data and a test set consisting of the remaining 15%. In the second way, the dataset is first divided into 8 adjoining sections, each of width of 25 K. In each section, 15% of the data points were randomly selected to join the test dataset. The remaining 85% of the data points form the training dataset. This strategy ensures that the statistical distributions of both the training and test datasets are similar to those of the entire dataset. We designate this second approach of dividing the dataset as "statistical splitting", while the first approach is termed "random splitting".

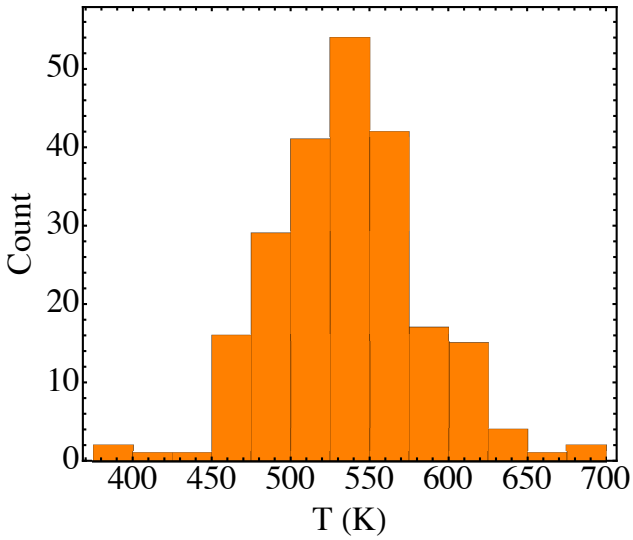

Figure 3.5: The distribution of T_g of polyimides collected from Refs. [142] and [143].

3.3.1.3 LASSO Regularization

For a given polyimide, there were 1342 features generated for the dianhydride group and the same number of features for the diamine group. Not all these features play important roles in affecting the glass transition temperature of a polyimide. Including irrelevant or partially relevant features can lead to overfitting behavior of the resulting predictive model and negatively impact its performance. Overfitting is a common problem faced by machine-learning methods and many techniques have been developed to address this problem. In our

approach, the importance of features were identified and ranked using LASSO regularization. At the end, a finite number of features were identified that control T_g of polyimides.

In a linear fitting, each estimated target value y_i could be represented as

$$y_i = \omega_0 + \sum_{j=1}^{p} x_{ij}\omega_j + \epsilon_i \ , \tag{3.3}$$

where ω_0 is a constant, ω_j is a fitting parameter representing the coefficient of the j-th feature in a linear mapping from features to target value, p is the number of features, and ϵ_i is the error of predicting the i-th data point. In a regular linear fitting scheme, the parameters ω_j can be found by minimizing the error function

$$\text{error} = \sum_{i=1}^{n}(y_i - \omega_0 - \sum_{j=1}^{p} x_{ij}\omega_j)^2 \ , \tag{3.4}$$

where n is the number of data points. In the LASSO regularization method, the error to be minimized is slightly modified as

$$\text{error} = \sum_{i=1}^{n}(y_i - \omega_0 - \sum_{j=1}^{p} x_{ij}\omega_j)^2 + \lambda \sum_{j}^{p} |\omega_j| \ , \tag{3.5}$$

where λ is a penalty factor. The advantage of the LASSO regularization method is that the coefficient of irrelevant and low-importance features can be shrunk to zero, which is an effective way of removing those features. If λ is 0, then there will no shrinkage of any of the 1342 features, and a LASSO regularization becomes a linear regression.[148] A big positive value of λ indicates that the majority of the features will be removed. In the LASSO regularization method, λ is called a hyper-parameter which cannot be learned directly. In our implementation, the value of λ was exhaustively searched from 0.01 to 2.0 in increments of 0.04. Our results reveal that the typical value of λ is between 0.3 and 1.6. For such λ, most

features are removed after a LASSO regularization. At the end, 197 features with nonzero coefficients remain in the final predictive model of T_g (see Appendix D) and the features with zero coefficients are removed during LASSO regularization. Out of 197 features, only about 12 features are actually important as indicated by their relatively large coefficients. The summation of the absolute value of coefficients of the largest 12 features is larger than the summation of the rest 185 features. Many of them can be easily justified on the basis of the available experimental evidence.

3.3.1.4 Bagging

Although the dataset on T_g of polyimides has more data points than those in many previous studies on other classes of polymers,[71–75], it is still a small set in the perspective of machine learning. The performance of the predictive model can exhibit significant fluctuations depending on how the dataset is split into a training and a test set. To reduce such variations, we utilized a bagging approach in the learning process.[149]

In the bagging approach, a dataset is randomly split into a training set and a test set. The machine learning procedure described above, including a LASSO regularization and an optimization process, is followed to generate a predictive model of T_g using the training set. The whole process is then repeated by splitting the dataset into a new training set and a new test set. After N_m repetitions, N_m models are generated. The performance of each model is quantified by its error defined as

$$\mathrm{Err}(k) = \sqrt{\frac{1}{n_k}\sum_{i=1}^{n_k}\left(T_g^k(i)_{\mathrm{predicted}} - T_g^k(i)_{\mathrm{experimental}}\right)^2}\,, \tag{3.6}$$

where k is the index of the model and n_k is the number of data points in the test set for the k-th model.

With the error associated with each model calculated, a weight, $\mathrm{W}(k)$, was assigned to the k-th model according to

$$\mathrm{W}(k) = \frac{(\mathrm{Err}(k))^{-1}}{\sum_{j=1}^{N_m} (\mathrm{Err}(j))^{-1}}. \tag{3.7}$$

The choice of the weight function in Eq. (3.7) guarantees that a model with a better performance, i.e., a smaller error in predicting the data in the corresponding test set, has a larger weight in the final predictive model. The final predictive model of T_g is the linear combination of N_m models weighted by $\mathrm{W}(k)$ as in

$$T_\mathrm{g} = \sum_{k=1}^{N_m} \mathrm{W}(k) \cdot T_\mathrm{g}^k \ . \tag{3.8}$$

In the context of machine learning, this bagging procedure is often used to render the learned model to be more robust and stable.

3.3.2 Model Training and Test

3.3.2.1 Various Ways of Training Predictive Model of Glass Transition Temperature

We implemented the machine learning approach and tested the resulting predictive model of T_g in four different ways. As discussed earlier, the dataset includes 151 data points from Ref. [142]. In the first and second way, this dataset was randomly split into a test set containing 15% of the data points. The remaining 85% of the data formed the training set. In the first way, the training set was used to train the predictive model of T_g via a LASSO regularization but bagging was not used. In the LASSO regularization, the fitting parameter was exhaustively searched using a grid search method. The performance of the predictive

model was quantified using the error of predicting the data points in the test set that never entered the training process. The entire procedure was repeated 1000 times and therefore 1000 models were generated. We analyzed the distribution of the errors of these models to predict the test set, which provided a metric of the stability and performance of the first way of training the predictive model of T_g.

In the second way, "random splitting" was still used as in the first way but the bagging approach was used to train the predictive model of T_g with each training set. In bagging, a training set was randomly split further into a training subset (85%) and a test subset (15%). A LASSO regularization was applied to the training subset to obtain a model. This model was used to predict the test subset and the error of prediction was used to decide the weight (i.e., performance) of the model. The random splitting of the training set into two subsets was repeated 40 times, i.e., $N_m = 40$. The linear combination of these models yielded a blended predictive model of T_g. This blended model was used to predict the test set that never entered the training process and the associated error of prediction was taken as the gauge of the model's performance. The entire procedure was repeated 1000 times to generate 1000 blended predictive models.

The third and fourth ways were similar to the first and second ones except that "statistical splitting" discussed in Sec. 3.3.1.2 was used instead of "random splitting". In the third way, bagging was not used while the fourth way was a combination of "statistical splitting" and bagging. In all these ways, the first level of splitting was repeated 1000 times, resulting in 1000 models. When bagging was used, the second-level splitting of the initial training set into a training and a test subset was always repeated 40 times and therefore, $N_m = 40$.

3.3.2.2 Performance of Predictive Model of Glass Transition Temperature

In this section, we show the performance of the predictive models of T_{g} generated using the training methods described previously. In the plots shown below, each dot represents one polyimide. For each dot, the x-coordinate indicates the target, which is the actual value of T_{g} determined experimentally. The y-coordinate indicates the predicted value of T_{g} from a machine-learning-based model. The blue line indicates $y = x$. The closer a dot to the blue line, the better the performance of the predictive model. In Figs. 3.6 to 3.9, orange dots represent the data used in training the predictive model while the green triangles represent the data used in testing the model, which do not enter the training process.

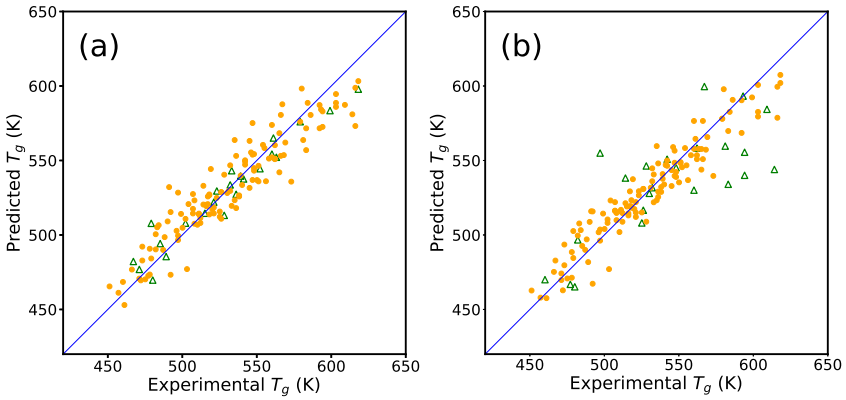

Figure 3.6: Performance of the (a) best and (b) worst model from training method #1 ("random splitting" + no bagging).

In our study, 1000 predictive models were generated using each training method. These models were ranked by their errors of predicting the test datasets that were not used to train the models. The performance of the best and worst predictive model of T_{g} derived in the first manner of implementing the machine learning approach described previously (i.e., "random splitting" + no bagging) is shown in Fig. 3.6. The error of using the best model to predict the training set is 14.37 K while the error is 10.78 K if the model is used to predict

the test dataset. The errors of the worst model are 12.03 K for the training set and 29.62 K for the test set, respectively.

The performance of the best and worst predictive model of T_g trained with the second method (i.e., "random splitting" + bagging) is shown in Fig. 3.7. The error of predicting the training set is 14.26 K for the best model and is 12.08 K for the worst model. The best model has an error of 10.65 K when it is used to predict the test dataset while the error is much larger at 29.61 K for the worst model.

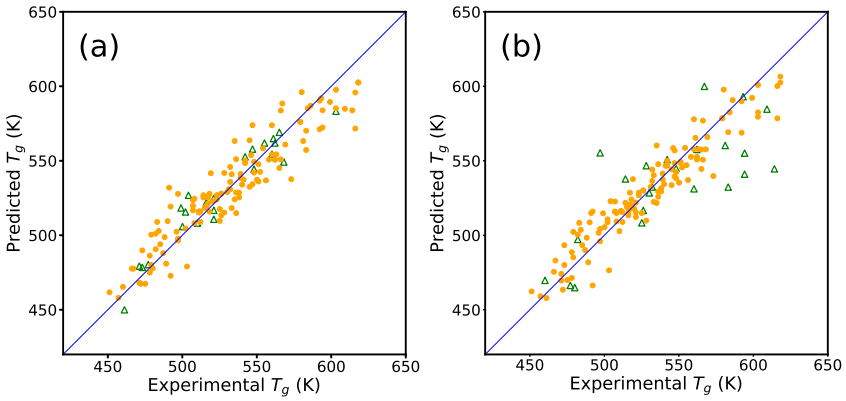

Figure 3.7: Performance of the (a) best and (b) worst model from training method #2 ("random splitting" + bagging)

Figure 3.8 shows the performance of the best and worst models trained using the third method (i.e., "statistical splitting" + no bagging). The best model has an error of 13.65 K of predicting the training set and of 9.79 K for the test dataset. The worst model has a smaller error at 9.66 K of predicting the training set while the error is much larger at 30.05 K when the test dataset was used to check the performance of the predictive model. This large discrepancy of errors of predicting the trainig and test dataset is a reflection of the overfitting issue faced by many machine-learning approaches. Below we show that using bagging can effectively address this issue.

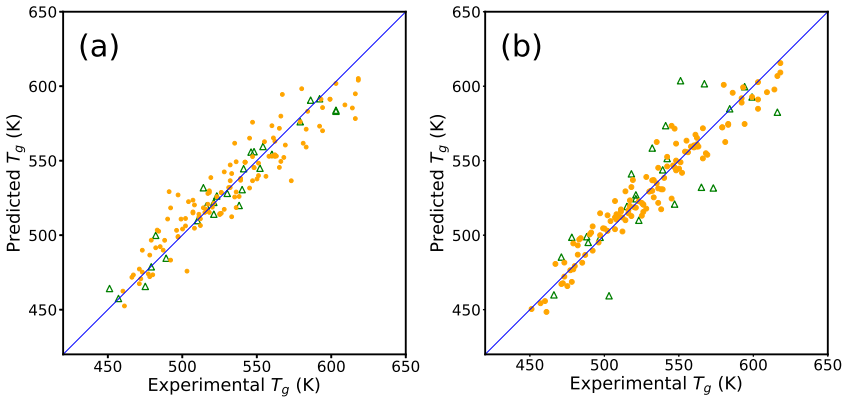

Figure 3.8: Performance of the (a) best and (b) worst model from training method #3 ("statistical splitting" + no bagging).

The bagging approach introduced in Sec. 3.3.1.4 can be used to improve the stability of a machine-learning-trained predictive model. In Fig. 3.9, we show the performance of the best and worst model trained with the fourth method that combines "statistical splitting" of the dataset with a bagging approach. The errors of the best model of predicting the training and test dataset are 13.80 K and 10.21 K, respectively. For the worst model, the corresponding errors are 11.86 K and 27.37 K, as shown in Fig. 3.9(b).

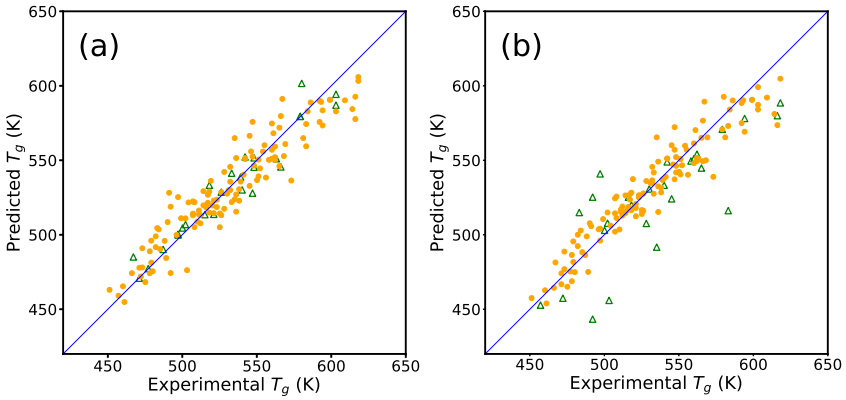

Figure 3.9: Performance of the (a) best and (b) worst model from training method #4 ("statistical splitting" + bagging).

To compare the various ways of training the predictive model of T_g, we performed a statistical analysis of the errors of the 1000 models when they were used to predict the test dataset. The average and standard deviation of these errors are included in Table 3.3. The results show that when bagging is used, both average and standard deviation of the errors are reduced. Bagging thus enhances the stability of the machine-learning-trained model. Furthermore, the training methods in which "statistical splitting" is used to make the training dataset more statistically representative of the entire dataset also lead to predictive models with better performance. The trends indicate that the best training method is to use "statistical splitting" coupled with bagging, i.e, the fourth method.

Table 3.3: Performance metrics of the predictive models of T_g

Training Method	Average Error (K)	Standard Deviation of Error (K)	Correlation
Method #1	18.58	3.09	0.71
Method #2	18.31	2.98	0.84
Method #3	18.17	2.87	0.68
Method #4	17.98	2.62	0.83

In each splitting of the entire dataset into a training and a test set, a predictive model of T_g was generated. This model is a linear mapping from all Z features generated for the dianhydide and diamine groups to T_g, with the coefficient of k-th feature denoted as $M(k)$. A larger absolute value of $M(k)$ implies that the corresponding k-th feature is more strongly correlated to T_g. For two models, a correlation can thus be defined as

$$c_{ij} = \frac{2\sum_{k=1}^{Z} M_i(k)M_j(k)}{\sum_{k=1}^{Z}\left[M_i^2(k) + M_j^2(k)\right]} , \tag{3.9}$$

where i and j are the indices of the models. If the two models are identical, then $c_{ij} = 1$. If the two models are anticorrelated with $M_i(k) = -M_j(k)$, then $c_{ij} = -1$. If a training method of the predictive model of T_g is stable, then different splittings will lead to models that are highly correlated, with the correlation between the models close to 1.

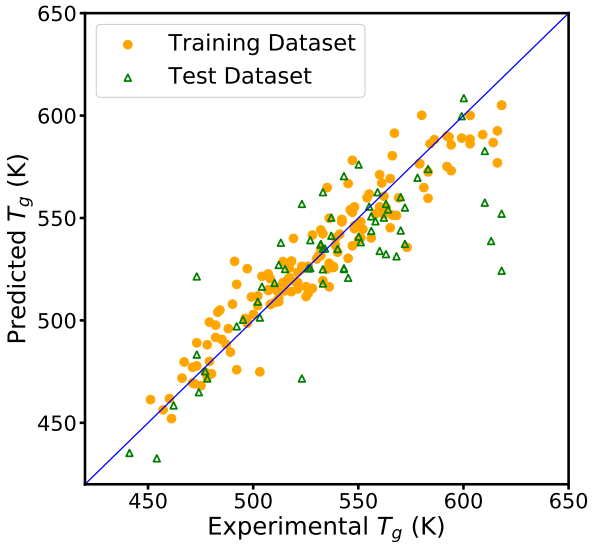

Figure 3.10: Performance of the predictive model of T_g trained with 151 data points from Ref. [142] and tested against 63 data points from Ref. [143].

We computed the correlations of all pairs out of the 1000 models generated with one of the four training methods discussed previously, using Eq. (3.9). The average correlation for each training method is included in Table 3.3. It is clear that the bagging method significantly increases the correlation between the resulting predictive models and thus enhances the stability of the training process.

Finally, we used the fourth method (i.e., "statistical splitting" + bagging) to train a predictive model of T_g with all 151 data points from Ref. [142]. The model was then used to predict the 63 data points from Ref. [143]. Since the training and test datasets in this case are from two different sources, this test serves as an independence check of the training method. The resulting predictive model of T_g has an average error of 25.5 K of predicting the test set, as shown in Fig. 3.10. However, large errors mainly occur for high T_g around 615 K. In the lower range of T_g, Fig. 3.10 indicates that the predictive model performs well in terms of predicting the independent dataset from a different source. The coefficients of features that

enter the predictive model are included in Appendix D.

3.4 Conclusions

Our results show that the PCFF force field combined with Mulliken charges can be used in all-atom MD simulations to compute and estimate T_g of polyimides. The determination of T_g can be achieved by computing either the polymer density, ρ, or diffusion coefficients of gas molecules, D, in the polymer matrix as a function of temperature. For temperatures lower or higher than T_g, ρ exhibits a linear dependence on T but the slopes are different. D, on the other hand, depends on T as $\mathrm{Log}D \propto 1/T$ and the linear coefficients are again different for $T < T_g$ and $T > T_g$. The comparison shows that in practice, D can be more reliably computed and used to give a more accurate estimate of T_g. However, several limitations of using all-atom MD simulations to compute T_g should be noted. First, the cooling rate used in MD simulations is typically several orders of magnitude larger than experimental rates. Secondly, the molecular weight of the polymers modeled in all-atom MD simulations is usually smaller than experimental values by a factor of 10 to 100. Lastly, the PCFF force field is a generic force field for polymers and not specifically designed and optimized for polyimides. All these issues point to the need of going beyond all-atom MD simulations and of seeking a predictive model that can be used to quickly estimate T_g of polyimides.

A predictive model of T_g of polyimides can be obtained by applying machine-learning algorithms to analyze the available experimental and simulation data on T_g. We demonstrate a machine-learning approach to systematically derive such predictive models, including using a SMILES notation to designate a polymer, feature generation by reading in the SMILES notation, removal of irrelevant and low-importance features through LASSO regularization, and improving and optimizing the predictive models via bagging. For polyimides, we have

explored 4 different training methods to construct a predictive model of T_g of polyimides using data collected from Ref. [142] and found that the best model is obtained if the entire dataset is split into the training and test sets that are statistically representative of the entire set and if bagging is used to improve the stability of the predictive model. We further demonstrate that this model can be successfully applied to accurately predict the results on T_g reported in Ref. [143], which are never used to train the predictive model. In the future, it is interesting to further improve the predictive model of T_g by training it with a larger dataset and taking into account the differences in the experimental conditions under which T_g is measured. It is also interesting to explore if similar predictive models of other physical quantities of interest, such as the dielectric constant and mechanical moduli, can be developed for polyimides.

Chapter 4

Polymerization of Branched Polyetherimides: Comparison between Monte Carlo Simulation and Flory-Stockmayer Theory

This chapter is based on a paper that is currently under review:

Chengyuan Wen, Roy Odle, and Shengfeng Cheng, "Polymerization of Branched Polyetherimides: Comparison between Monte Carlo Simulation and Flory-Stockmayer Theory".

I designed and built all the models for Monte Carlo simulations. I performed all the calculations on the basis of the Flory-Stockmayer theory. I performed all the data analyses and prepared figures. Dr. Odle from SABIC contributed some ideas and provided experimental insights. All authors contributed to the writing of this paper. My contributions to this paper were under Dr. Cheng's supervision.

4.1 Introduction

The molecular weight distribution and architecture are two important characteristics of a system of polymer chains [150]. They strongly affect material properties such as dynamic moduli, fracture toughness, glass transition temperature, and viscosity [150–152]. Experimental methods for an accurate determination of molecular weight distributions are thus of great interest [152–154]. Theoretically, it is also highly desirable if the molecular weight distribution of a polymer can be predicted *a priori* based on the knowledge of the polymerization reaction without even synthesizing the polymer. Such a theoretical method will be a valuable tool not only useful for understanding experimental measurements but also beneficial for other theories and models aiming to predict polymer properties. For example, Nichetti and Manas-Zloczower proposed a theoretical model to predict viscosity of a polymer melt based on its molecular weight distribution that was determined by fitting the gel permeation chromatography data to statistical distribution functions [155]. A method to quickly generate molecular weight distributions before polymers are synthesized thus might be able to advance the predictive capability of such theories.

A theory on the constitution and molecular size distribution of a step-growth polymer was proposed by Flory and Stockmayer many years ago [156–161] and has been frequently used to determine the gel point. Flory studied the polymerization of bifunctional monomers mixed with trifunctional and tetra functional branching units and made two fundamental assumptions. First, the same functional group has the same probability to react with another group and this probability is not affected by the length of the polymer to which the functional group belongs as well as the position of the functional group on that polymer. Secondly, ring polymers are not formed. Stockmayer extended the theory to branching units with arbitrary functionalities and derived the Stockmayer formula for the number of a polymer

with a given composition, though ring structures were still excluded. The predictions of the Flory-Stockmayer theory, including the gel point and the average molecular weight, have been tested experimentally [162–168]. However, the entire distribution is hard to probe experimentally and often only some average molecular weight is measured. Practically, it is also difficult to directly predict the molecular weight distribution using the Flory-Stockmayer theory because of mathematical complexity of computing the numbers of all possible molecules in a branched polymer system. Furthermore, the Flory-Stockmayer theory is only expected to be valid below the gel point. Beyond the gel point, the formation of cyclics and closed loops in network structures (i.e., branching) becomes significant and the theory may fail [169, 170].

Monte Carlo (MC) methods are a class of techniques based on random sampling to numerically solve problems that have a probabilistic interpretation [171]. MC methods have broad applications in polymer science [172, 173], especially in polymer reaction engineering [173]. Johnson and O'Driscoll used MC simulation to study sequence distributions in step-growth copolymerization [174]. Tobita applied MC simulation to a wide range of polymerization problems, including free-radical cross-linking copolymerization [175], emulsion polymerization [176], the modification of polymer via crosslinking and degradation [177], long-chain branching and random scission [178], and living radical polymerization [179, 180]. Hadicke and Stutz used an amine-cured epoxy as an example to compare the structure of step-growth networks obtained by MC simulation and branching theory [181]. He *et al.* applied a MC method to simulate self-condensing vinyl polymerization in the presence of multifunctional initiators and probed the role of reactive rate constants [182, 183]. Rouault and Milchev [184] and He *et al.* [185] performed MC simulations to study the kinetic and chain length distributions in living polymerization. Prescott used a MC model to show that chain-length dependent termination plays a significant role in living/controlled free-radical polymeriza-

tion systems containing reversible transfer agents [186]. In a series of papers, Al-Harthi *et al.* used dynamic MC simulations to study atom-transfer radical polymerizations [187–190]. Polanowski *et al.* [191, 192] and Bannister *et al.* [193] used MC methods to study the branching and gelation in living copolymerization of vinyl and divinyl monomers. Recently, Lyu *et al.* used a similar model to study the atom transfer radical and the conventional free radical polymerization of divinyl monomers and checked the applicability of the Flory-Stockmayer theory in such systems [169, 170]. Gao *et al.* used kinetic MC methods to simulate free radical copolymerization processes and discussed how to accelerate such simulations using scaling approaches [194, 195]. Meimaroglou *et al.* proposed a MC algorithm to calculate the molecular weight distribution for linear polymers and the bivariate molecular weight-long chain branching distribution for highly branched polymers [196]. They also used MC simulation to investigate the molecular, topological, and solution properties of highly branched low-density polyethylene [197] and the ring-opening homopolymerization of L,L-Lactide [198]. Iedema *et al.* developed a MC simulation model including both branching and random scission to calculate the molecular weight and branching distribution and compared their calculations to experimental measurements on high-density polyethylene [199]. Yaghini and Iedema compared the results on low-density polyethylene from such MC simulations to the predictions of a multiradical model based on a Galerkin finite element approach [200].

One important application of MC simulations is to quickly compute molecular weight distributions [169, 170, 172, 175–180, 184, 185, 194, 195, 199–205]. In MC simulations, all structures including rings and networks allowed by a polymerization reaction can be produced [206], no matter the system is below or beyond the gel point [207]. Various schemes can also be implemented for the kinetics of the polymerization, which thus allows us to test the specific assumptions made by a theory. The Gillespie algorithm can be used to speed up the kinetics of a reaction *in silico* and enable a reactive system to quickly reach

a steady state [208, 209]. In this chapter, we develop a MC simulation model based on the Gillespie algorithm to study the polymerization of polyetherimides in the presence of chain terminators and branching agents. The results from the MC simulations are used to test the Flory-Stockmayer theory including its assumption on the reaction rates.

This chapter is organized as follows. In Sec. 4.2 the Flory-Stockmayer theory is introduced and the technical challenge of computing molecular weight distributions with this theory is discussed. In Sec. 4.3 we describe the MC model of the polymerization process of branched polyetherimides in detail. In Sec. 4.4, comparisons of results from the MC simulations and the Flory-Stockmayer theory are presented. Practically, computations of molecular weight distributions can be only be executed for a small system either with the Flory-Stockmayer theory or the MC model. We thus include a discussion on the effect of a finite system size in this section. Although the emphasis is on stoichiometric, fully reacted systems, those that are only partially reacted and/or nonstoichiometric are also discussed in Sec. 4.4. Finally, a brief summary is provided in Sec. 4.5.

4.2 Flory-Stockmayer Theory of Step-Growth Polymers

Flory and Stockmayer considered a general step-growth polymer that consists of two types of monomers, A and B. All reactions occur between A and B. There are i type-A monomers denoted as A_1, A_2, ..., A_i. To simplify notation, we also use A_q with $q \in \{1, 2, ..., i\}$ to denote the number of A_q monomers. Similarly, there are j type-B monomers and the corresponding numbers are B_1, B_2, ..., B_j, respectively. The symbol f_q denotes the functionality of an A_q monomer, where $q \in \{1, 2, ..., i\}$, i.e., there are f_q functional groups on an A_q monomer that can form bonds with the corresponding functional groups on a B_h monomer, where $h \in \{1, 2, ..., j\}$. The functionality of a B_h monomer is denoted as g_h. The Flory-Stockmayer

theory can be applied to a polymerized state where a fraction p_A of all the functional groups on the type-A monomers have reacted with a fraction p_B of all the functional groups on the type-B monomers. Therefore,

$$p_A \sum_{q=1}^{i} f_q A_q = p_B \sum_{h=1}^{j} g_h B_h. \tag{4.1}$$

In this chapter, we call the systems with $\sum_{q=1}^{i} f_q A_q = \sum_{h=1}^{j} g_h B_h$ and thus $p_A = p_B$ as stoichiometric systems while those with $\sum_{q=1}^{i} f_q A_q \neq \sum_{h=1}^{j} g_h B_h$ and $p_A \neq p_B$ as nonstoichiometric. The systems with p_A or p_B, or both, equal to 1 are fully reacted.

We use $N\{m,n\}$ to denote the number of molecules formed by m_q monomers of sub-type A_q and n_h monomers of sub-type B_h, with q running from 1 to i and h running from 1 to j. Here $\{m,n\}$ is a shorthand of $\{m_1, m_2., ..., m_i, n_1, n_2, ..., n_j\}$, which denotes the monomer composition of a given molecule. The Flory-Stockmayer theory predicts that

$$\begin{aligned} N\{m,n\} = K & \frac{\left(\sum_{q=1}^{i} f_q m_q - \sum_{q=1}^{i} m_q\right)!}{\left(\sum_{q=1}^{i} f_q m_q - \sum_{q=1}^{i} m_q - \sum_{h=1}^{j} n_h + 1\right)!} \\ & \times \frac{\left(\sum_{h=1}^{j} g_h n_h - \sum_{h=1}^{j} n_h\right)!}{\left(\sum_{h=1}^{j} g_h n_h - \sum_{h=1}^{j} n_h - \sum_{q=1}^{i} m_q + 1\right)!} \\ & \times \prod_{q=1}^{i} \frac{x_q^{m_q}}{m_q!} \prod_{h=1}^{j} \frac{y_h^{n_h}}{n_h!} \end{aligned} \tag{4.2}$$

with

$$x_q = \frac{f_q A_q}{\sum_{l=1}^{i} f_l A_l} \frac{p_B \left(1-p_A\right)^{f_q-1}}{\left(1-p_B\right)} \ , \tag{4.3}$$

$$y_h = \frac{g_h B_h}{\sum_{l=1}^{j} g_l B_l} \frac{p_A \left(1-p_B\right)^{g_h-1}}{1-p_A} \ , \tag{4.4}$$

$$K = \frac{\left(1-p_A\right)\left(1-p_B\right)}{p_B} \sum_{q=1}^{i} f_q A_q = \frac{\left(1-p_A\right)\left(1-p_B\right)}{p_A} \sum_{h=1}^{j} g_h B_h \ . \tag{4.5}$$

Equation (4.2) is called the Stockmayer formula, which gives the number of molecules of any monomer compositions. However, practically it is difficult to compute the molecular weight distribution from the Stockmayer formula, as all the possible combinations for $\{m_1, m_2., ..., m_i, n_1, n_2, ..., n_j\}$ have to be taken into account. Since m_q runs from 1 to A_q for $q \in \{1, 2, ..., i\}$ and n_h runs from 1 to A_h for $h \in \{1, 2, ..., j\}$, the total number of possible molecules is $\prod_{q=1}^{i} A_q! \times \prod_{h=1}^{j} B_h!$. This number is huge when there are many sub-types (i.e., large i and j) and/or large numbers (i.e., large A_q and B_h) of monomers involved in a polymerization.

For a molecule with composition $\{m, n\}$, the total number of monomers is $\sum_{q=1}^{i} m_q + \sum_{h=1}^{j} n_h$. Since the Flory-Stockmayer theory does not consider rings, the total number of bonds in this molecule must be $\sum_{q=1}^{i} m_q + \sum_{h=1}^{j} n_h - 1$. When $p_A = p_B = 1$, all the functional groups have reacted and in a given molecule the total number of the functional groups on all the type-A monomers is equal to the total number of the functional groups on all the type-B monomers. This number must also be equal to the total number of bonds in that molecule. Namely, for $p_A = p_B = 1$ there are two identities,

$$\sum_{q=1}^{i} f_q m_q = \sum_{q=1}^{i} m_q + \sum_{h=1}^{j} n_h - 1 \tag{4.6}$$

and

$$\sum_{h=1}^{j} g_h n_h = \sum_{h=1}^{j} n_h + \sum_{q=1}^{i} m_q - 1. \tag{4.7}$$

These two identities can help us simplify the Stockmayer formula for stoichiometric, fully reacted systems. Note that the terms in Eq. (4.2) involving $(1-p_A)$ and $(1-p_B)$ appear as

$$(1-p_A)^{\sum_{q=1}^{i} f_q m_q - \sum_{q=1}^{i} m_q - \sum_{h=1}^{j} n_h + 1}$$

and

$$(1-p_B)^{\sum_{h=1}^{j} g_h n_h - \sum_{h=1}^{j} n_h - \sum_{q=1}^{i} m_q + 1} \quad .$$

These terms can be dropped out because of Eqs. (4.6) and (4.7). As a result, for fully reacted stoichiometric systems with $p_A = p_B = 1$ the Stockmayer formula is simplified as

$$N\{m,n\} = K \left(\sum_{q=1}^{i} f_q m_q - \sum_{q=1}^{i} m_q \right)! \left(\sum_{h=1}^{j} g_h n_h - \sum_{h=1}^{j} n_h \right)! \prod_{q=1}^{i} \frac{x_q^{m_q}}{m_q!} \prod_{h=1}^{j} \frac{y_h^{n_h}}{n_h!} \tag{4.8}$$

with

$$x_q = \frac{f_q A_q}{\sum_{l=1}^{i} f_l A_l} \ , \tag{4.9}$$

$$y_h = \frac{g_h B_h}{\sum_{l=1}^{j} g_l B_l} \ , \tag{4.10}$$

$$K = \sum_{q=1}^{i} f_q A_q = \sum_{h=1}^{j} g_h B_h \ . \tag{4.11}$$

Computing $N\{m,n\}$ is not easy as it contains many factorials. The calculation can be

expedited using the Stirling approximation

$$\log n! \approx \log\left(\sqrt{2\pi n}\right) + n\log\left(\frac{n}{e}\right) + \log\left(1 + \frac{1}{12n}\right) \tag{4.12}$$

Then for fully reacted stoichiometric systems, the Stockmayer formula can be approximated logarithmically as

$$\begin{aligned}\log N\{m,n\} \approx &\log K + \log\left(\sum_{q=1}^{i} f_q m_q - \sum_{q=1}^{i} m_q\right)! + \log\left(\sum_{h=1}^{j} g_h n_h - \sum_{h=1}^{j} n_h\right)! \\ &+ \sum_{q=1}^{i}\left(m_q \log x_q - \log m_q!\right) + \sum_{h=1}^{j}\left(n_h \log y_h - \log n_h!\right)\end{aligned} \tag{4.13}$$

The computation of the molecular weight distribution from $N\{m,n\}$ can be further accelerated by noting that not all the combinations $\{m,n\}$ will yield a molecule. For a fully reacted stoichiometric system where $p_A = p_B = 1$, Eqs. (4.1), (4.6), and (4.7) can be used to reduce the total number of $\{m,n\}$. For the branched polyetherimides considered in this chapter (see Sec. 4.3), $f_1 = 1$, $f_2 = 2$, $g_1 = 2$ and $g_2 = 3$. The constraints become

$$m_1 + 2m_2 = 2n_1 + 3n_2 \tag{4.14}$$

and

$$m_2 = n_1 + n_2 - 1 \tag{4.15}$$

Combining Eqs. (4.14) and (4.15), we get

$$m_1 = n_2 + 2 \tag{4.16}$$

Equations (4.15) and (4.16) indicate that m_1 and m_2 are totally constrained by n_1 and n_2 in an allowed composition. Furthermore, since $m_2 \geq 0$, n_1 and n_2 cannot be zero at the same time. The time complexity to enumerate all possible molecules is thus $\mathcal{O}(B_1 B_2)$, which is about $\mathcal{O}(Z^2)$ with Z being the system size (i.e., the total number of monomers prior to polymerization). This time complexity is acceptable for small systems. However, if there are more types of monomers, then the time complexity will increase exponentially as $\mathcal{O}(Z^w)$, where w is the number of monomer types. For not fully reacted or nonstoichiometric systems where p_A or p_B are less than 1, we lose the constrains that help reduce the number of possible $\{m, n\}$ and then computing the molecular weight distribution from $N\{m, n\}$ has to use Eq. (4.2) and will become more challenging, even though the Stirling approximation may still be used. In these situations, the MC model described below will serve as a solution as it does not suffer from such limitations and the time complexity of computing the molecular weight distribution with MC simulations is always $\mathcal{O}(Z)$.

4.3 Monte Carlo Model of Polymerization of Branched Polyetherimides

Four types of monomers are involved in the formation of branched polyetherimides, including 4,4'-bisphenol A dianhydride (BPADA), m-phenylenediamine (MPD), phthalic anhydride (PA), and tris[4-(4-aminophenoxy)phenyl] ethane (TAPE).[210, 211] The chemical structures of these monomers are shown in Fig. 4.1. The involved reaction is the condensation reaction between an amine group on MPD or TAPE and a carboxylic anhydride group on BPADA or PA. In the notation of the Flory-Stockmayer theory, PA is monomer A_1 with $f_1 = 1$, BPADA is monomer A_2 with $f_2 = 2$, MPD is monomer B_1 with $g_1 = 2$, and TAPE is monomer B_2 with $g_2 = 3$. Out of these monomers, PA is an end capper to terminate a chain and TAPE

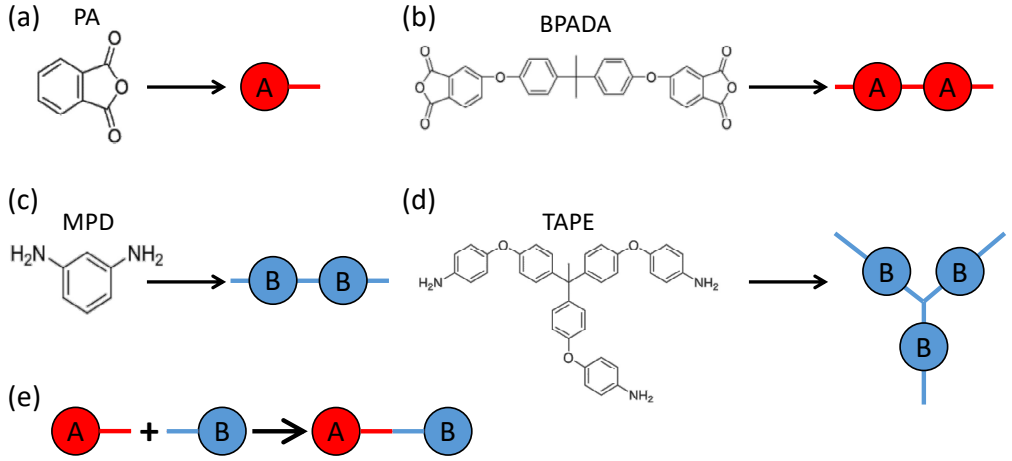

Figure 4.1: (a)-(d): The representation of the four types of monomers of branched polyetherimides in the MC simulation model. Each functional group containing one amine is mapped to a B bead. Each functional group containing one carboxylic anhydride is mapped to an A bead. (e): Each A bead can form a bond with a B bead, mimicking the condensation reaction between an amine group and a carboxylic anhydride group in the polymerization of polyetherimides.

is a trifunctional branching agent. Fig. 4.1 shows the representation of these monomers in our MC model. Each functional group containing one carboxylic anhydride is mapped to an A bead and that containing one amine is mapped to a B bead. Each A bead can react with a B bead to form a bond (i.e., $A + B \rightarrow AB$), which describes the condensation reaction between an amine group and a carboxylic anhydride group.

In the formation of branched polyetherimides consisting of the above 4 types of monomers, there are 4 possible reactions, as sketched in Fig. 4.2. Reaction 1 is between BPADA and MPD, which leads to the formation of a polyetherimide backbone. Reaction 2 is between BPADA and TAPE that results in branching. Reaction 3 is between PA and MPD, which terminates a polyetherimide chain. Reaction 4 is between PA and TAPE, which consumes one amine group on TAPE and effectively reduces its functionality by 1.

With the mapping in Fig. 4.1 and the reaction scheme in Fig. 4.2, we perform MC simulations to study the polymerization of branched polyetherimides. We adopt the Gillespie algorithm

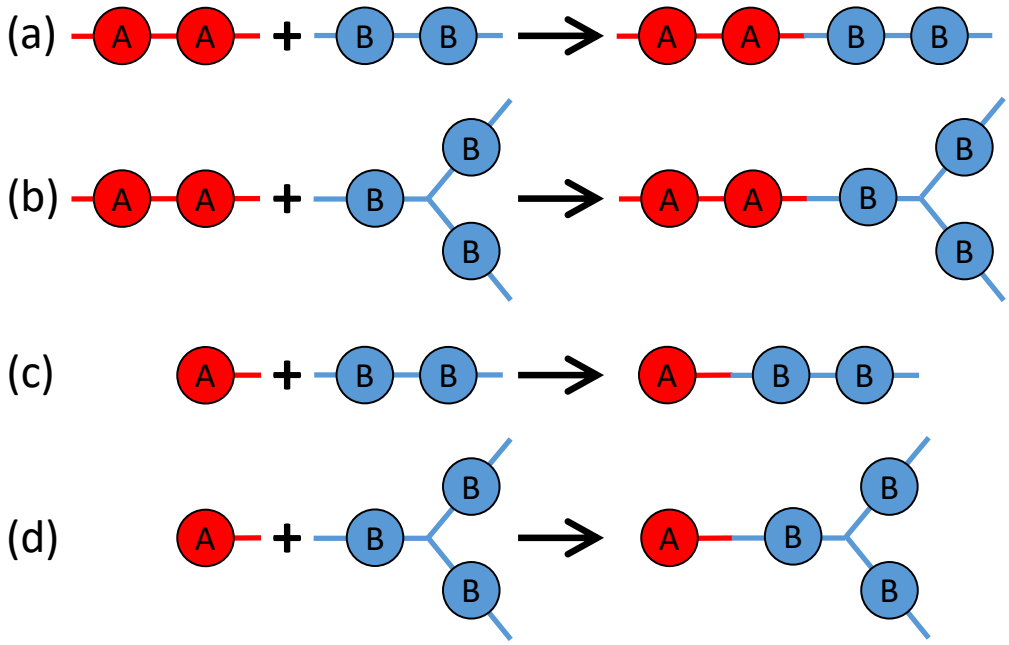

Figure 4.2: The four reactions occurring in the polymerization of branched polyetherimides: (a) Reaction 1: BPADA + MPD, (b) Reaction 2: BPADA + TAPE, (c) Reaction 3: PA + MPD, and (d) Reaction 4: PA + TAPE.

to speed up MC simulations. Since only the final chain structures are concerned, we neglect the random process in the typical Gillespie algorithm that determines the time interval after which the next reaction occurs. We only keep the random process of picking a reaction at a time. At each MC step, all four reactions will have a probability to occur and the reaction rate of a particular reaction is determined by a rate constant and the concentration of the two types of monomers involved in that reaction. Mathematically, the probability of reaction l is proportional to

$$P_l(L_l + R_l \rightarrow L_l R_l) = k_l n_{L_l} n_{R_l} \ , \tag{4.17}$$

where L_l (R_l) represents the reactant consisting of A (B) beads, k_l is a rate constant, n_{L_l} (n_{R_l}) is a quantity that depends on the concentration of the reactant L_l (R_l), and $l \in \{1, 2, 3, 4\}$ indexes the possible reactions sketched in Fig. 4.2. Specifically, L_1 and L_2 are

BPADA, L_3 and L_4 are PA, R_1 and R_3 are MPD, and R_2 and R_4 are TAPE. Since all the 4 reactions can be reduced to the reaction between an A bead and a B bead (i.e., the reaction between a functional group containing one amine and another functional group containing one carboxylic anhydride) as shown in Fig. 4.1(e), k_l will be set as a constant k for all the 4 reactions and n_{L_l} and n_{R_l} can be expressed as

$$\begin{aligned} n_{L_1} &= n_{L_2} = 2n_{\text{BPADA}} \ , \\ n_{L_3} &= n_{L_4} = n_{\text{PA}} \ , \\ n_{R_1} &= n_{R_3} = 2n_{\text{MPD}} \ , \\ n_{R_2} &= n_{R_4} = 3n_{\text{TAPE}} \ , \end{aligned} \tag{4.18}$$

where n_{BPADA}, n_{PA}, n_{MPD}, and n_{TAPE} are the concentrations of monomers available for reactions (i.e., monomers with at least one unreacted functional group). In other words, n_{L_l} (n_{R_l}) is the concentration in terms of the number of A (B) beads on the reactant L_l (R_l). The particular reason of writing n_{L_l} and n_{R_l} in this way will be discussed in Sec. 4.4.1.

At each MC step, the probability of Reaction l to be chosen is equal to $P_l / \sum_{q=1}^{4} P_q$. After a reaction is selected, a pair of L_l and R_l that have unreacted functional groups (i.e., with unreacted A and B beads, respectively) is randomly chosen to react. Then the system status is updated, including the bond information between the monomers and the identity of monomers with unreacted functional groups. The MC process is repeated for the updated system until no more reactions can occur or the system has reached a desired extent of reaction. The flow chart of the MC simulation model is shown in Fig. 4.3. Note that in this model, we made a simplification by not allowing backward reactions, which means that once formed, the bond between an A bead and a B bead cannot break. However, the model permits the formation of both rings and networks.

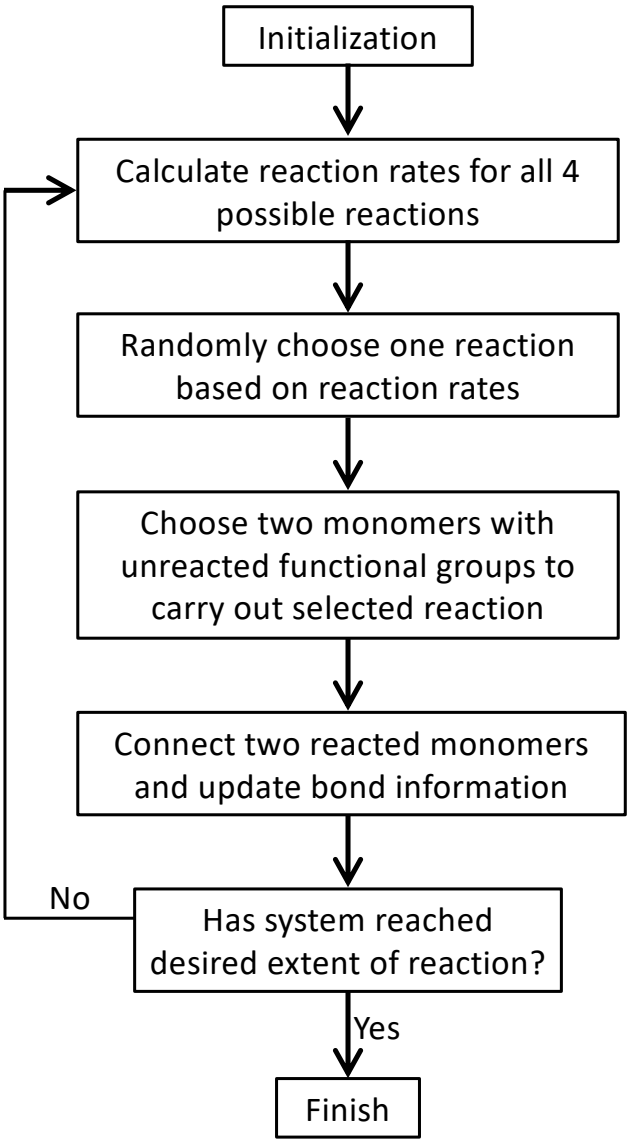

Figure 4.3: The flow chart of the MC simulation model

4.4 Results and Discussion

4.4.1 Rate Constant k

Equation (4.18) indicates that the reaction rate P_l is based on the concentrations of the functional groups (i.e., A beads or B beads) on the reactants involved in that reaction. However, P_l can also be computed from the concentrations of the available reactants themselves, i.e., the monomer concentrations. In this case, the reaction rate P_l can be written in the same way as in Eq. (4.17) but with Eq. (4.18) replaced by

$$\begin{aligned} n_{L_1} &= n_{L_2} = n_{\text{BPADA}} \ , \\ n_{L_3} &= n_{L_4} = n_{\text{PA}} \ , \\ n_{R_1} &= n_{R_3} = n_{\text{MPD}} \ , \\ n_{R_2} &= n_{R_4} = n_{\text{TAPE}} \ . \end{aligned} \tag{4.19}$$

To check which way of computing the reaction rates yields results that are more applicable to realistic systems, we performed a test with a simple system consisting of only PA and TAPE monomers, as shown in Table 4.1. For this system, there are only 4 possible final products, including single TAPEs and TAPEs connected with one, two, or three PAs, respectively.

Table 4.1: System used for checking the way to compute the reaction rates.

Monomer	PA	BPADA	MPD	TAPE
Number	2000	0	0	1000

Figure 4.4 shows the results on the probability distribution of the 4 final products for the system in Table 4.1, for which gelation is not a concern. The comparison shows that the results from the MC simulations based on Eq. (4.18) agree with the Flory-Stockmayer theory

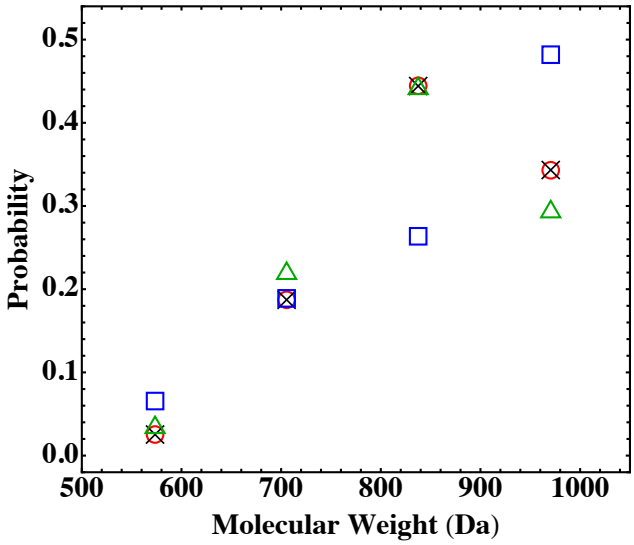

Figure 4.4: Probability distribution of the 4 possible final products for the system in Table 4.1. Results are from the Flory-Stockmayer theory (red circles), MC simulations based on Eq. (4.18) (black crosses), MC simulations based on Eq. (4.19) (blue squares), and a simple statistical model discussed in the main text (green triangles). The MC results are averages of 10,000 runs.

while those based on Eq. (4.19) do not. Furthermore, for the molar ratio in Table 4.1, all the anhydride groups on the PA monomers are reacted and each amine group on a TAPE monomer has a 2/3 chance to be reacted in a fully reacted system. A simple statistical analysis shows that the probabilities for a TAPE monomer to react with 0, 1, 2, and 3 PAs are $\left(\frac{1}{3}\right)^3$, $3 \times \frac{2}{3} \times \left(\frac{1}{3}\right)^2$, $3 \times \left(\frac{2}{3}\right)^2 \times \frac{1}{3}$, and $\left(\frac{2}{3}\right)^3$, respectively. These results are also plotted in Fig. 4.4 and are close to those from the Flory-Stockmayer theory and the MC simulations based on Eq. (4.18). The small differences are due to the fact that the theory and simulations consider a finite system while the statistical model assumes an infinite system. We conclude that the reactions rates based on Eq. (4.18) should be used in the MC simulations. From now on all the data presented in this chapter are computed with Eq. (4.18) for the reaction rates. In the next two sections, Sec. 4.4.2 and Sec. 4.4.3, we focus on fully reacted stoichiometric systems where $p_A = p_B = 1$. We discuss partially reacted stoichiometric systems where $p_A = p_B < 1$ in Sec. 4.4.4 and nonstoichiometric systems where $p_A \neq p_B$ in Sec. 4.4.5.

4.4.2 Fully Reacted Stoichiometric Systems

For the branched polyetherimides considered in this chapter, type A monomers are BPADA and PA and type B monomers are MPD and TAPE, with $f_1 = 1$, $f_2 = 2$, $g_1 = 2$ and $g_2 = 3$. From the Flory theory [156], the gel point is $\alpha_c = 1/(g_2 - 1) = 1/2$. However, for the systems at hand the expression of α, which characterizes the level of cross-linking, has to be modified because each PA monomer as a chain terminator has only one functional group. The modified expression is

$$\begin{aligned}\alpha &= \sum_{q=0}^{\infty} [p_A(1-\rho_1)p_B(1-\rho_2)]^q \, p_A(1-\rho_1)p_B\rho_2 \\ &= p_A p_B \frac{(1-\rho_1)\rho_2}{1 - p_A p_B(1-\rho_1)(1-\rho_2)} \ ,\end{aligned} \tag{4.20}$$

where ρ_1 is the fraction of functional groups on the terminators (i.e., PA monomers) with respect to all the functional groups on type A monomers and ρ_2 is the fraction of functional groups on the branching agents (i.e., TAPE monomers) with respect to all the functional groups on type B monomers. For a fully reacted stoichiometric system, $p_A = p_B = 1$ and this expression can be simplified as

$$\alpha = \frac{(1-\rho_1)\rho_2}{\rho_1 + \rho_2 - \rho_1\rho_2} \ . \tag{4.21}$$

We can vary the numbers of monomers to tune ρ_1 and ρ_2, thus putting the fully reacted system below, around, or beyond the gel point. Three such systems are listed in Table 4.2, where ρ_2 is changed by varying the numbers of MPD and TAPE monomers. In the MC simulations of these stoichiometric systems, we set $p_A = p_B = 1$, thus allowing the systems to be fully reacted.

Table 4.2: Three fully reacted, stoichiometric systems below, around, and beyond the gel point. The first column is the system label. The next 4 columns list the number of each type of monomers. The values of ρ_1 and ρ_2 are determined from the monomer numbers. The value of α is computed using Eq. (4.20). The average molecular weights, M_n, M_w, and M_z, are from the MC simulations.

	PA	BPADA	MPD	TAPE	p_A	p_B	α	M_n	σ_{M_n}	M_w	σ_{M_w}	M_z	σ_{M_z}
$S_<$	50	670	680	10	1	1	0.36643	19120	1045	52126	17245	78671	32596
$S_\simeq$	50	670	671	16	1	1	0.48064	22000	1460	77353	30668	116227	49890
$S_>$	50	670	620	50	1	1	0.74307	51055	12536	330124	45800	369588	34765

Table 4.3: Four fully reacted, stoichiometric systems all below the gel point but with the size increased proportionally without changing the fraction of each type of monomers. The entries have the same format as in Table 4.2. The subscript of the system label in the first column indicates the size ratio with respect to the base system, S_1.

	PA	BPADA	MPD	TAPE	p_A	p_B	α	M_n	σ_{M_n}	M_w	σ_{M_w}	M_z	σ_{M_z}
S_1	10	134	136	2	1	1	0.36643	15759	3096	30840	10296	37355	12745
S_{10}	100	1340	1360	20	1	1	0.36643	19799	566	59940	19181	101321	45570
S_{50}	500	6700	6800	100	1	1	0.36643	20361	121	73582	19585	161904	82601
S_{80}	800	10720	10880	160	1	1	0.36643	20417	80	75980	17403	177919	85751

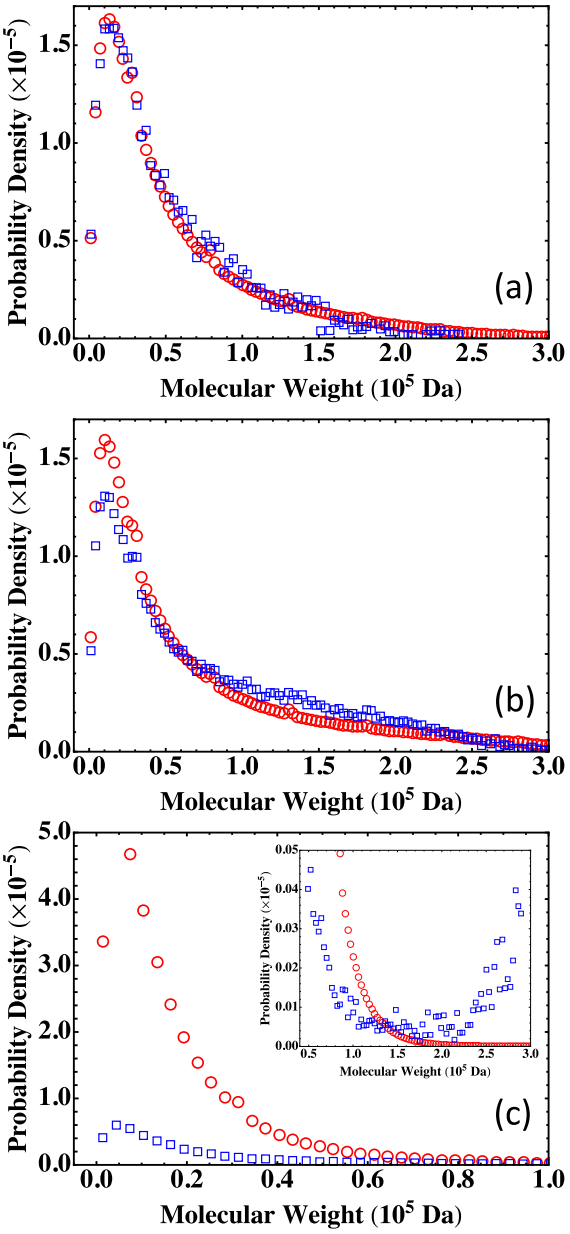

Figure 4.5: Molecular weight distribution for the three systems in Table 4.2: (a) $S_<$, (b) $S_\simeq$, and (c) $S_>$. The results are for the Flory-Stockmayer theory (red circles) and the MC simulations (blue squares). The MC results are averages of 10,000 runs.

The results on the molecular weight distribution from the Flory-Stockmayer theory and the MC simulations are shown in Fig. 4.5. The comparison shows that for a system below the gel point such as $S_<$ (Fig. 4.5(a)), the MC results agree well with the Flory-Stockmayer theory. For $S_\simeq$ which is close to the gel point, the Flory-Stockmayer theory overestimates the fraction of low molecular weight polymers and underestimates the fraction of high molecular weight species when compared to the results from the MC simulations, as shown in Fig. 4.5(b). The discrepancy between the Flory-Stockmayer theory and the MC results becomes more dramatic for systems above the gel point. For $S_>$, $\alpha = 0.74308$, way above the critical gel point $\alpha_c = 0.5$. The Flory-Stockmayer theory predicts a probability density that is about 8 times of the MC result in the region of low molecular weight from 0 to about 0.5×10^5 Da, as shown in Fig. 4.5(c). However, the MC simulations show a significant fraction of polymers in the region of molecular weight higher than about 1.5×10^5 Da and these high molecular weight polymers are completely overlooked by the Flory-Stockmayer theory, as shown in the inset of Fig. 4.5(c). This discrepancy is not surprising as beyond the gel point, polymers with a large network structure are expected and closed loops can frequently emerge in such polymers. The Flory-Stockmayer theory does not consider the formation of rings and thus cannot accurately predict the molecular weight distribution for systems above the gel point.

4.4.3 Effect of System Size

In experiments, the amount of monomers involved is at the order of moles, i.e., at the order of 10^{23}. It is thus practically impossible to directly compute the molecular weight distribution from the Stockmayer formula (Eq. (4.2)) for such macroscopic systems. These systems are also out of the reach of MC simulations that typically deals with a few hundred to a few thousand monomers. A natural question we can ask is: if we keep the molar ratios unchanged but reduce the numbers of participating monomers, can we use either the Flory-Stockmayer

theory or the MC simulations to generate a molecular weight distribution that is applicable to a macroscopic system? To answer this question, we test 4 additional systems listed in Table 4.3. The smallest system has 10 PA, 134 BPADA, 146 MPD, and 2 TAPE and is denoted as S_1. Then the numbers of monomers are increased 10, 50, and 80 fold by keeping the ratios to generate systems S_{10}, S_{50}, and S_{80}. The subscript of the system label thus reflects the size ratio with respect to the smallest system, S_1. In this notation, the system $S_<$ in Table 4.2 is equivalent to S_5. All these systems are still below the gel point when fully reacted.

The molecular weight distributions predicted by the Flory-Stockmayer theory for S_1, $S_<$, S_{10}, S_{50}, and S_{80}, including the probability density and the cumulative probability, are shown in Fig. 4.6. The main panels are for the region of low molecular weight and the insets show the data in the high molecular weight region. The data show that when the system size is increased, the curves of the molecular weight distribution converge quickly. There is a clear difference between the data for S_1 and those for $S_<$ (i.e., S_5). However, the difference between $S_<$ and S_{80} is very small in the low molecular weight region and only discernible in the tail of the distribution in the region of high molecular weight (see the insets of Fig. 4.6). Furthermore, the results for S_{50} and S_{80} are almost indistinguishable in the entire region of molecular weight relevant to experiments, indicating that these systems are already large enough such that the molecular weight distribution is not affected by the finite system size any more.

Since S_1, $S_<$, S_{10}, S_{50}, and S_{80} are all below the gel point, we expect the results from the Flory-Stockmayer theory and the MC simulations on the molecular weight distribution to agree. The comparison between the two is shown in Fig. 4.7 for S_1 and S_{80}. For S_1, some difference is observed between the prediction of the Flory-Stockmayer theory and the MC result because of the small size of this system. An excellent agreement is found between the

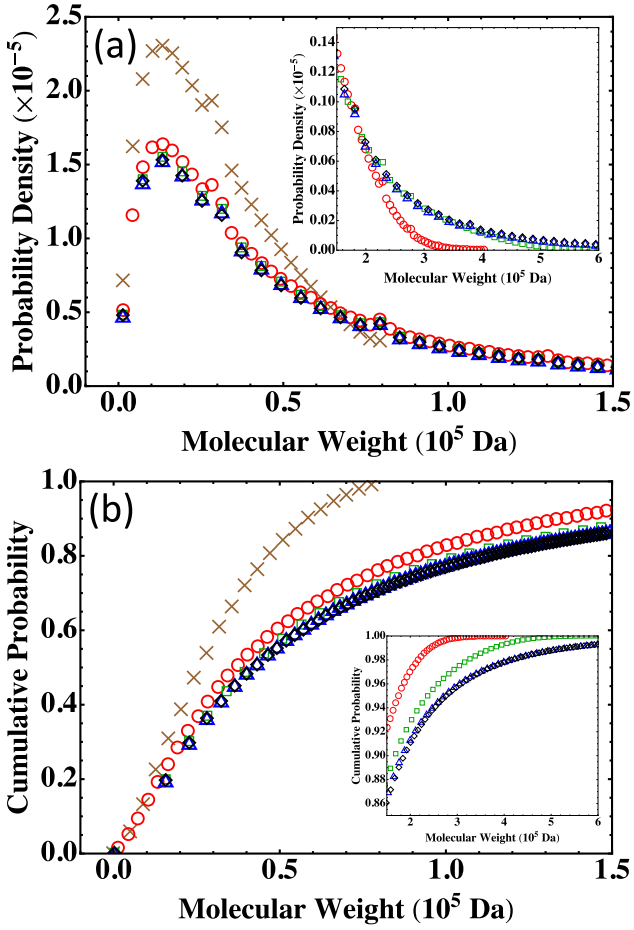

Figure 4.6: Molecular weight distribution predicted by the Flory-Stockmayer theory for systems with different sizes: (a) Probability density and (b) cumulative probability. The main panels show the data in the low molecular weight region while the insets show the data in the high molecular weight region. Data are for S_1 (brown crosses), $S_<$ (red circles), S_{10} (green squares), S_{50} (blue triangles), and S_{80} (black diamonds).

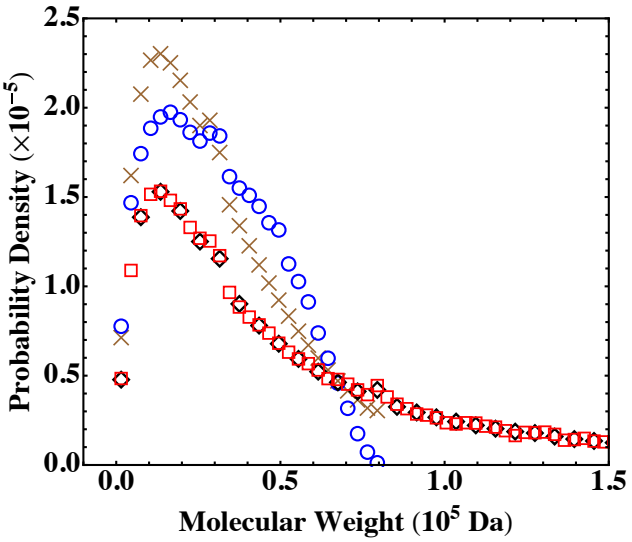

Figure 4.7: Molecular weight distribution with data from for the Flory-Stockmayer theory for S_1 (brown crosses), the MC simulations for S_1 (blue circles), the Flory-Stockmayer theory for S_{80} (black diamonds), and the MC simulations for S_{80} (red squares). The MC results are averages of 10,000 runs.

theory and simulations for S_{80}. Similar agreements are also found for S_{10} and S_{50}. A good agreement is already discussed earlier for $S_<$ as shown in Fig. 4.5(a). These comparisons once again confirm that the Flory-Stockmayer theory provides a good description of the molecular weight distribution for systems well below the gel point, where ring formation is not a big concern. Below the gel point, both the Flory-Stockmayer theory and the MC simulations can be applied to a system containing only a few hundred to a few thousand monomers but having the same molar ratios of monomers as a macroscopic system to accurately predict the molecular weight distribution. As discussed earlier, the Flory-Stockmayer theory starts to fail when a system approaches or goes above the gel point. However, in these situations the MC simulations can still be used to quickly generate a molecular weight distribution that is applicable to an experimental system.

4.4.4 Partially Reacted Stoichiometric Systems

Up to this point, we mainly focus on fully reacted stoichiometric systems as it is possible to compute the molecular weight distribution using the Stockmayer formula even for a system with a relatively large size such as S_{80}. In this and next section we show that the conclusions reached so far also apply to partially reacted and/or nonstoichiometric systems. However, because of the practical difficulty of using the Stockmayer formula to compute the molecular weight distribution when either p_A or p_B, or both, are less than 1, we use small systems with sizes similar to S_5 to illustrate the main point.

In this section we discuss partially reacted stoichiometric systems where $\sum_{q=1}^{i} f_q A_q = \sum_{h=1}^{j} g_h B_h$ but $p_A = p_B < 1$. Five such systems with the same size as $S_>$ are listed in Table 4.4 where the values of p_A and p_B are increased from 0.95 to 0.99. The corresponding values of α changes from about 0.42 to about 0.65, thus enclosing the gelation transition at $\alpha_c = 0.5$.

The results on the molecular weight distribution from the Flory-Stockmayer theory and MC simulations at various values of p_A and p_B are shown in Figs. 4.8(a) and (b), respectively. The molecular weight distribution predicted by the Flory-Stockmayer theory seems to be relatively insensitive to the values of p_A and p_B. However, the MC results show that when the value of p_A and p_B is increased, the probability density in the low molecular weight region is reduced (see Fig. 4.8(b)) while that in the high molecular weight region is enhanced (see the inset of Fig. 4.8(b)). This systematic trend is expected as when the extent of reaction is larger, more polymers with higher molecular weights are anticipated to form.

To compare the predictions of the Flory-Stockmayer theory and the MC simulations on the molecular weight distribution, in Fig. 4.8(c) their differences are shown for various p_A and p_B. It is clear that when p_A and p_B are small, the systems are below the gel point and the

Table 4.4: Five partially reacted, stoichiometric systems (i.e., $p_A = p_B < 1$). The entries have the same format as in Table 4.2. The superscript of the system label indicates the values of p_A and p_B. The first two are below and the rest three are beyond the gel point.

	PA	BPADA	MPD	TAPE	p_A	p_B	α	M_n	σ_{M_n}	M_w	σ_{M_w}	M_z	σ_{M_z}
$S^{0.95}$	50	670	620	50	0.95	0.95	0.419416	5965	99	24779	10468	47952	26051
$S^{0.96}$	50	670	620	50	0.96	0.96	0.462209	7386	165	36488	15428	69368	34683
$S^{0.97}$	50	670	620	50	0.97	0.97	0.512934	9655	308	58237	25507	103512	47567
$S^{0.98}$	50	670	620	50	0.98	0.98	0.574009	13780	729	105576	45386	165690	65778
$S^{0.99}$	50	670	620	50	0.99	0.99	0.648951	22710	2213	200514	60560	266624	65880

Table 4.5: Three partially reacted, nonstoichiometric systems (i.e., $p_A \neq p_B$ and both are less than 1) below, around, and beyond the gel point. The entries have the same format as in Table 4.2.

	PA	BPADA	MPD	TAPE	p_A	p_B	α	M_n	σ_{M_n}	M_w	σ_{M_w}	M_z	σ_{M_z}
$S^n_<$	50	670	664	50	0.99	0.93	0.445	7225	158	33130	13759	62388	31644
$S^n_\approx$	50	670	649	50	0.99	0.95	0.502	9530	299	54269	24699	96424	46537
$S^n_>$	50	670	634	50	0.99	0.97	0.569	13824	719	103074	44117	162837	64716

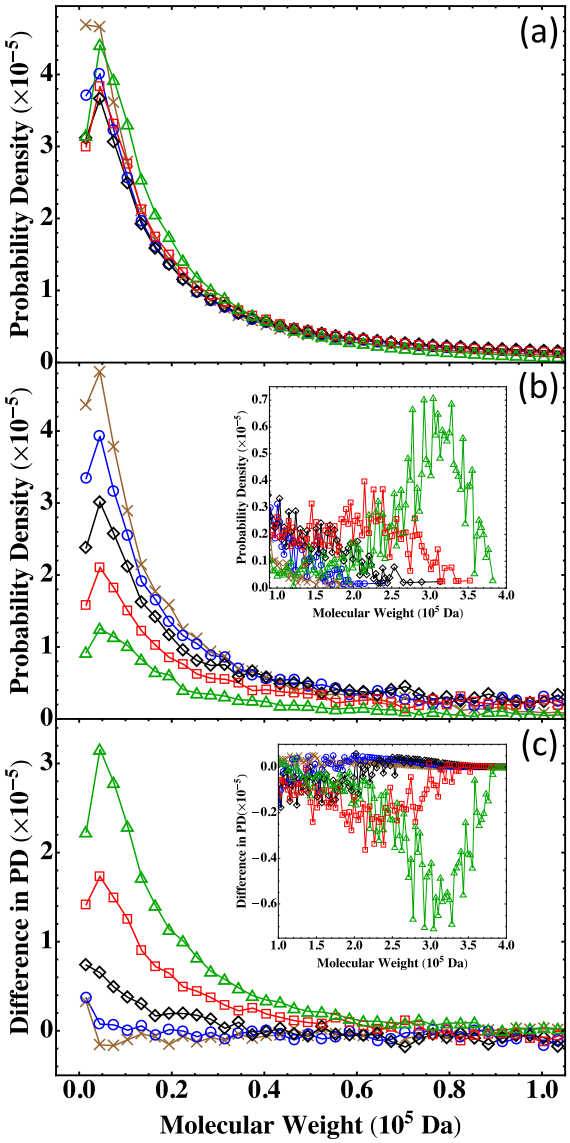

Figure 4.8: Molecular weight distribution (a) from the Flory-Stockmayer theory and (b) from the MC simulations. The inset of (b) shows the MC results in the high molecular weight region. (c) Difference between the results from the Flory-Stockmayer theory and the MC simulations on the probability density (PD). The inset of (c) shows the difference in the high molecular weight region. The data are for $S^{0.95}$ (brown crosses), $S^{0.96}$ (blue circles), $S^{0.97}$ (black diamonds), $S^{0.98}$ (red squares), and $S^{0.99}$ (green triangles). The MC results are averages of 10,000 runs.

results from the theory and simulations agree, as for $S^{0.95}$ and $S^{0.96}$. The difference becomes noticeable when the system approaches the gel point, such as $S^{0.97}$. For $S^{0.98}$ and $S^{0.99}$, they are above the gel point and clear differences in the probability density from the theory and simulations can be noted in both low (see Fig. 4.8(c)) and high (see the inset of Fig. 4.8(c)) molecular weight regions. The results for the partially reacted stoichiometric systems thus corroborate the conclusion that the Flory-Stockmayer theory only applies to systems well below the gel point. However, the MC simulations can be used to compute the molecular weight distribution for any systems no matter they are below, around, or above the gel point.

4.4.5 Nonstoichiometric Systems

We finally discuss nonstoichiometric systems where $\sum_{q=1}^{i} f_q A_q \neq \sum_{h=1}^{j} g_h B_h$ and $p_A \neq p_B$. Three systems with sizes similar to $S_>$ are shown in Table 4.5. We fix the value of p_A at 0.99 but vary p_B from 0.93 to 0.97. The numbers of monomers of PA, BPADA, and TAPE are all fixed. The number of MPD is varied according to Eq. (4.1). Specifically, when the number of MPD is reduced, the values of p_B, ρ_2, and α are all increased. For the three systems in Table 4.5, $S^n_<$ is below, $S^n_\simeq$ is around, and $S^n_>$ is above the gel point. Here the superscript n in the system labels indicates that these systems are nonstoichiometric.

The results on the molecular weight distribution for the three nonstoichiometric systems are plotted in Fig. 4.9. For $S^n_<$ which is below the gel point, the MC results agree with the prediction of the Flory-Stockmayer theory, as shown in Fig. 4.9(a). The two start to differ when a system approaches the gel point. An example is shown in Fig. 4.9(b) for $S^n_\simeq$ with $\alpha = 0.501584$. For this system, the Flory-Stockmayer theory overestimates the probability of low molecular weight polymers while underestimates the probability in the region of molecular weight higher than about 0.5×10^5 Da (see the inset of Fig. 4.9(b)).

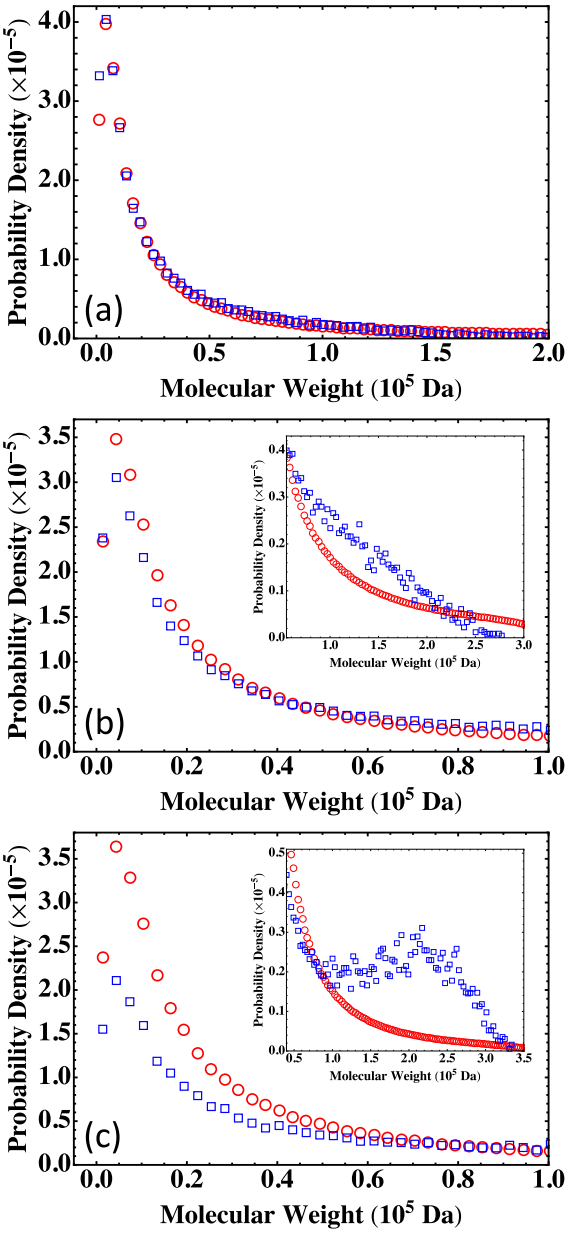

Figure 4.9: Molecular weight distribution for the three systems in Table 4.5: (a) $S^n_<$, (b) $S^n_\simeq$, and (c) $S^n_>$. The results are for the Flory-Stockmayer theory (red circles) and the MC simulations (blue squares). The MC results are averages of 50,000 runs.

For $S_>^n$ which is above the gel point, the MC results on the probability density are smaller than those calculated with the Flory-Stockmayer theory when the molecular weight is lower than about 1×10^5 Da (Fig. 4.9(c)) but higher at higher molecular weights (see the inset of Fig. 4.9(c)). For $S_>^n$ the molecular weight distribution has a second peak around 2×10^5 Da, while the Flory-Stockmayer theory predicts a monotonically decaying distribution in this region. The results on nonstoichiometric systems thus one more time indicate that the Flory-Stockmayer theory only applies to systems well below the gel point, for which the formation of cyclic polymers or closed loops is negligible.

4.5 Conclusions

We have used MC simulations to study the polymerization of branched polyetherimides from BPADA (backbone monomer), MPD (backbone monomer), PA (chain terminator), and TAPE (branching agent). All the reactions for this system can be reduced to a condensation reaction between an amine group and a carboxylic anhydride group and thus can be characterized by one reaction rate. Our work shows that in the MC model, the reaction rate should be computed using the concentrations of the functional groups on the monomers involved in a specific reaction, not the concentrations of the monomers themselves. The MC results are compared to the predictions of the Flory-Stockmayer theory. A practical approach of using the Flory-Stockmayer theory to compute molecular weight distributions has been suggested. We find that both the Flory-Stockmayer theory and the MC simulations accurately predict the molecular weight distribution for systems well below the gel point that is set by the functionality of the branching agent, though ring formation is not considered by the Flory-Stockmayer theory but allowed in MC simulations. The agreement between the theory and simulations thus indicates that ring formation is negligible for systems well below

the gel point. However, for systems close to or above the gel point, the Flory-Stockmayer theory is not applicable as many cyclic polymers can be produced and ring structures can form in highly branched networks. For these systems, the MC simulations can still be used to quickly compute the molecular weight distribution that can be used to describe experimental measurements including average molecular weights.

Our tests indicate that in the MC simulations, a system with only a few hundred to a few thousand monomers but the same molar ratios of participating monomers is large enough to yield converging results on the molecular weight distribution for the region of molecular weight relevant to typical experiments (from 0 to about 3×10^5 Da in the case of polyetherimides). These conclusions have been thoroughly confirmed with simulations for fully reacted, partially reacted, stoichiometric, and nonstoichiometric systems. The MC model is expected to be applicable to a wide range of step-growth polymers.

Chapter 5

Summary and Future Prospects

All-atom molecular dynamics (MD) simulations is a great tool for researchers to study molecules, molecular complexes, and materials. However, there are obvious obstacles due to the bottleneck of computational efficiency. Recently, a research group from Google published a chapter about a quantum computer.[212] They did a simulation which took 200 seconds on their quantum processors but would take about 10,000 years on a traditional classical computer. If computational efficiency really improves by such an order, more than 10^{11}, for molecular modeling, then the obstacles we meet now in the simulation of a material will automatically be solved. However, we do not know when this kind of techniques could be regularly applied to material modeling. We have to find other ways to deal with the limitations of all-atom MD simulations, which usually mean trading some accuracy or atomistic details for speed.

In this thesis, we used three different approaches to study various aspects of polyetherimides. First of all, we tested a new approach to build a coarse-grained model of polyetherimides on the basis of the force-matching principle, which is similar to the conditional reversible work (CRW) method. However, in our approach we modified the entropic correction term to

enable us to deal with planar structures such as aromatic rings, which are coarse-grained into spherical beads, and maintain a reasonable accuracy in the coarse-grained force field. These highly aspherical structures are difficult to deal with using the CRW and force-matching methods.[46] Calculations of the polyetherimide density at various temperatures with our coarse-grained model showed that the model has a good temperature-transferability. The coarse-grained model is also able to reproduce the mechanical moduli of polyetherimides from all-atom MD simulations within a factor of about 2, much improved from the coarse-grained models of other polymers reported in the literature. Furthermore, this factor remains roughly unchanged when the coarse-grained model is applied to a different temperature. Our results thus show that the key for a coarse-grained model to capture the mechanical properties of a polymer at various temperatures simultaneously is to capture its thermal expansion properties, which provides a useful guideline for developing more accurate coarse-grained models of polyetherimides and other polymers as well in the future. For example, recently we have applied the same approach to construct a coarse-grained model of epoxy and the model has accurately reproduced its mechanical moduli as compared to the results from all-atom MD simulations and experiments.

Although we have achieved a certain level of success with the development of the coarse-grained model of polyetherimides, there are still many challenges to address in future work. For example, how can we improve the coarse-grained model to accurately reproduce the mechanical properties of polyetherimides while at the same time, the structural property of the polymer (e.g., the pair correlation functions) is reproduced as well? Another challenge is that although the coarse-grained model allows us to model a larger system (by a factor about 10 to 20) at a longer time scale (by a factor about 5) compared to all-atom MD simulations, it is still not sufficient for us to capture the shear rates in experimental and industrial conditions. The lowest shear rate achievable even with the coarse-grained model is at the

order of 10^7 s^{-1}, which is still several orders of magnitude higher than typical experimental rates. Therefore, we need to either further coarse-grain the system or introduce some novel techniques to reach shear rates that overlap with the experimentally probed range.

Our second effort is to explore various strategies to quickly compute or estimate the glass transition temperature (T_g) of polyimides, of which polyetherimides are a subclass. We showed that a predictive model of T_g can be established by analyzing the data on T_g of polyimides reported in the literature. The input of this model is the chemical structure of a polyimide polymer and other properties of the system such as molecular weights. Then a series of features of the polymer can be generated and quantified and the predictive model can be used to predict T_g of the polymer on the basis of these features. We showed that this model can be trained with available data and then used to estimate T_g for new polyimide polymers only in several minutes. Similar models can also be constructed for other physical properties of polymers. Our tests also showed that for this type of predictive models, if the size of the training dataset is small, then a bagging approach can be used to make the models more stable.

The predictions of T_g from the machine-learning-based model are compared to experimental data not used in the training process as well as results computed in all-atom MD simulations. We showed that in addition to the frequently used density-temperature curve, the diffusion coefficients of small gas molecules can also be quickly computed in all-atom MD simulations and used to determine T_g. These methods are used to calculate T_g for a few polyetherimides that are not yet synthesized. The results are expected to guide future experiments to synthesize and characterize these new polyetherimides that may be potentially useful. When more data become available, the predictive model can be further trained to improve the accuracy of its prediction and the revelation of the most important features of polyimides that are strongly correlated to their glass transition temperatures.

Finally, we built a Monte Carlo (MC) model on the basis of the Gillespie algorithm to study the formation of branched polymers via condensation polymerization. The model takes into account the presence of end cappers that can terminate a branch and multi-functional branching agents and can quickly yield information on the molecular weight distribution of the branched chains. We used this model to study a wide range of systems, including partially reacted, fully reacted, stoichiometric, and nonstoichiometric ones. Below the gel point of a system, the MC results agree well with the predictions of the Flory-Stockmayer theory of branched polymers. However, above the gel point, the theory fails while the MC model still produce sensible molecular weight distributions. A software based on this model has been developed and is currently in usage at SABIC to guide the experimental design of polymerizing systems for the production of polyetherimides with desired molecular weight distributions. In the future, it may be interesting to extend the MC model to describe the polymerization of more complex branching polymers. One example can be a system with mixed branching agents with different functionalities.

Bibliography

[1] M. T. Bogert and R. R. Renshaw, J. Am. Chem. Soc. **30**, 1135 (1908).

[2] R. O. Johnson and H. S. Burlhis, J. Polym. Sci.: Polym. Symp. **70**, 129 (1983).

[3] D. R. Heath and J. Wirth, "Process for making polyetherimides and products derived therefrom," (1971), US Patent App. 3/838,097.

[4] C. H. Sheppard and H. R. Lubowitz, "Polyetherimide oligomer," (1989), US Patent 4,851,495.

[5] L. Colucci-mizenko and A. Kugler, "Polymeric implantable medical devices and surgical instruments," (2018), US Patent App. 15/574,507.

[6] N. Yamaguchi, R. R. Odle, and S. Mishra, "Polyetherimide and polyetherimide sulfone blends having automotive lighting applications," (2013), US Patent 8,545,988.

[7] K. Cao and G. Liu, Macromolecules **50**, 2016 (2017).

[8] S.-H. Hsiao, H.-M. Wang, P.-C. Chang, Y.-R. Kung, and T.-M. Lee, Journal of Polymer Science Part A: Polymer Chemistry **51**, 2925 (2013).

[9] B. V. Tawade, A. D. Kulkarni, and P. P. Wadgaonkar, Polymer International **64**, 1770 (2015).

BIBLIOGRAPHY

[10] J. J. de Pablo, N. E. Jackson, M. A. Webb, L.-Q. Chen, J. E. Moore, D. Morgan, R. Jacobs, T. Pollock, D. G. Schlom, E. S. Toberer, J. Analytis, I. Dabo, D. M. DeLongchamp, G. A. Fiete, G. M. Grason, G. Hautier, Y. Mo, K. Rajan, E. J. Reed, E. Rodriguez, V. Stevanovic, J. Suntivich, K. Thornton, and J.-C. Zhao, NPJ Comput. Mater. **5**, 41 (2019).

[11] K. Kremer and G. S. Grest, Journal of Physics: Condensed Matter **2**, SA295 (1990).

[12] M. Matsui, S. C. Parker, and M. Leslie, American Mineralogist **85**, 312 (2000).

[13] Q. Hu, S. Viswanadham, R. Joshi, K. H. Schoenbach, S. J. Beebe, and P. Blackmore, Physical Review E **71**, 031914 (2005).

[14] A. Hospital, J. R. Goñi, M. Orozco, and J. L. Gelpí, Advances and Applications in Bioinformatics and Chemistry: AABC **8**, 37 (2015).

[15] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **27**, 1208 (1957).

[16] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **31**, 459 (1959).

[17] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **33**, 1439 (1960).

[18] B. Alder, J. Chem. Phys. **40**, 2724 (1964).

[19] W. G. Hoover and B. J. Alder, J. Chem. Phys. **46**, 686 (1967).

[20] B. Alder, W. Hoover, and D. Young, J. Chem. Phys. **49**, 3688 (1968).

[21] T. Einwohner and B. Alder, J. Chem. Phys. **49**, 1458 (1968).

[22] B. Alder and C. Hecht, J. Chem. Phys. **50**, 2032 (1969).

[23] B. Alder, D. Gass, and T. Wainwright, J. Chem. Phys. **53**, 3813 (1970).

[24] M. Levitt, Journal of Molecular Biology **170**, 723 (1983).

[25] M. Levitt, Journal of Molecular Biology **168**, 595 (1983).

[26] C. Brooks III, A. Brünger, and M. Karplus, Biopolymers: Original Research on Biomolecules **24**, 843 (1985).

[27] A. T. Brünger, G. M. CLoRE, A. M. Gronenborn, and M. Karplus, Proceedings of the National Academy of Sciences **83**, 3801 (1986).

[28] R. Elber and M. Karplus, Science **235**, 318 (1987).

[29] J. Gao, K. Kuczera, B. Tidor, and M. Karplus, Science **244**, 1069 (1989).

[30] G. S. Grest and K. Kremer, Phys. Rev. A **33**, 3628 (1986).

[31] J. Clarke and D. Brown, Mol. Simul. **3**, 27 (1989).

[32] D. Rigby and R.-J. Roe, J. Chem. Phys. **87**, 7285 (1987).

[33] H. Takeuchi, R.-J. Roe, and J. E. Mark, J. Chem. Phys. **93**, 9042 (1990).

[34] P. Khalatur, Y. G. Papulov, and A. Pavlov, Molecular Physics **58**, 887 (1986).

[35] H. Sun, Macromolecules **28**, 701 (1995).

[36] H. Sun, S. J. Mumby, J. R. Maple, and A. T. Hagler, Journal of the American Chemical Society **116**, 2978 (1994).

[37] S. L. Mayo, B. D. Olafson, and W. A. Goddard, Journal of Physical Chemistry **94**, 8897 (1990).

[38] M. O. Steinhauser, *Computational multiscale modeling of fluids and solids* (Springer, 2017).

BIBLIOGRAPHY

[39] P. D. Gujrati and A. I. Leonov, *Modeling and simulation in polymers* (John Wiley & Sons, 2010).

[40] M. Levitt and A. Warshel, Nature **253**, 694 (1975).

[41] K. Kremer and G. S. Grest, J. Chem. Phys. **92**, 5057 (1990).

[42] T. E. Ouldridge, A. A. Louis, and J. P. Doye, J. Chem. Phys. **134**, 02B627 (2011).

[43] A. Soper, Chem. Phys. **202**, 295 (1996).

[44] S. Izvekov and G. A. Voth, J. Chem. Phys. **123**, 134105 (2005).

[45] W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, J. Chem. Phys. **128**, 244114 (2008).

[46] Y. Wang, W. Noid, P. Liu, and G. A. Voth, Phys. Chem. Chem. Phys. **11**, 2002 (2009).

[47] E. Brini, V. Marcon, and N. F. van der Vegt, Phys. Chem. Chem. Phys. **13**, 10468 (2011).

[48] M. Hanke, Journal of Statistical Physics **170**, 536 (2018).

[49] A. Villa, C. Peter, and N. F. van der Vegt, J. Chem. Theory Comp. **6**, 2434 (2010).

[50] W. Noid, J. Chem. Phys. **139**, 09B201_1 (2013).

[51] B. Hess, H. Bekker, H. J. Berendsen, and J. G. Fraaije, J. Comput. Chem. **18**, 1463 (1997).

[52] B. Hess, J. Chem. Theory Comp. **4**, 116 (2008).

[53] B. Hess, C. Holm, and N. van der Vegt, Phys. Rev. Lett. **96**, 147801 (2006).

[54] D. Fritz, V. A. Harmandaris, K. Kremer, and N. F. van der Vegt, Macromolecules **42**, 7579 (2009).

[55] G. Deichmann and N. F. van der Vegt, J. Chem. Theory Comp. **13**, 6158 (2017).

[56] J. Han, R. H. Gee, and R. H. Boyd, Macromolecules **27**, 7781 (1994).

[57] B. F. Abu-Sharkh, Computational and Theoretical Polymer Science **11**, 29 (2001).

[58] H. Morita, K. Tanaka, T. Kajiyama, T. Nishi, and M. Doi, Macromolecules **39**, 6233 (2006), https://doi.org/10.1021/ma052632h .

[59] J. Buchholz, W. Paul, F. Varnik, and K. Binder, J. Chem. Phys. **117**, 7364 (2002).

[60] C. F. Fan, T. Çagin, W. Shi, and K. A. Smith, Macromol. Theo. Simul. **6**, 83 (1997).

[61] A. Deazle, I. Hamerton, C. Heald, and B. Howlin, Polymer International **41**, 151 (1996).

[62] I. Hamerton, B. J. Howlin, P. Klewpatinond, H. J. Shortley, and S. Takeda, Polymer **47**, 690 (2006).

[63] J. Pozuelo and J. Baselga, Polymer **43**, 6049 (2002).

[64] N. Hu, R. Chen, and A. Hsu, Polymer International **55**, 872 (2006).

[65] Y. Wang, W. Wang, Z. Zhang, L. Xu, and P. Li, European Polymer Journal **75**, 36 (2016).

[66] J.-L. Barrat, J. Baschnagel, and A. Lyulin, Soft Matter **6**, 3430 (2010).

[67] D. G. Luchinsky, H. Hafiychuk, V. Hafiychuk, and K. R. Wheeler, NASA Technical Report , NASA/TM (2018).

BIBLIOGRAPHY

[68] S. V. Lyulin, S. V. Larin, A. A. Gurtovenko, V. M. Nazarychev, S. G. Falkovich, V. E. Yudin, V. M. Svetlichnyi, I. V. Gofman, and A. V. Lyulin, Soft Matter **10**, 1224 (2014).

[69] B. G. Sumpter and D. W. Noid, Macromol. Theo. Simul. **3**, 363 (1994).

[70] S. J. Joyce, D. J. Osguthorpe, J. A. Padgett, and G. J. Price, Journal of the Chemical Society, Faraday Transactions **91**, 2491 (1995).

[71] X. Yu, X. Wang, X. Li, J. Gao, and H. Wang, Macromol. Theo. Simul. **15**, 94 (2006).

[72] X. Chen, L. Sztandera, and H. M. Cartwright, International Journal of Intelligent Systems **23**, 22 (2008).

[73] L. Ning, Journal of Materials Science **44**, 3156 (2009).

[74] J. Xu, L. Zhu, D. Fang, L. Liu, W. Xu, and Z. Li, Fibers and Polymers **13**, 352 (2012).

[75] J.-F. Pei, C.-Z. Cai, Y.-M. Zhu, and B. Yan, Macromol. Theo. Simul. **22**, 52 (2013).

[76] M. Chen, F. Jabeen, B. Rasulev, M. Ossowski, and P. Boudjouk, Journal of Polymer Science Part B: Polymer Physics **56**, 877 (2018).

[77] K. P. Eurenius, D. C. Chatfield, B. R. Brooks, and M. Hodoscek, Int. J. Quantum Chem. **60**, 1189 (1996).

[78] D. C. Chatfield, K. P. Eurenius, and B. R. Brooks, Journal of Molecular Structure: THEOCHEM **423**, 79 (1998).

[79] H. M. Senn and W. Thiel, Current Opinion In Chemical Biology **11**, 182 (2007).

[80] M. Shoji, H. Isobe, and K. Yamaguchi, Chemical Physics Letters **636**, 172 (2015).

[81] A. C. Van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, J. Phys. Chem. A **105**, 9396 (2001).

[82] J. Tersoff, Physical Review B **37**, 6991 (1988).

[83] M. Zheng, X. Li, and L. Guo, Journal of Molecular Graphics and Modelling **41**, 1 (2013).

[84] M. Szwarc, M. Levy, and R. Milkovich, J. Am. Chem. Soc. **78**, 2656 (1956).

[85] J. He, H. Zhang, J. Chen, and Y. Yang, Macromolecules **30**, 8010 (1997).

[86] M. Al-Harthi, J. B. Soares, and L. C. Simon, Macromol. Mater. Engr. **291**, 993 (2006).

[87] J. He, L. Li, and Y. Yang, Macromol. Theo. Simul. **9**, 463 (2000).

[88] Y. Luo and B. Yu, Polymer - Plastics Technology and Engineering **43**, 1299 (2005).

[89] H. Tobita and N. Hamashima, Macromol. Theo. Simul. **9**, 453 (2000).

[90] H. Tobita, Journal of Polymer Science Part B: Polymer Physics **31**, 1363 (1993).

[91] A. K. Tripathi and D. C. Sundberg, Macromol. Theo. Simul. **24**, 52 (2015).

[92] S. Hamzehlou, N. Ballard, P. Carretero, M. Paulis, J. M. Asua, Y. Reyes, and J. R. Leiza, Polymer **55**, 4801 (2014).

[93] A. Johnson and K. O'Driscoll, European Polymer Journal **20**, 979 (1984).

[94] L. T. Hillegers, M. Kapnistos, A. Nijenhuis, J. J. Slot, and P. A. Steeman, Macromol. Theo. Simul. **20**, 219 (2011).

[95] W. Wright, Materials Design **12**, 222 (1991).

BIBLIOGRAPHY

[96] M. F. Rigana, P. Thirukumaran, K. Shanthi, and M. Sarojadevi, RSC Advances **6**, 33249 (2016).

[97] I. Kaya and M. Kamacı, Journal of Applied Polymer Science **135**, 46573 (2018).

[98] V. Nazarychev, S. Larin, A. Lyulin, T. Dingemans, J. Kenny, and S. Lyulin, Polymers **9**, 548 (2017).

[99] S. Lyulin, A. Gurtovenko, S. Larin, V. Nazarychev, and A. Lyulin, Macromolecules **46**, 6357 (2013).

[100] A. Shokuhfar and B. Arab, J. Mol. Modeling **19**, 3719 (2013).

[101] J. Ennari, L.-O. Pietilä, V. Virkkunen, and F. Sundholm, Polymer **43**, 5427 (2002).

[102] Y. Jin and F. Yuan, Composites Sci. Techno. **63**, 1507 (2003).

[103] X. Chen, C. Yuan, C. K. Wong, and G. Zhang, Journal of Molecular Modeling **18**, 2333 (2012).

[104] P. Bhattacharya, N. K. Geitner, S. Sarupria, and P. C. Ke, Phys. Chem. Chem. Phys. **15**, 4477 (2013).

[105] D. Reith, M. Pütz, and F. Müller-Plathe, Journal of Computational Chemistry **24**, 1624 (2003).

[106] T. D. Potter, J. Tasche, E. L. Barrett, M. Walker, and M. R. Wilson, Liquid Crystals **44**, 1979 (2017).

[107] J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, P. B. Moore, and M. L. Klein, The Journal of Physical Chemistry B **105**, 9785 (2001).

[108] W. Shinoda, R. DeVane, and M. L. Klein, Molecular Simulation **33**, 27 (2007).

[109] F. Ercolessi and J. B. Adams, Europhy. Lett. **26**, 583 (1994).

[110] D. Fritz, V. A. Harmandaris, K. Kremer, and N. F. A. van der Vegt, Macromolecules **42**, 7579 (2009).

[111] Scienomics, (2004-2012).

[112] T. A. Halgren, Journal of Computational Chemistry **17**, 520 (1996).

[113] Q. Xiao and H. Guo, Physical Chemistry Chemical Physics **18**, 29808 (2016).

[114] M. O. Sinnokrot, E. F. Valeev, and C. D. Sherrill, J. Am. Chem. Soc. **124**, 10887 (2002).

[115] D. Wilson, H. D. Stenzenberger, P. M. Hergenrother, F. W. Harris, T. Takekoshi, P. R. Young, R. Escott, H. Satou, H. Suzuki, D. Makino, and C. E. Sroog, *Polyimides* (Springer, 1990).

[116] K. Mittal and S. of Plastics Engineers. Mid-Hudson Section, *Polyimides: Synthesis, Characterization, and Applications*, (Technical Conference on Polyimides) No. v. 1 (Springer US, 1984).

[117] M. Mohammadi, J. Davoodi, *et al.*, European Polymer Journal **91**, 121 (2017).

[118] Y. Sun, L. Chen, L. Cui, Y. Zhang, and X. Du, Comput. Mater. Sci. **143**, 240 (2018).

[119] L. Alzate-Vargas, M. E. Fortunato, B. Haley, C. Li, C. M. Colina, and A. Strachan, Modelling and Simulation in Materials Science and Engineering **26**, 065007 (2018).

[120] S. G. Falkovich, S. V. Lyulin, V. M. Nazarychev, S. V. Larin, A. A. Gurtovenko, N. V. Lukasheva, and A. V. Lyulin, Journal of Polymer Science Part B: Polymer Physics **52**, 640 (2014).

BIBLIOGRAPHY

[121] J. Xia, S. Liu, P. K. Pallathadka, M. L. Chng, and T.-S. Chung, Industrial & Engineering Chemistry Research **49**, 12014 (2010).

[122] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).

[123] H. Sun, S. J. Mumby, J. R. Maple, and A. T. Hagler, J. Am. Chem. Soc. **116**, 2978 (1994).

[124] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, "Gaussian 09, revision e.01," (2013), gaussian Inc. Wallingford CT.

[125] D. Bashford, in *Thermoplastics* (Springer, 1997) pp. 470–473.

[126] Q. Yang, X. Chen, Z. He, F. Lan, and H. Liu, RSC Adv. **6**, 12053 (2016).

[127] M. Li, X. Y. Liu, J. Q. Qin, and Y. Gu, Express Polym. Lett. **3**, 665 (2009).

[128] S. V. Lyulin, S. V. Larin, A. A. Gurtovenko, V. M. Nazarychev, S. G. Falkovich, V. E.

Yudin, V. M. Svetlichnyi, I. V. Gofman, and A. V. Lyulin, Soft Matter **10**, 1224 (2014).

[129] M. Minelli, M. G. De Angelis, and D. Hofmann, Fluid Phase Equilibria **333**, 87 (2012).

[130] H. Yoshimizu, S. Ohta, T. Asano, T. Suzuki, and Y. Tsujita, Polym. J **44**, 821 (2012).

[131] K.-q. Yu, Z.-s. Li, and J. Sun, Macromol. Theory Simul. **10**, 624 (2001).

[132] P. Meares, Transactions of the Faraday Society **53**, 101 (1957).

[133] C. Kumins and J. Roteman, Journal of Polymer Science **55**, 683 (1961).

[134] M. Menzinger and R. Wolfgang, Angewandte Chemie International Edition in English **8**, 438 (1969).

[135] D. Heyes, Chemical Physics **82**, 285 (1983).

[136] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, J. Chem. Phys. **103**, 4613 (1995).

[137] X. Fang, Z. Yang, S. Zhang, L. Gao, and M. Ding, Macromolecules **35**, 8708 (2002).

[138] S.-H. Hsiao, G.-S. Liou, and S.-H. Chen, Journal of Polymer Science Part A: Polymer Chemistry **36**, 1657 (1998).

[139] T. Takahashi, S. Takabayashi, and H. Inoue, High Performance Polymers **10**, 33 (1998).

[140] Q. Li, X. Fang, Z. Wang, L. Gao, and M. Ding, Journal of Polymer Science Part A: Polymer Chemistry **41**, 3249 (2003).

[141] M. Zhang, Z. Wang, L. Gao, and M. Ding, Journal of Polymer Science Part A: Polymer Chemistry **44**, 959 (2006).

[142] M. Ding, Progress in Polymer Science **32**, 623 (2007).

[143] W. Liu, Polymer Engineering & Science **50**, 1547 (2010).

[144] "Molecular descriptor and fingerprint generator," https://www.alvascience.com/alvadesc/, accessed: 12-04-2019.

[145] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, npj Comput. Mater. **3**, 54 (2017).

[146] D. J. Audus and J. J. de Pablo, ACS Macro Letters **6**, 1078 (2017).

[147] J. S. Peerless, N. J. B. Milliken, T. J. Oweida, M. D. Manning, and Y. G. Yingling, Adv. Theo. Simul. **2**, 1800129 (2019), https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.201800129 .

[148] R. Tibshirani, Journal of the Royal Statistical Society: Series B (Methodological) **58**, 267 (1996).

[149] L. Breiman, Machine Learning **24**, 123 (1996).

[150] M. Rubinstein and R. H. Colby, *Polymer Physics* (Oxford University Press, New York, NY, 2003).

[151] R. W. Nunes, J. R. Martin, and J. F. Johnson, Polym. Eng. Sci. **22**, 205 (1982).

[152] Suneel, D. M. A. Buzza, D. J. Groves, T. C. B. McLeish, D. Parker, A. J. Keeney, and W. J. Feast, Macromolecules **35**, 9605 (2002).

[153] D. W. Mead, J. Rheol. **38**, 1797 (1994).

[154] N. H. Williamson, M. Nydén, and M. Röding, J. Magn. Reson. **267**, 54 (2016).

[155] D. Nichetti and I. Manas-Zloczower, J. Rheol. **42**, 951 (1998).

[156] P. J. Flory, J. Am. Chem. Soc. **63**, 3083 (1941).

[157] P. J. Flory, J. Am. Chem. Soc. **63**, 3091 (1941).

[158] P. J. Flory, J. Am. Chem. Soc. **63**, 3096 (1941).

[159] W. H. Stockmayer, J. Chem. Phys. **11**, 45 (1943).

[160] W. H. Stockmayer, J. Chem. Phys. **12**, 125 (1944).

[161] W. H. Stockmayer, J. Polym. Sci. **IX**, 69 (1952).

[162] L. H. Peebles, *Molecular Weight Distributions in Polymers* (Interscience Publishers, 1971).

[163] A. Matsumoto, "Free-radical crosslinking polymerization and copolymerization of multivinyl compounds," in *Synthesis and Photosynthesis* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1995) pp. 41–80.

[164] A. Matsumoto, Y. Kitaguchi, and O. Sonoda, Macromolecules **32**, 8336 (1999).

[165] I. Bannister, N. C. Billingham, S. P. Armes, S. P. Rannard, and P. Findlay, Macromolecules **39**, 7483 (2006).

[166] H. Gao, K. Min, and K. Matyjaszewski, Macromolecules **40**, 7763 (2007).

[167] K. M. Schultz, A. D. Baldwin, K. L. Kiick, and E. M. Furst, Macromolecules **42**, 5310 (2009).

[168] J. Rosselgong, S. P. Armes, W. Barton, and D. Price, Macromolecules **42**, 5919 (2009).

[169] J. Lyu, Y. Gao, Z. Zhang, U. Greiser, H. Tai, and W. Wang, Sci. China **61**, 319 (2018).

BIBLIOGRAPHY

[170] J. Lyu, Y. Gao, Z. Zhang, U. Greiser, P. Polanowski, J. K. Jeszka, K. Matyjaszewski, H. Tai, and W. Wang, Macromolecules **51**, 6673 (2018).

[171] D. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, New York, NY, USA, 2005).

[172] H.-P. Hsu and P. Grassberger, J. Stat. Phys. **144**, 597 (2011).

[173] A. L. T. Brandão, J. B. P. Soares, J. C. Pinto, and A. L. Alberton, Macromol. React. Eng. **9**, 141 (2015).

[174] A. F. Johnson and K. F. O'Driscoll, European Polymer Journal **20**, 979 (1984).

[175] H. Tobita, Macromolecules **26**, 5427 (1993).

[176] H. Tobita, Acta Polym. **46**, 185 (1995).

[177] H. Tobita, Polymer **36**, 2585 (1995).

[178] H. Tobita, J. Polym. Sci. B: Polym. Phys. **39**, 391 (2001).

[179] H. Tobita, Macromol. Theo. Simul. **15**, 23 (2006).

[180] H. Tobita and F. Yanase, Macromol. Theo. Simul. **16**, 476 (2007).

[181] E. Hädicke and H. Stutz, J. Appl. Polym. Sci. **85**, 929 (2002).

[182] X. He, H. Liang, and C. Pan, Macromo. Theo. Simul. **10**, 196 (2001).

[183] X. He, H. Liang, and C. Pan, Polymer **44**, 6697 (2003).

[184] Y. Rouault and A. Milchev, Phys. Rev. E **55**, 2020 (1997).

[185] J. He, H. Zhang, J. Chen, and Y. Yang, Macromolecules **30**, 8010 (1997).

[186] S. W. Prescott, Macromolecules **36**, 9608 (2003).

[187] M. Al-Harthi, J. B. Soares, and L. C. Simon, Macromol. Mater. Eng. **291**, 993 (2006).

[188] M. Al-Harthi, J. B. P. Soares, and L. C. Simon, Macromol. React. Eng. **1**, 95 (2007).

[189] M. A. Al-Harthi, J. K. Masihullah, S. H. Abbasi, and J. B. P. Soares, Macromol. Theo. Simul. **18**, 307 (2009).

[190] M. Al-Harthi, M. J. Khan, S. H. Abbasi, and J. B. P. Soares, Macromol. React. Eng. **3**, 148 (2009).

[191] P. Polanowski, J. K. Jeszka, and K. Matyjaszewski, Polymer **51**, 6084 (2010).

[192] P. Polanowski, J. K. Jeszka, W. Li, and K. Matyjaszewski, Polymer **52**, 5092 (2011).

[193] I. Bannister, N. C. Billingham, and S. P. Armes, Soft Matter **5**, 3495 (2009).

[194] H. Gao, L. H. Oakley, I. A. Konstantinov, S. G. Arturo, and L. J. Broadbelt, Ind. Eng. Chem. Res. **54**, 11975 (2015).

[195] H. Gao, L. J. Broadbelt, I. A. Konstantinov, and S. G. Arturo, AIChE J. **63**, 4013 (2017).

[196] D. Meimaroglou, A. Krallis, V. Saliakas, and C. Kiparissides, Macromolecules **40**, 2224 (2007).

[197] D. Meimaroglou, P. Pladis, A. Baltsas, and C. Kiparissides, Chem. Eng. Sci. **66**, 1685 (2011).

[198] D. Meimaroglou, P. Pladis, and C. Kiparissides, Macromol. React. Eng. **11**, 1600039 (2017).

[199] P. Iedema, K. Remerie, M. van der Ham, and E. Biemond, Polymer **54**, 4093 (2013).

BIBLIOGRAPHY

[200] N. Yaghini and P. D. Iedema, Chem. Eng. Sci. **130**, 310 (2015).

[201] A. Milchev, J. P. Wittmer, and D. P. Landau, Phys. Rev. E **61**, 2959 (2000).

[202] Z. Gao and J. He, Macromol. React. Eng. **9**, 431 (2015).

[203] I. M. Maafa, J. B. P. Soares, and A. Elkamel, Macromol. React. Eng. **1**, 364 (2007).

[204] J. B. P. Soares and A. E. Hamielec, Macromol. React. Eng. **1**, 53 (2007).

[205] S. Hamzehlou, Y. Reyes, and J. R. Leiza, Macromolecules **46**, 9064 (2013).

[206] A. H. Fawcett, R. A. W. Mee, and F. V. McBride, Macromolecules **28**, 1481 (1995).

[207] A. K. Tripathi, J. G. Tsavalas, and D. C. Sundberg, Macromolecules **48**, 184 (2015).

[208] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).

[209] D. T. Gillespie, Annu. Rev. Phys. Chem. **58**, 35 (2007).

[210] R. R. Odle, G. Liu, K. Cao, T. E. Long, and J. M. Dennis, "Telechelic poly (imide) oligomers, methods of manufacture, and uses thereof," (2019), US Patent App. 16/347,105.

[211] R. R. Odle, K. Cao, G. Liu, T. E. Long, J. M. Dennis, and R. J. Mondschein, "Polyimide and polyetherimide from metal containing oligomers," (2019), US Patent App. 16/236,717.

[212] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, C. Giustina, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. Klimov, and S. Knysh, Nature **574**, 505 (2019).

Appendices

Appendix A

Equivalence between "Recentering" and Constraint-Force Schemes of Fixing a Center of Mass

Here we prove the equivalence between the "recentering" and constraint-force schemes of fixing the center of mass of a group of atoms. We first examine the movement of each atom in the "recentering" approach. We use $\vec{f_i}$ to denote the total force on the i-th atom in a group from its interactions with all other atoms in the system. Then after an infinitesimal time dt, the velocity of this atom becomes

$$\vec{v}_i(t+dt) = \vec{v}_i + \frac{\vec{f}_i}{m_i}dt \ , \tag{A.1}$$

and the displacement of this atom is

$$d\vec{r}_i = \vec{v}_i dt + \frac{1}{2}\frac{\vec{f}_i}{m_i}(dt)^2 \ , \tag{A.2}$$

where $\vec{v}_i$ is the velocity of the i-th atom at time t, and m_i is its mass. To fix the center of mass, we require $\vec{v}_i$ to satisfy $\sum_i m_i \vec{v}_i = 0$, where the summation is over all atoms in the group under consideration. The displacement of the group's center of mass before "recentering" is therefore

$$\begin{aligned} d\vec{R} &\equiv \frac{\sum_i m_i d\vec{r}_i}{\sum_i m_i} \\ &= \frac{\sum_i m_i \vec{v}_i + \frac{1}{2} dt \sum_i \vec{f}_i}{\sum_i m_i} dt \\ &= \frac{1}{2} \frac{\sum_i \vec{f}_i}{\sum_i m_i} (dt)^2. \end{aligned} \tag{A.3}$$

During "recentering", each atom in the group is displaced by $-d\vec{R}$ to move the group's center of mass back to its starting location and the renormalized displacement of the i-th atom becomes

$$d\vec{r}_{i,R} = d\vec{r}_i - d\vec{R} = \vec{v}_i dt + \frac{1}{2} \frac{\vec{f}_i}{m_i} (dt)^2 - \frac{1}{2} \frac{\sum_i \vec{f}_i}{\sum_i m_i} (dt)^2 \ . \tag{A.4}$$

It is easy to prove that $\sum_i (m_i \times d\vec{r}_{i,R}) = 0$, indicating that the center of mass is fixed. The velocity of the center of mass, $\vec{v}_c$, before "recentering" is

$$\vec{v}_c = \frac{\sum_i \vec{f}_i}{\sum_i m_i} dt \ . \tag{A.5}$$

When this velocity is subtracted from the velocity of each atom in the group, the renormalized velocity of the i-th atom becomes

$$\vec{v}_{i,R}(t + dt) = \vec{v}_i(t + dt) - \vec{v}_c = \vec{v}_i + \frac{\vec{f}_i}{m_i} dt - \frac{\sum_i \vec{f}_i}{\sum_i m_i} dt. \tag{A.6}$$

The velocity of the group's center of mass after "recentering" is reduced to zero, i.e., $\sum_i [m_i \times \vec{v}_{i,R}(t+dt)] = 0$. As expected, "recentering" renders the center of mass of the group to be fixed.

Next, we consider the constraint-force approach by applying an extra constraining force, $\vec{f}_{i,C}$, to the i-th atom in the group. The velocity of the i-th atom after an infinitesimal time dt is

$$\vec{v}_{i,C}(t+dt) = \vec{v}_i + \frac{\vec{f}_i + \vec{f}_{i,C}}{m_i} dt \ , \tag{A.7}$$

and its corresponding displacement

$$d\vec{r}_{i,C} = \vec{v}_i dt + \frac{1}{2} \frac{\vec{f}_i + \vec{f}_{i,C}}{m_i} (dt)^2 \ . \tag{A.8}$$

It is easy to show that if

$$\vec{f}_{i,C} \equiv -m_i \frac{\sum_j \vec{f}_j}{\sum_j m_j} \ , \tag{A.9}$$

then

$$d\vec{r}_{i,C} = d\vec{r}_{i,R} \quad \text{and} \quad \vec{v}_{i,C}(t+dt) = \vec{v}_{i,R}(t+dt) \ . \tag{A.10}$$

This proves that the "recentering" and constraint-force schemes are equivalent. Furthermore, it can be noted that

$$\sum_i \vec{f}_{i,C} = -\sum_i \vec{f}_i \ , \tag{A.11}$$

and

$$\sum_i \vec{r}_i \times \vec{f}_{i,C} = 0 \ , \tag{A.12}$$

where $\vec{r}_i$ is the position vector of the i-th atom relative to the group's center of mass. Eqs. (A.11) and (A.12) are two natural requirements of the constraint-force scheme of fixing a center of mass. That is, the total force from constraints should balance the total force exerted on all atoms in the group by other atoms in the system, which makes the acceleration of the center of mass to be zero. When the initial velocity of the center of mass is zero, it is naturally fixed. Furthermore, the constraint forces should have zero torque on the group to which they are applied. The rotation of the group around its center of mass is purely determined by the interactions with atoms in other groups.

Appendix B

Coarse-Grained Force Field of Polyetherimides

Here we include the final coarse-grained force field of polyetherimides that can be read by LAMMPS.

MORSE_bpada1		Potential	Force
N	165		
1	0.2	450567.5987	173536.5087
2	2.6	34079.97783	173536.5087
3	2.7	20474.05998	98581.84842
4	2.8	12673.5501	57428.34916
5	3	4821.107251	21096.07932
6	3.1	3105.991677	13206.23217
7	3.2	2027.620131	8361.198749
8	3.3	1340.826846	5374.666954
9	3.4	898.431758	3473.234796
10	3.5	608.394438	2327.511604
11	3.6	414.199362	1556.389917
12	3.7	284.26692	1042.258935
13	3.8	196.695774	709.163989
14	3.9	136.711158	490.528319
15	4	95.231113	339.072577
16	4.1	66.343256	238.684579
17	4.2	45.925492	169.670684
18	4.3	31.536818	118.10281
19	4.4	21.458717	83.459208
20	4.5	14.365118	58.412762
21	4.6	9.359685	41.695911
22	4.7	5.799103	29.515714

23	4.8	3.256502	21.336324
24	4.9	1.447155	14.850611
25	5	0.150903	11.074421
26	5.1	-0.794454	7.832718
27	5.2	-1.474566	5.769523
28	5.3	-1.972275	4.184662
29	5.4	-2.333487	3.039574
30	5.5	-2.594561	2.181904
31	5.6	-2.77816	1.490087
32	5.7	-2.901411	0.974923
33	5.8	-2.981689	0.630645
34	5.9	-3.030474	0.345054
35	6	-3.051992	0.085308
36	6.1	-3.051123	-0.102696
37	6.2	-3.033512	-0.249531
38	6.3	-3.004751	-0.325682
39	6.4	-2.967916	-0.411019
40	6.5	-2.922365	-0.5
41	6.6	-2.867514	-0.597015
42	6.7	-2.80193	-0.714667
43	6.8	-2.723553	-0.852884
44	6.9	-2.629989	-1.018395
45	7	-2.519995	-1.181482
46	7.1	-2.396118	-1.296063
47	7.2	-2.26338	-1.358693

48	7.3	-2.126489	-1.379114
49	7.4	-1.990805	-1.334577
50	7.5	-1.859046	-1.300598
51	7.6	-1.731984	-1.240649
52	7.7	-1.609689	-1.205239
53	7.8	-1.490972	-1.169118
54	7.9	-1.376285	-1.124621
55	8	-1.267333	-1.054414
56	8.1	-1.164972	-0.992801
57	8.2	-1.069617	-0.914309
58	8.3	-0.981632	-0.845384
59	8.4	-0.900826	-0.770739
60	8.5	-0.82724	-0.700971
61	8.6	-0.759718	-0.649467
62	8.7	-0.69766	-0.591693
63	8.8	-0.640556	-0.550389
64	8.9	-0.587775	-0.505229
65	9	-0.53925	-0.465279
66	9.1	-0.494782	-0.424082
67	9.2	-0.454116	-0.389228
68	9.3	-0.416982	-0.35346
69	9.4	-0.383242	-0.321348
70	9.5	-0.352648	-0.290518
71	9.6	-0.324956	-0.263322
72	9.7	-0.299839	-0.239018

73	9.8	-0.277016	-0.217454
74	9.9	-0.256256	-0.197746
75	10	-0.237345	-0.180476
76	10.1	-0.220079	-0.164841
77	10.2	-0.204284	-0.151047
78	10.3	-0.189802	-0.138606
79	10.4	-0.176536	-0.126716
80	10.5	-0.164374	-0.116514
81	10.6	-0.153171	-0.107554
82	10.7	-0.14285	-0.098854
83	10.8	-0.133346	-0.091235
84	10.9	-0.124578	-0.084112
85	11	-0.116488	-0.0777
86	11.1	-0.109004	-0.071971
87	11.2	-0.102076	-0.066587
88	11.3	-0.095648	-0.061976
89	11.4	-0.089678	-0.057433
90	11.5	-0.084129	-0.053539
91	11.6	-0.078963	-0.04978
92	11.7	-0.07416	-0.046287
93	11.8	-0.069688	-0.043146
94	11.9	-0.065514	-0.040343
95	12	-0.061618	-0.03757
96	12.1	-0.057987	-0.035058
97	12.2	-0.054594	-0.032788

98	12.3	-0.051422	-0.030655
99	12.4	-0.048455	-0.028688
100	12.5	-0.045677	-0.026868
101	12.6	-0.043074	-0.025194
102	12.7	-0.040634	-0.023619
103	12.8	-0.038347	-0.022116
104	12.9	-0.0362	-0.020826
105	13	-0.034179	-0.019599
106	13.1	-0.032278	-0.018408
107	13.2	-0.030492	-0.017321
108	13.3	-0.02881	-0.016321
109	13.4	-0.027225	-0.015379
110	13.5	-0.025731	-0.014495
111	13.6	-0.024322	-0.013683
112	13.7	-0.022994	-0.012871
113	13.8	-0.021742	-0.012169
114	13.9	-0.020559	-0.0115
115	14	-0.01944	-0.010886
116	14.1	-0.018383	-0.010256
117	14.2	-0.017384	-0.00972
118	14.3	-0.016437	-0.009212
119	14.4	-0.015541	-0.008709
120	14.5	-0.014694	-0.008242
121	14.6	-0.013891	-0.007814
122	14.7	-0.013129	-0.007414

123	14.8	-0.012408	-0.007018
124	14.9	-0.011724	-0.006669
125	15	-0.011073	-0.006334
126	15.1	-0.010456	-0.006023
127	15.2	-0.009868	-0.005736
128	15.3	-0.009307	-0.005472
129	15.4	-0.008772	-0.005225
130	15.5	-0.008262	-0.004988
131	15.6	-0.007774	-0.004772
132	15.7	-0.007307	-0.004562
133	15.8	-0.006861	-0.004362
134	15.9	-0.006434	-0.004179
135	16	-0.006026	-0.003984
136	16.1	-0.005636	-0.003803
137	16.2	-0.005264	-0.00364
138	16.3	-0.004909	-0.003464
139	16.4	-0.004571	-0.003297
140	16.5	-0.00425	-0.00313
141	16.6	-0.003945	-0.002969
142	16.7	-0.003656	-0.002808
143	16.8	-0.003383	-0.002657
144	16.9	-0.003124	-0.002514
145	17	-0.00288	-0.002376
146	17.1	-0.002649	-0.002246
147	17.2	-0.00243	-0.00213

148	17.3	-0.002222	-0.00202
149	17.4	-0.002025	-0.001915
150	17.5	-0.001839	-0.001817
151	17.6	-0.001662	-0.001721
152	17.7	-0.001494	-0.001632
153	17.8	-0.001336	-0.001544
154	17.9	-0.001185	-0.00146
155	18	-0.001043	-0.00138
156	18.1	-0.000909	-0.00131
157	18.2	-0.000782	-0.001236
158	18.3	-0.000661	-0.00117
159	18.4	-0.000548	-0.001101
160	18.5	-0.000441	-0.001037
161	18.6	-0.00034	-0.000974
162	18.7	-0.000246	-0.000907
163	18.8	-0.000158	-0.000848
164	18.9	-0.000076	-0.000789
165	19	0	-0.000741
MORSE_bpada1bpada3			
N	163		
1	0.2	4342939.75	1895882.802
2	2.3	361585.8671	1895882.802
3	2.5	117668.5471	543290.3977

4	2.8	21202.70787	99815.19739
5	2.9	13250.45583	59229.84353
6	3	8497.964113	35819.99074
7	3.2	3520.843805	13951.21234
8	3.3	2382.30122	8819.639368
9	3.4	1654.802303	5730.338963
10	3.7	525.133587	1800.785814
11	3.8	373.057108	1240.743753
12	3.9	267.837302	863.652367
13	4	193.667261	619.748451
14	4.1	139.963194	454.332897
15	4.2	101.121945	322.492086
16	4.3	73.402758	231.891647
17	4.4	53.418096	167.801589
18	4.5	38.732161	125.917114
19	4.6	27.797309	92.779934
20	4.7	19.749762	68.170992
21	4.8	13.793234	50.959577
22	4.9	9.390644	37.092214
23	5	6.231309	26.094487
24	5.1	3.97954	18.940896
25	5.2	2.327707	14.095762
26	5.3	1.09845	10.489387
27	5.4	0.185923	7.761148
28	5.5	-0.490558	5.76848

29	5.6	-0.996756	4.355467
30	5.7	-1.373567	3.180768
31	5.8	-1.6492	2.331881
32	5.9	-1.850705	1.698232
33	6	-1.993581	1.159273
34	6.1	-2.089743	0.763976
35	6.2	-2.148622	0.413594
36	6.3	-2.177077	0.155516
37	6.4	-2.183428	-0.0285
38	6.5	-2.172723	-0.185587
39	6.6	-2.147331	-0.322265
40	6.7	-2.111002	-0.404313
41	6.8	-2.068393	-0.447867
42	6.9	-2.019591	-0.528179
43	7	-1.96587	-0.546231
44	7.1	-1.909813	-0.574915
45	7.2	-1.851273	-0.59587
46	7.3	-1.791825	-0.593104
47	7.4	-1.731392	-0.61555
48	7.5	-1.669803	-0.616223
49	7.6	-1.607035	-0.639148
50	7.7	-1.542985	-0.641852
51	7.8	-1.478456	-0.648721
52	7.9	-1.413825	-0.6439
53	8	-1.34941	-0.644397

54	8.1	-1.28523	-0.63922
55	8.2	-1.222276	-0.619855
56	8.3	-1.160612	-0.613421
57	8.4	-1.100434	-0.590146
58	8.5	-1.042403	-0.570465
59	8.6	-0.986093	-0.555732
60	8.7	-0.931427	-0.537584
61	8.8	-0.878693	-0.517107
62	8.9	-0.827786	-0.501022
63	9	-0.778585	-0.483009
64	9.1	-0.731276	-0.463163
65	9.2	-0.685831	-0.445736
66	9.3	-0.642306	-0.424778
67	9.4	-0.600753	-0.406264
68	9.5	-0.560892	-0.390957
69	9.6	-0.522717	-0.372547
70	9.7	-0.486472	-0.352352
71	9.8	-0.452086	-0.335371
72	9.9	-0.419682	-0.312705
73	10	-0.389397	-0.292995
74	10.1	-0.361149	-0.271983
75	10.2	-0.334976	-0.251468
76	10.3	-0.31083	-0.231457
77	10.4	-0.288568	-0.213778
78	10.5	-0.268036	-0.196862

79	10.6	-0.249175	-0.180356
80	10.7	-0.231842	-0.166305
81	10.8	-0.215829	-0.153965
82	10.9	-0.201083	-0.140954
83	11	-0.187537	-0.129959
84	11.1	-0.175041	-0.119949
85	11.2	-0.163506	-0.110759
86	11.3	-0.152848	-0.102397
87	11.4	-0.142984	-0.094895
88	11.5	-0.133854	-0.087699
89	11.6	-0.125397	-0.081439
90	11.7	-0.117543	-0.07564
91	11.8	-0.110257	-0.070086
92	11.9	-0.1035	-0.065041
93	12	-0.097202	-0.060919
94	12.1	-0.09134	-0.056327
95	12.2	-0.08589	-0.052672
96	12.3	-0.080801	-0.049116
97	12.4	-0.076045	-0.045999
98	12.5	-0.071597	-0.042962
99	12.6	-0.067448	-0.040026
100	12.7	-0.063567	-0.037577
101	12.8	-0.059928	-0.035208
102	12.9	-0.056513	-0.033103
103	13	-0.053306	-0.031035

104	13.1	-0.050297	-0.029136
105	13.2	-0.047477	-0.027267
106	13.3	-0.04483	-0.025669
107	13.4	-0.042343	-0.024083
108	13.5	-0.039998	-0.022802
109	13.6	-0.037787	-0.021425
110	13.7	-0.03571	-0.020113
111	13.8	-0.033757	-0.018953
112	13.9	-0.031913	-0.017923
113	14	-0.030169	-0.016948
114	14.1	-0.028522	-0.016008
115	14.2	-0.026967	-0.015077
116	14.3	-0.0255	-0.01426
117	14.4	-0.024113	-0.013497
118	14.5	-0.022799	-0.012765
119	14.6	-0.021556	-0.012099
120	14.7	-0.02038	-0.011421
121	14.8	-0.019268	-0.010824
122	14.9	-0.018213	-0.010267
123	15	-0.017211	-0.009787
124	15.1	-0.016258	-0.009265
125	15.2	-0.015355	-0.008798
126	15.3	-0.014498	-0.008339
127	15.4	-0.013685	-0.00792
128	15.5	-0.012913	-0.007525

129	15.6	-0.012179	-0.007155
130	15.7	-0.011481	-0.00681
131	15.8	-0.010816	-0.006479
132	15.9	-0.010184	-0.006174
133	16	-0.009581	-0.005883
134	16.1	-0.009008	-0.005576
135	16.2	-0.008463	-0.005326
136	16.3	-0.007942	-0.005087
137	16.4	-0.007445	-0.004863
138	16.5	-0.00697	-0.004631
139	16.6	-0.006518	-0.004412
140	16.7	-0.006087	-0.0042
141	16.8	-0.005676	-0.004033
142	16.9	-0.005281	-0.003848
143	17	-0.004905	-0.003672
144	17.1	-0.004546	-0.003522
145	17.2	-0.004202	-0.003359
146	17.3	-0.003873	-0.003217
147	17.4	-0.003559	-0.003065
148	17.5	-0.003259	-0.002937
149	17.6	-0.002971	-0.002811
150	17.7	-0.002696	-0.002695
151	17.8	-0.002432	-0.002586
152	17.9	-0.002179	-0.002478
153	18	-0.001936	-0.002369

154	18.1	-0.001704	-0.002272
155	18.2	-0.001482	-0.002176
156	18.3	-0.001269	-0.00209
157	18.4	-0.001064	-0.002002
158	18.5	-0.000868	-0.001922
159	18.6	-0.00068	-0.00184
160	18.7	-0.000499	-0.001769
161	18.8	-0.000326	-0.001697
162	18.9	-0.00016	-0.001632
163	19	0	-0.001562
MORSE_bpada1mpd			
N	169		
1	0.2	20655216	10522528.23
2	1.9	2766918.009	10522528.23
3	2.4	56716.37181	318278.3178
4	2.5	32147.25484	173104.0215
5	2.6	18666.8471	96504.13332
6	2.7	11077.92457	55274.31719
7	2.8	6704.267588	32198.82247
8	2.9	4132.347418	19239.58094
9	3	2587.604713	11655.27316
10	3.1	1645.873137	7179.358343
11	3.2	1062.157292	4494.958569

12	3.3	694.76087	2852.969873
13	3.4	460.197959	1838.288345
14	3.5	308.208075	1201.509328
15	3.6	208.44062	793.839779
16	3.7	142.042831	534.116001
17	3.8	97.423283	358.274959
18	3.9	67.323979	243.711112
19	4	46.782601	167.11646
20	4.1	32.650999	115.515582
21	4.2	22.874432	80.015753
22	4.3	16.076439	55.944105
23	4.4	11.324707	39.090534
24	4.5	7.986124	27.681135
25	4.6	5.613289	19.775548
26	4.7	3.913696	14.21631
27	4.8	2.686843	10.320762
28	4.9	1.792022	7.575654
29	5	1.133431	5.596175
30	5.1	0.644837	4.17569
31	5.2	0.281746	3.086143
32	5.3	0.010572	2.337325
33	5.4	-0.193757	1.749258
34	5.5	-0.347299	1.321579
35	5.6	-0.463893	1.010304
36	5.7	-0.553024	0.772324

37	5.8	-0.620723	0.581658
38	5.9	-0.670848	0.420829
39	6	-0.705868	0.279577
40	6.1	-0.727714	0.157339
41	6.2	-0.738207	0.052522
42	6.3	-0.738168	-0.053288
43	6.4	-0.728458	-0.140913
44	6.5	-0.71095	-0.209264
45	6.6	-0.68736	-0.262521
46	6.7	-0.659481	-0.295061
47	6.8	-0.628675	-0.321053
48	6.9	-0.595754	-0.33737
49	7	-0.561586	-0.34599
50	7.1	-0.526921	-0.347317
51	7.2	-0.492116	-0.348778
52	7.3	-0.45734	-0.346747
53	7.4	-0.423009	-0.339877
54	7.5	-0.389458	-0.331146
55	7.6	-0.35685	-0.321002
56	7.7	-0.325427	-0.307465
57	7.8	-0.295489	-0.291288
58	7.9	-0.267298	-0.272537
59	8	-0.241065	-0.252117
60	8.1	-0.216805	-0.23309
61	8.2	-0.194522	-0.212576

62	8.3	-0.174206	-0.193751
63	8.4	-0.155672	-0.176928
64	8.5	-0.138826	-0.159985
65	8.6	-0.123619	-0.144154
66	8.7	-0.109888	-0.130469
67	8.8	-0.097484	-0.117613
68	8.9	-0.086317	-0.105716
69	9	-0.07627	-0.095225
70	9.1	-0.067228	-0.085613
71	9.2	-0.059096	-0.077026
72	9.3	-0.051796	-0.068981
73	9.4	-0.045251	-0.061918
74	9.5	-0.039388	-0.055335
75	9.6	-0.034153	-0.049371
76	9.7	-0.029483	-0.044036
77	9.8	-0.025319	-0.039239
78	9.9	-0.021612	-0.034894
79	10	-0.018322	-0.03091
80	10.1	-0.015414	-0.027248
81	10.2	-0.012854	-0.023964
82	10.3	-0.010611	-0.020893
83	10.4	-0.008655	-0.018232
84	10.5	-0.006956	-0.015744
85	10.6	-0.005494	-0.013497
86	10.7	-0.004248	-0.011427

87	10.8	-0.003198	-0.009569
88	10.9	-0.002324	-0.007899
89	11	-0.001612	-0.006337
90	11.1	-0.001049	-0.004932
91	11.2	-0.00062	-0.003642
92	11.3	-0.000314	-0.002494
93	11.4	-0.000117	-0.001437
94	11.5	-0.000023	-0.000451
95	11.6	0	0
96	11.7	0	0
97	11.8	0	0
98	11.9	0	0
99	12	0	0
100	12.1	0	0
101	12.2	0	0
102	12.3	0	0
103	12.4	0	0
104	12.5	0	0
105	12.6	0	0
106	12.7	0	0
107	12.8	0	0
108	12.9	0	0
109	13	0	0
110	13.1	0	0
111	13.2	0	0

112	13.3	0	0
113	13.4	0	0
114	13.5	0	0
115	13.6	0	0
116	13.7	0	0
117	13.8	0	0
118	13.9	0	0
119	14	0	0
120	14.1	0	0
121	14.2	0	0
122	14.3	0	0
123	14.4	0	0
124	14.5	0	0
125	14.6	0	0
126	14.7	0	0
127	14.8	0	0
128	14.9	0	0
129	15	0	0
130	15.1	0	0
131	15.2	0	0
132	15.3	0	0
133	15.4	0	0
134	15.5	0	0
135	15.6	0	0
136	15.7	0	0

137	15.8	0	0
138	15.9	0	0
139	16	0	0
140	16.1	0	0
141	16.2	0	0
142	16.3	0	0
143	16.4	0	0
144	16.5	0	0
145	16.6	0	0
146	16.7	0	0
147	16.8	0	0
148	16.9	0	0
149	17	0	0
150	17.1	0	0
151	17.2	0	0
152	17.3	0	0
153	17.4	0	0
154	17.5	0	0
155	17.6	0	0
156	17.7	0	0
157	17.8	0	0
158	17.9	0	0
159	18	0	0
160	18.1	0	0
161	18.2	0	0

162	18.3	0	0
163	18.4	0	0
164	18.5	0	0
165	18.6	0	0
166	18.7	0	0
167	18.8	0	0
168	18.9	0	0
169	19	0	0
MORSE_bpada1o			
N	176		
1	0.2	19266211.04	12658060.28
2	1.6	1544926.65	12658060.28
3	1.7	657015.2998	5100166.723
4	1.8	293758.1833	2164975.607
5	1.9	137366.0436	962867.1866
6	2	66910.14805	446250.7254
7	2.1	33844.38034	215064.6288
8	2.2	17725.23294	107318.319
9	2.3	9588.735115	55411.63754
10	2.4	5349.19647	29379.13536
11	2.5	3072.915878	16146.4765
12	2.6	1811.207377	9087.693525
13	2.7	1093.651978	5263.414443

14	2.8	675.31023	3103.420529
15	2.9	425.80932	1886.59767
16	3	272.854131	1172.506101
17	3.1	177.80505	728.475531
18	3.2	117.590163	475.822207
19	3.3	78.292209	310.136871
20	3.4	52.97531	196.201108
21	3.5	36.661968	130.065724
22	3.6	25.419496	94.783723
23	3.7	17.443926	64.727671
24	3.8	12.152638	41.098099
25	3.9	8.658006	28.794541
26	4	6.214795	20.069687
27	4.1	4.505633	14.113549
28	4.2	3.287769	10.243731
29	4.3	2.405907	7.393512
30	4.4	1.763506	5.454501
31	4.5	1.292041	3.974798
32	4.6	0.93353	3.195423
33	4.7	0.63456	2.783984
34	4.8	0.364888	2.609446
35	4.9	0.129791	2.0925
36	5	-0.050592	1.515149
37	5.1	-0.177966	1.032332
38	5.2	-0.262571	0.659779

39	5.3	-0.313745	0.363693
40	5.4	-0.340365	0.168716
41	5.5	-0.350511	0.034196
42	5.6	-0.349358	-0.057255
43	5.7	-0.340625	-0.117405
44	5.8	-0.326729	-0.160508
45	5.9	-0.309506	-0.183951
46	6	-0.29074	-0.191376
47	6.1	-0.271403	-0.195361
48	6.2	-0.251902	-0.194659
49	6.3	-0.232528	-0.19283
50	6.4	-0.213448	-0.188771
51	6.5	-0.19495	-0.181188
52	6.6	-0.177283	-0.172147
53	6.7	-0.160622	-0.161066
54	6.8	-0.145109	-0.149206
55	6.9	-0.130788	-0.137208
56	7	-0.11765	-0.125547
57	7.1	-0.105661	-0.114241
58	7.2	-0.094769	-0.103593
59	7.3	-0.084891	-0.093973
60	7.4	-0.075953	-0.084792
61	7.5	-0.067888	-0.076505
62	7.6	-0.06062	-0.068858
63	7.7	-0.054068	-0.062173

64	7.8	-0.048166	-0.055873
65	7.9	-0.042854	-0.050368
66	8	-0.038068	-0.045345
67	8.1	-0.033764	-0.040731
68	8.2	-0.029896	-0.03663
69	8.3	-0.02642	-0.032889
70	8.4	-0.0233	-0.029519
71	8.5	-0.0205	-0.026481
72	8.6	-0.017989	-0.023735
73	8.7	-0.015741	-0.021233
74	8.8	-0.01373	-0.018978
75	8.9	-0.011932	-0.016973
76	9	-0.010327	-0.015143
77	9.1	-0.008896	-0.013476
78	9.2	-0.007622	-0.011989
79	9.3	-0.006492	-0.010625
80	9.4	-0.00549	-0.009406
81	9.5	-0.004606	-0.00828
82	9.6	-0.003829	-0.007261
83	9.7	-0.003149	-0.006338
84	9.8	-0.002557	-0.005499
85	9.9	-0.002045	-0.004738
86	10	-0.001606	-0.004047
87	10.1	-0.001233	-0.003418
88	10.2	-0.000919	-0.002853

89	10.3	-0.00066	-0.002332
90	10.4	-0.00045	-0.001859
91	10.5	-0.000286	-0.001429
92	10.6	-0.000163	-0.001037
93	10.7	-0.000077	-0.000683
94	10.8	-0.000025	-0.000358
95	10.9	-0.000003	-0.000066
96	11	0	0
97	11.1	0	0
98	11.2	0	0
99	11.3	0	0
100	11.4	0	0
101	11.5	0	0
102	11.6	0	0
103	11.7	0	0
104	11.8	0	0
105	11.9	0	0
106	12	0	0
107	12.1	0	0
108	12.2	0	0
109	12.3	0	0
110	12.4	0	0
111	12.5	0	0
112	12.6	0	0
113	12.7	0	0

114	12.8	0	0
115	12.9	0	0
116	13	0	0
117	13.1	0	0
118	13.2	0	0
119	13.3	0	0
120	13.4	0	0
121	13.5	0	0
122	13.6	0	0
123	13.7	0	0
124	13.8	0	0
125	13.9	0	0
126	14	0	0
127	14.1	0	0
128	14.2	0	0
129	14.3	0	0
130	14.4	0	0
131	14.5	0	0
132	14.6	0	0
133	14.7	0	0
134	14.8	0	0
135	14.9	0	0
136	15	0	0
137	15.1	0	0
138	15.2	0	0

139	15.3	0	0
140	15.4	0	0
141	15.5	0	0
142	15.6	0	0
143	15.7	0	0
144	15.8	0	0
145	15.9	0	0
146	16	0	0
147	16.1	0	0
148	16.2	0	0
149	16.3	0	0
150	16.4	0	0
151	16.5	0	0
152	16.6	0	0
153	16.7	0	0
154	16.8	0	0
155	16.9	0	0
156	17	0	0
157	17.1	0	0
158	17.2	0	0
159	17.3	0	0
160	17.4	0	0
161	17.5	0	0
162	17.6	0	0
163	17.7	0	0

164	17.8	0	0
165	17.9	0	0
166	18	0	0
167	18.1	0	0
168	18.2	0	0
169	18.3	0	0
170	18.4	0	0
171	18.5	0	0
172	18.6	0	0
173	18.7	0	0
174	18.8	0	0
175	18.9	0	0
176	19	0	0
MORSE_bpada1tape3			
N	161		
1	0.2	22259208.1	10667288.01
2	2.1	1991360.873	10667288.01
3	2.3	652013.4498	2726186.221
4	2.7	57388.08109	246940.6225
5	3	12615.03195	51546.37179
6	3.1	8454.218142	31669.90435
7	3.2	5879.327694	19827.90462
8	3.6	1150.738615	3815.040771

9	3.7	826.47018	2670.327924
10	3.8	600.978799	1839.499703
11	3.9	443.307082	1313.934644
12	4.1	237.981754	739.318629
13	4.2	176.24184	495.47966
14	4.3	132.423396	380.889215
15	4.4	99.73586	272.861506
16	4.5	75.880693	204.241828
17	4.6	58.050496	152.362109
18	4.7	44.912825	110.391308
19	4.8	34.20432	103.778809
20	4.9	25.399617	72.31524
21	5	18.927517	57.126758
22	5.1	13.801111	45.401368
23	5.2	10.034407	29.932712
24	5.3	7.416672	22.421999
25	5.4	5.363563	18.64017
26	5.5	3.661119	15.408703
27	5.6	2.291106	11.991569
28	5.7	1.191089	10.008767
29	5.8	0.31659	7.481213
30	5.9	-0.360899	6.068567
31	6	-0.92337	5.180857
32	6.1	-1.384923	4.050196
33	6.2	-1.738263	3.016614

34	6.3	-2.00856	2.389324
35	6.4	-2.204999	1.539448
36	6.5	-2.33015	0.96357
37	6.6	-2.404116	0.515754
38	6.7	-2.443391	0.269758
39	6.8	-2.455335	-0.030883
40	6.9	-2.44182	-0.239427
41	7	-2.409606	-0.404846
42	7.1	-2.362691	-0.533447
43	7.2	-2.305854	-0.603298
44	7.3	-2.242403	-0.66572
45	7.4	-2.174245	-0.697445
46	7.5	-2.103917	-0.709112
47	7.6	-2.030817	-0.752899
48	7.7	-1.954747	-0.768495
49	7.8	-1.877057	-0.785311
50	7.9	-1.798714	-0.781537
51	8	-1.721283	-0.76709
52	8.1	-1.645397	-0.750625
53	8.2	-1.571225	-0.732823
54	8.3	-1.499046	-0.710753
55	8.4	-1.429878	-0.672615
56	8.5	-1.364331	-0.638313
57	8.6	-1.301675	-0.614803
58	8.7	-1.241449	-0.589728

59	8.8	-1.18335	-0.57224
60	8.9	-1.127869	-0.537391
61	9	-1.074703	-0.525923
62	9.1	-1.023123	-0.505682
63	9.2	-0.973324	-0.490298
64	9.3	-0.924794	-0.480296
65	9.4	-0.877327	-0.469059
66	9.5	-0.831046	-0.456542
67	9.6	-0.786071	-0.442969
68	9.7	-0.742618	-0.42608
69	9.8	-0.70097	-0.406884
70	9.9	-0.661039	-0.391735
71	10	-0.622878	-0.371486
72	10.1	-0.586698	-0.352111
73	10.2	-0.552361	-0.334642
74	10.3	-0.519975	-0.313071
75	10.4	-0.489545	-0.295537
76	10.5	-0.461004	-0.27528
77	10.6	-0.434146	-0.261886
78	10.7	-0.408515	-0.250721
79	10.8	-0.384075	-0.23809
80	10.9	-0.360784	-0.227723
81	11	-0.338487	-0.218212
82	11.1	-0.317285	-0.205835
83	11.2	-0.297378	-0.192305

84	11.3	-0.278803	-0.179204
85	11.4	-0.26142	-0.168456
86	11.5	-0.245246	-0.15502
87	11.6	-0.230332	-0.143253
88	11.7	-0.216478	-0.133837
89	11.8	-0.203551	-0.124691
90	11.9	-0.191553	-0.115279
91	12	-0.180466	-0.106458
92	12.1	-0.170188	-0.099108
93	12.2	-0.16059	-0.092853
94	12.3	-0.151632	-0.086305
95	12.4	-0.143289	-0.080558
96	12.5	-0.135506	-0.075088
97	12.6	-0.128234	-0.070363
98	12.7	-0.121424	-0.065836
99	12.8	-0.115047	-0.061704
100	12.9	-0.109074	-0.057753
101	13	-0.103465	-0.054422
102	13.1	-0.098189	-0.051099
103	13.2	-0.093231	-0.048059
104	13.3	-0.088574	-0.045083
105	13.4	-0.084191	-0.042574
106	13.5	-0.08005	-0.040256
107	13.6	-0.07614	-0.037948
108	13.7	-0.072454	-0.035761

109	13.8	-0.068973	-0.033862
110	13.9	-0.065677	-0.03207
111	14	-0.062546	-0.030534
112	14.1	-0.059577	-0.028845
113	14.2	-0.056758	-0.027539
114	14.3	-0.054072	-0.026184
115	14.4	-0.051528	-0.024701
116	14.5	-0.049107	-0.023706
117	14.6	-0.046794	-0.022551
118	14.7	-0.044594	-0.021455
119	14.8	-0.042497	-0.020492
120	14.9	-0.040493	-0.019586
121	15	-0.038575	-0.018765
122	15.1	-0.036743	-0.01789
123	15.2	-0.034994	-0.017085
124	15.3	-0.033321	-0.016382
125	15.4	-0.031718	-0.015665
126	15.5	-0.030183	-0.01503
127	15.6	-0.02871	-0.014431
128	15.7	-0.027298	-0.013821
129	15.8	-0.025939	-0.013355
130	15.9	-0.02463	-0.01282
131	16	-0.023373	-0.012317
132	16.1	-0.022164	-0.011867
133	16.2	-0.020999	-0.011433

134	16.3	-0.019877	-0.011012
135	16.4	-0.018796	-0.010599
136	16.5	-0.017755	-0.010218
137	16.6	-0.01675	-0.009896
138	16.7	-0.015776	-0.009572
139	16.8	-0.014837	-0.009225
140	16.9	-0.013929	-0.008922
141	17	-0.01305	-0.008663
142	17.1	-0.012198	-0.008377
143	17.2	-0.011374	-0.008108
144	17.3	-0.010576	-0.007856
145	17.4	-0.009802	-0.007612
146	17.5	-0.009052	-0.007391
147	17.6	-0.008324	-0.007167
148	17.7	-0.007618	-0.006959
149	17.8	-0.006932	-0.006763
150	17.9	-0.006265	-0.006575
151	18	-0.005616	-0.006396
152	18.1	-0.004986	-0.006213
153	18.2	-0.004373	-0.006054
154	18.3	-0.003775	-0.005892
155	18.4	-0.003193	-0.005746
156	18.5	-0.002627	-0.00559
157	18.6	-0.002075	-0.005448
158	18.7	-0.001536	-0.005316

159	18.8	-0.001012	-0.005182
160	18.9	-0.0005	-0.005056
161	19	0	-0.004936
MORSE_bpada3			
N	160		
1	0.2	25673816105	15556273792
2	1.4	7006287555	15556273792
3	2.3	1879749.906	9076885.642
4	2.6	301704.9755	1443413.892
5	2.8	109847.3355	475162.5084
6	3	45426.0298	169050.549
7	3.1	31799.94662	103471.1146
8	3.6	3126.509892	11222.63232
9	3.8	1496.545793	5077.008675
10	4	743.357333	2454.875922
11	4.1	534.818757	1715.895587
12	4.2	388.229889	1215.881774
13	4.3	284.090286	866.910284
14	4.4	209.729525	620.304936
15	4.5	156.243855	449.408481
16	4.6	116.686675	341.735122
17	4.7	87.4851	242.296374
18	4.8	66.299389	181.41784

19	4.9	50.271719	139.135567
20	5	38.165995	102.978908
21	5.1	28.929983	81.741323
22	5.2	21.749627	61.865807
23	5.3	16.364184	45.843059
24	5.4	12.323211	34.976388
25	5.5	9.169527	28.097292
26	5.6	6.690473	21.483795
27	5.7	4.789637	16.532919
28	5.8	3.318071	12.8984
29	5.9	2.175409	9.954853
30	6	1.290858	7.736168
31	6.1	0.599964	6.081696
32	6.2	0.057494	4.767703
33	6.3	-0.366608	3.714348
34	6.4	-0.698861	2.930711
35	6.5	-0.958817	2.268401
36	6.6	-1.159867	1.752605
37	6.7	-1.31579	1.365847
38	6.8	-1.435992	1.038201
39	6.9	-1.528452	0.810995
40	7	-1.599086	0.601684
41	7.1	-1.651155	0.439695
42	7.2	-1.688596	0.309126
43	7.3	-1.7124	0.166958

44	7.4	-1.722198	0.028998
45	7.5	-1.719307	-0.086813
46	7.6	-1.706047	-0.178383
47	7.7	-1.683898	-0.264604
48	7.8	-1.655239	-0.308579
49	7.9	-1.621511	-0.365976
50	8	-1.582399	-0.416267
51	8.1	-1.539848	-0.434742
52	8.2	-1.494949	-0.463236
53	8.3	-1.448167	-0.472405
54	8.4	-1.400131	-0.488313
55	8.5	-1.350521	-0.503896
56	8.6	-1.300966	-0.48721
57	8.7	-1.251775	-0.496599
58	8.8	-1.202269	-0.493534
59	8.9	-1.153528	-0.481277
60	9	-1.105952	-0.470247
61	9.1	-1.058839	-0.472005
62	9.2	-1.011876	-0.467261
63	9.3	-0.965146	-0.467334
64	9.4	-0.918694	-0.461716
65	9.5	-0.872697	-0.458222
66	9.6	-0.827649	-0.44274
67	9.7	-0.783733	-0.435568
68	9.8	-0.740884	-0.421412

69	9.9	-0.699603	-0.404211
70	10	-0.659884	-0.390176
71	10.1	-0.621805	-0.37141
72	10.2	-0.585689	-0.350906
73	10.3	-0.551451	-0.333843
74	10.4	-0.519171	-0.311765
75	10.5	-0.48867	-0.298249
76	10.6	-0.459684	-0.281467
77	10.7	-0.432201	-0.268191
78	10.8	-0.406034	-0.255162
79	10.9	-0.381215	-0.241204
80	11	-0.357699	-0.229121
81	11.1	-0.33549	-0.215066
82	11.2	-0.31465	-0.201731
83	11.3	-0.295112	-0.189027
84	11.4	-0.276869	-0.175842
85	11.5	-0.25986	-0.16432
86	11.6	-0.244008	-0.152721
87	11.7	-0.229324	-0.140975
88	11.8	-0.215695	-0.131593
89	11.9	-0.203015	-0.122021
90	12	-0.191234	-0.113588
91	12.1	-0.180213	-0.10684
92	12.2	-0.169933	-0.098755
93	12.3	-0.160399	-0.091924

94	12.4	-0.151507	-0.085918
95	12.5	-0.1432	-0.080214
96	12.6	-0.135437	-0.075052
97	12.7	-0.128171	-0.070265
98	12.8	-0.121369	-0.065788
99	12.9	-0.114997	-0.06164
100	13	-0.109014	-0.058014
101	13.1	-0.103392	-0.054435
102	13.2	-0.09812	-0.051005
103	13.3	-0.093154	-0.048306
104	13.4	-0.088462	-0.045536
105	13.5	-0.084044	-0.042835
106	13.6	-0.079877	-0.040502
107	13.7	-0.075935	-0.03834
108	13.8	-0.07221	-0.036159
109	13.9	-0.068691	-0.034215
110	14	-0.065358	-0.032445
111	14.1	-0.062198	-0.030758
112	14.2	-0.059205	-0.029105
113	14.3	-0.056367	-0.027662
114	14.4	-0.053663	-0.0264
115	14.5	-0.051094	-0.024984
116	14.6	-0.048651	-0.023892
117	14.7	-0.046317	-0.022776
118	14.8	-0.044094	-0.021678

119	14.9	-0.04198	-0.020618
120	15	-0.039963	-0.019706
121	15.1	-0.038037	-0.018816
122	15.2	-0.036196	-0.018016
123	15.3	-0.034437	-0.017153
124	15.4	-0.032758	-0.016435
125	15.5	-0.031146	-0.015802
126	15.6	-0.029597	-0.015169
127	15.7	-0.028112	-0.014542
128	15.8	-0.026689	-0.013918
129	15.9	-0.025324	-0.013388
130	16	-0.024011	-0.012857
131	16.1	-0.022752	-0.012338
132	16.2	-0.02154	-0.011886
133	16.3	-0.020375	-0.011421
134	16.4	-0.019253	-0.011022
135	16.5	-0.018171	-0.010615
136	16.6	-0.017128	-0.01024
137	16.7	-0.016121	-0.009897
138	16.8	-0.015149	-0.009544
139	16.9	-0.014212	-0.009211
140	17	-0.013306	-0.008897
141	17.1	-0.01243	-0.008622
142	17.2	-0.011582	-0.00834
143	17.3	-0.010762	-0.008068

144	17.4	-0.009968	-0.007806
145	17.5	-0.009199	-0.007572
146	17.6	-0.008453	-0.007347
147	17.7	-0.00773	-0.007114
148	17.8	-0.007029	-0.006903
149	17.9	-0.006349	-0.006701
150	18	-0.005688	-0.006514
151	18.1	-0.005046	-0.006326
152	18.2	-0.004423	-0.00615
153	18.3	-0.003816	-0.005977
154	18.4	-0.003226	-0.005818
155	18.5	-0.002653	-0.005661
156	18.6	-0.002094	-0.00551
157	18.7	-0.00155	-0.005362
158	18.8	-0.001021	-0.005234
159	18.9	-0.000504	-0.005102
160	19	0	-0.004973
MORSE_bpada3mpd			
N	175		
1	0.2	2353413125	1659269761
2	1.5	196362435.6	1659269761
3	1.6	81888359.25	630211764.7
4	1.7	37685968.77	253836044.8

5	1.9	7516200.444	47861638.43
6	2.1	1663400.654	10666359.48
7	2.2	864643.5623	5308782.353
8	2.3	462929.256	2725503.773
9	2.4	254673.9555	1439602.238
10	2.5	143666.0929	780555.0152
11	2.6	82953.41901	433698.4617
12	2.7	48950.37477	246362.4232
13	2.8	29487.05837	142903.9049
14	2.9	18114.78095	84541.6435
15	3	11336.41523	51025.67088
16	3.1	7217.253076	31357.57213
17	3.2	4668.7287	19612.9154
18	3.3	3068.345705	12394.74451
19	3.4	2049.316849	7985.83262
20	3.5	1387.656997	5247.364418
21	3.6	951.615768	3473.460155
22	3.7	661.344707	2331.961056
23	3.8	464.136256	1612.20797
24	3.9	328.177165	1106.973845
25	4	234.120089	774.167684
26	4.1	167.995958	548.314944
27	4.2	121.227776	387.048679
28	4.3	88.074364	276.019572
29	4.4	64.343114	198.605432

30	4.5	47.167432	144.908198
31	4.6	34.625543	105.929578
32	4.7	25.434227	77.896756
33	4.8	18.66163	57.555181
34	4.9	13.752772	40.621977
35	5	10.208828	30.256901
36	5.1	7.56452	22.629254
37	5.2	5.582371	17.013734
38	5.3	4.085361	12.926453
39	5.4	2.941877	9.943231
40	5.5	2.063721	7.6199
41	5.6	1.387036	5.913794
42	5.7	0.863696	4.553003
43	5.8	0.46047	3.511515
44	5.9	0.14961	2.7057
45	6	-0.090047	2.087423
46	6.1	-0.276946	1.650557
47	6.2	-0.423568	1.281888
48	6.3	-0.537758	1.001906
49	6.4	-0.626643	0.775804
50	6.5	-0.694672	0.584769
51	6.6	-0.745307	0.427941
52	6.7	-0.780999	0.285904
53	6.8	-0.803616	0.166428
54	6.9	-0.815123	0.063713

55	7	-0.817436	-0.01746
56	7.1	-0.811795	-0.095361
57	7.2	-0.799684	-0.146854
58	7.3	-0.782786	-0.191101
59	7.4	-0.76193	-0.226019
60	7.5	-0.738072	-0.251135
61	7.6	-0.711914	-0.272027
62	7.7	-0.684193	-0.282401
63	7.8	-0.655442	-0.29262
64	7.9	-0.626335	-0.289514
65	8	-0.59725	-0.292194
66	8.1	-0.567953	-0.293735
67	8.2	-0.538602	-0.293299
68	8.3	-0.509229	-0.294149
69	8.4	-0.47999	-0.290647
70	8.5	-0.451083	-0.287482
71	8.6	-0.422537	-0.283441
72	8.7	-0.394429	-0.278712
73	8.8	-0.366942	-0.271043
74	8.9	-0.340313	-0.261519
75	9	-0.314772	-0.24931
76	9.1	-0.290531	-0.235502
77	9.2	-0.267635	-0.222435
78	9.3	-0.246088	-0.208489
79	9.4	-0.22602	-0.192886

80	9.5	-0.20751	-0.177312
81	9.6	-0.190476	-0.163363
82	9.7	-0.174756	-0.151028
83	9.8	-0.160302	-0.138052
84	9.9	-0.14706	-0.126801
85	10	-0.134898	-0.116441
86	10.1	-0.12376	-0.106301
87	10.2	-0.11357	-0.097517
88	10.3	-0.104241	-0.089062
89	10.4	-0.095681	-0.082135
90	10.5	-0.087817	-0.075135
91	10.6	-0.080606	-0.06909
92	10.7	-0.073976	-0.063514
93	10.8	-0.067876	-0.058476
94	10.9	-0.062259	-0.053864
95	11	-0.057094	-0.049449
96	11.1	-0.052357	-0.045278
97	11.2	-0.048005	-0.041774
98	11.3	-0.043991	-0.038501
99	11.4	-0.04029	-0.035512
100	11.5	-0.036885	-0.032598
101	11.6	-0.033747	-0.030167
102	11.7	-0.030856	-0.027641
103	11.8	-0.028194	-0.025601
104	11.9	-0.025735	-0.02358

105	12	-0.023469	-0.021747
106	12.1	-0.021382	-0.019983
107	12.2	-0.019455	-0.018561
108	12.3	-0.017675	-0.017032
109	12.4	-0.016042	-0.015634
110	12.5	-0.014543	-0.014349
111	12.6	-0.013161	-0.013297
112	12.7	-0.011885	-0.01222
113	12.8	-0.010712	-0.011247
114	12.9	-0.009633	-0.010325
115	13	-0.008644	-0.009459
116	13.1	-0.007735	-0.008707
117	13.2	-0.006902	-0.007968
118	13.3	-0.006138	-0.007303
119	13.4	-0.00544	-0.006662
120	13.5	-0.004803	-0.006067
121	13.6	-0.004223	-0.005537
122	13.7	-0.003694	-0.005041
123	13.8	-0.003215	-0.004552
124	13.9	-0.00278	-0.004134
125	14	-0.002389	-0.003704
126	14.1	-0.002036	-0.003351
127	14.2	-0.001719	-0.002978
128	14.3	-0.001437	-0.002677
129	14.4	-0.001186	-0.002341

130	14.5	-0.000965	-0.002063
131	14.6	-0.000773	-0.001778
132	14.7	-0.000607	-0.001559
133	14.8	-0.000462	-0.001331
134	14.9	-0.00034	-0.001118
135	15	-0.000239	-0.000894
136	15.1	-0.000158	-0.00072
137	15.2	-0.000095	-0.000544
138	15.3	-0.000049	-0.000374
139	15.4	-0.000019	-0.000225
140	15.5	-0.000004	-0.000079
141	15.6	0	0
142	15.7	0	0
143	15.8	0	0
144	15.9	0	0
145	16	0	0
146	16.1	0	0
147	16.2	0	0
148	16.3	0	0
149	16.4	0	0
150	16.5	0	0
151	16.6	0	0
152	16.7	0	0
153	16.8	0	0
154	16.9	0	0

155	17	0	0
156	17.1	0	0
157	17.2	0	0
158	17.3	0	0
159	17.4	0	0
160	17.5	0	0
161	17.6	0	0
162	17.7	0	0
163	17.8	0	0
164	17.9	0	0
165	18	0	0
166	18.1	0	0
167	18.2	0	0
168	18.3	0	0
169	18.4	0	0
170	18.5	0	0
171	18.6	0	0
172	18.7	0	0
173	18.8	0	0
174	18.9	0	0
175	19	0	0
MORSE_bpada3tape3			
N	176		

1	0.2	4.22816E+11	4.15111E+11
2	1.1	49216341055	4.15111E+11
3	1.3	4317556056	33876937868
4	1.5	533871037	3959912317
5	1.6	260674199.1	1504024440
6	1.9	17937142.82	114222601.9
7	2.1	3969361.691	25455209.45
8	2.2	2063157.003	12668884.31
9	2.3	1104520.94	6503836.951
10	2.4	607556.5267	3435451.316
11	2.5	342650.8186	1862662.846
12	2.6	197806.0244	1034233.038
13	2.7	116720.0181	587487.0893
14	2.8	70303.80399	340837.1923
15	2.9	43182.62649	201586.3577
16	3	27023.70948	121591.9824
17	3.1	17228.27641	74316.67894
18	3.2	11187.80721	46492.70515
19	3.3	7391.597514	29431.48874
20	3.4	4971.057207	18979.31739
21	3.5	3392.404386	12593.73903
22	3.6	2347.038692	8313.574864
23	3.7	1653.667247	5553.854033
24	3.8	1185.682704	3805.836821
25	3.9	861.987491	2668.067446

26	4	635.188447	1867.913444
27	4.1	475.638265	1323.09018
28	4.2	359.565492	998.365291
29	4.3	272.852454	735.895456
30	4.4	209.188591	537.381806
31	4.5	160.757046	431.249093
32	4.6	122.620058	331.490668
33	4.7	93.709605	246.718388
34	4.8	71.992567	187.622386
35	4.9	55.254424	147.140475
36	5	42.288154	112.184919
37	5.1	32.265753	88.263095
38	5.2	24.337386	70.304247
39	5.3	18.107206	54.299355
40	5.4	13.324516	41.35445
41	5.5	9.648889	32.158088
42	5.6	6.872145	23.376798
43	5.7	4.758898	18.888136
44	5.8	3.089671	14.496399
45	5.9	1.816503	10.966958
46	6	0.852357	8.315965
47	6.1	0.123156	6.26805
48	6.2	-0.41754	4.545888
49	6.3	-0.831618	3.735665
50	6.4	-1.154947	2.730915

51	6.5	-1.39572	2.084547
52	6.6	-1.579296	1.58697
53	6.7	-1.728043	1.387962
54	6.8	-1.857132	1.193832
55	6.9	-1.965039	0.964302
56	7	-2.051306	0.761043
57	7.1	-2.120183	0.61649
58	7.2	-2.170273	0.385304
59	7.3	-2.199966	0.208561
60	7.4	-2.210187	-0.004131
61	7.5	-2.203436	-0.130889
62	7.6	-2.183593	-0.265971
63	7.7	-2.152941	-0.347073
64	7.8	-2.112931	-0.453125
65	7.9	-2.064817	-0.509161
66	8	-2.012073	-0.54571
67	8.1	-1.955508	-0.585593
68	8.2	-1.896436	-0.595848
69	8.3	-1.836216	-0.608564
70	8.4	-1.775286	-0.610033
71	8.5	-1.714615	-0.603384
72	8.6	-1.653787	-0.613173
73	8.7	-1.592379	-0.614998
74	8.8	-1.530879	-0.615002
75	8.9	-1.469827	-0.606033

76	9	-1.409613	-0.598252
77	9.1	-1.350228	-0.589445
78	9.2	-1.291468	-0.585756
79	9.3	-1.233832	-0.566948
80	9.4	-1.177694	-0.555821
81	9.5	-1.122733	-0.543407
82	9.6	-1.068897	-0.533305
83	9.7	-1.016141	-0.521807
84	9.8	-0.964616	-0.5087
85	9.9	-0.914433	-0.494956
86	10	-0.865737	-0.478966
87	10.1	-0.818841	-0.458955
88	10.2	-0.774158	-0.434716
89	10.3	-0.731579	-0.416861
90	10.4	-0.690966	-0.395394
91	10.5	-0.652368	-0.376556
92	10.6	-0.615603	-0.358753
93	10.7	-0.58076	-0.338111
94	10.8	-0.54765	-0.324077
95	10.9	-0.516006	-0.30881
96	11	-0.48581	-0.295115
97	11.1	-0.456796	-0.285163
98	11.2	-0.429019	-0.270379
99	11.3	-0.402572	-0.258555
100	11.4	-0.377389	-0.245117

101	11.5	-0.353561	-0.231437
102	11.6	-0.331156	-0.216654
103	11.7	-0.310148	-0.203505
104	11.8	-0.290491	-0.189642
105	11.9	-0.272171	-0.176749
106	12	-0.255153	-0.163622
107	12.1	-0.239378	-0.151874
108	12.2	-0.22473	-0.141078
109	12.3	-0.211134	-0.130848
110	12.4	-0.198492	-0.121986
111	12.5	-0.186701	-0.113848
112	12.6	-0.175715	-0.105857
113	12.7	-0.165496	-0.098536
114	12.8	-0.155986	-0.091659
115	12.9	-0.147129	-0.085481
116	13	-0.13886	-0.079891
117	13.1	-0.131126	-0.074796
118	13.2	-0.123873	-0.070265
119	13.3	-0.117066	-0.065875
120	13.4	-0.110689	-0.061662
121	13.5	-0.104709	-0.057932
122	13.6	-0.099098	-0.054299
123	13.7	-0.093819	-0.051279
124	13.8	-0.088845	-0.048195
125	13.9	-0.084162	-0.045478

126	14	-0.079748	-0.042804
127	14.1	-0.07558	-0.040554
128	14.2	-0.071645	-0.038144
129	14.3	-0.067938	-0.035999
130	14.4	-0.06443	-0.034156
131	14.5	-0.061102	-0.032397
132	14.6	-0.057948	-0.030683
133	14.7	-0.054963	-0.029013
134	14.8	-0.05214	-0.027459
135	14.9	-0.049459	-0.026167
136	15	-0.046907	-0.02486
137	15.1	-0.044484	-0.023599
138	15.2	-0.042181	-0.022467
139	15.3	-0.039987	-0.021404
140	15.4	-0.037898	-0.020387
141	15.5	-0.035906	-0.019455
142	15.6	-0.034006	-0.018536
143	15.7	-0.032194	-0.01771
144	15.8	-0.030463	-0.0169
145	15.9	-0.028809	-0.01619
146	16	-0.027226	-0.01547
147	16.1	-0.025714	-0.014775
148	16.2	-0.024267	-0.014148
149	16.3	-0.022881	-0.013574
150	16.4	-0.021551	-0.013031

151	16.5	-0.020274	-0.012502
152	16.6	-0.01905	-0.011979
153	16.7	-0.017877	-0.011485
154	16.8	-0.01675	-0.011062
155	16.9	-0.015665	-0.01063
156	17	-0.014623	-0.010212
157	17.1	-0.013621	-0.009823
158	17.2	-0.012658	-0.009453
159	17.3	-0.011728	-0.009131
160	17.4	-0.010832	-0.00879
161	17.5	-0.009969	-0.008476
162	17.6	-0.009136	-0.008178
163	17.7	-0.008333	-0.007887
164	17.8	-0.007557	-0.007627
165	17.9	-0.006808	-0.007356
166	18	-0.006085	-0.007113
167	18.1	-0.005385	-0.006883
168	18.2	-0.004708	-0.006661
169	18.3	-0.004052	-0.006444
170	18.4	-0.003418	-0.006241
171	18.5	-0.002803	-0.006053
172	18.6	-0.002208	-0.005859
173	18.7	-0.001631	-0.005682
174	18.8	-0.001071	-0.005517
175	18.9	-0.000527	-0.005352

176	19	0	-0.005197
MORSE_mpd			
N	167		
1	0.2	235546.2805	94734.13701
2	2.5	17657.76534	94734.13701
3	2.6	10272.07279	52979.71395
4	2.7	6104.508702	30371.56781
5	2.8	3698.05604	17757.48542
6	2.9	2280.490265	10593.83009
7	3	1429.894716	6418.080889
8	3.1	911.237638	3955.060672
9	3.2	589.547352	2478.745053
10	3.3	386.765497	1576.892049
11	3.4	257.047637	1017.46515
12	3.5	172.950553	664.47652
13	3.6	117.737025	439.794034
14	3.7	81.014306	294.66035
15	3.8	56.293265	199.76048
16	3.9	39.454263	137.019552
17	4	27.858112	94.903478
18	4.1	19.803514	66.188478
19	4.2	14.168506	46.511671
20	4.3	10.192996	32.998527

21	4.4	7.35967	23.667994
22	4.5	5.316682	17.191782
23	4.6	3.824086	12.660126
24	4.7	2.71875	9.446591
25	4.8	1.889166	7.145099
26	4.9	1.258506	5.468104
27	5	0.77405	4.221006
28	5.1	0.400069	3.258618
29	5.2	0.11186	2.505553
30	5.3	-0.109248	1.91661
31	5.4	-0.278689	1.47222
32	5.5	-0.408807	1.130138
33	5.6	-0.508435	0.86241
34	5.7	-0.583691	0.64271
35	5.8	-0.638821	0.459903
36	5.9	-0.676113	0.285931
37	6	-0.697459	0.140984
38	6.1	-0.705186	0.013565
39	6.2	-0.701347	-0.090351
40	6.3	-0.68777	-0.181181
41	6.4	-0.666364	-0.246946
42	6.5	-0.639353	-0.293265
43	6.6	-0.608527	-0.323254
44	6.7	-0.575221	-0.342884
45	6.8	-0.540645	-0.348619

46	6.9	-0.505771	-0.348864
47	7	-0.471264	-0.341272
48	7.1	-0.437704	-0.329945
49	7.2	-0.405532	-0.313483
50	7.3	-0.375074	-0.295669
51	7.4	-0.34649	-0.276012
52	7.5	-0.319883	-0.256141
53	7.6	-0.295247	-0.236573
54	7.7	-0.272526	-0.217848
55	7.8	-0.251633	-0.200002
56	7.9	-0.232457	-0.183532
57	8	-0.214871	-0.168191
58	8.1	-0.19875	-0.154222
59	8.2	-0.183971	-0.141359
60	8.3	-0.170427	-0.129511
61	8.4	-0.158009	-0.118861
62	8.5	-0.146614	-0.109044
63	8.6	-0.136154	-0.100156
64	8.7	-0.126547	-0.091969
65	8.8	-0.11772	-0.084581
66	8.9	-0.1096	-0.077813
67	9	-0.102121	-0.071763
68	9.1	-0.09523	-0.066067
69	9.2	-0.088881	-0.060914
70	9.3	-0.083024	-0.056228

71	9.4	-0.077613	-0.051992
72	9.5	-0.072608	-0.048092
73	9.6	-0.067979	-0.0445
74	9.7	-0.063695	-0.041183
75	9.8	-0.059727	-0.038178
76	9.9	-0.056045	-0.035445
77	10	-0.052627	-0.032916
78	10.1	-0.049451	-0.030613
79	10.2	-0.046496	-0.028485
80	10.3	-0.043747	-0.026503
81	10.4	-0.041186	-0.024701
82	10.5	-0.038801	-0.023006
83	10.6	-0.036578	-0.021449
84	10.7	-0.034505	-0.020029
85	10.8	-0.032568	-0.018708
86	10.9	-0.030758	-0.017489
87	11	-0.029065	-0.016375
88	11.1	-0.02748	-0.015324
89	11.2	-0.025995	-0.014365
90	11.3	-0.024604	-0.013462
91	11.4	-0.023299	-0.012637
92	11.5	-0.022074	-0.011871
93	11.6	-0.020922	-0.011163
94	11.7	-0.019839	-0.010488
95	11.8	-0.018821	-0.009881

96	11.9	-0.017862	-0.009302
97	12	-0.016959	-0.008761
98	12.1	-0.016108	-0.008258
99	12.2	-0.015305	-0.007795
100	12.3	-0.014547	-0.007354
101	12.4	-0.013833	-0.006942
102	12.5	-0.013158	-0.006559
103	12.6	-0.01252	-0.006202
104	12.7	-0.011916	-0.005866
105	12.8	-0.011346	-0.005546
106	12.9	-0.010806	-0.005249
107	13	-0.010294	-0.004984
108	13.1	-0.009809	-0.004725
109	13.2	-0.009348	-0.004482
110	13.3	-0.008912	-0.004252
111	13.4	-0.008497	-0.00404
112	13.5	-0.008103	-0.003838
113	13.6	-0.007729	-0.003651
114	13.7	-0.007373	-0.003472
115	13.8	-0.007034	-0.003309
116	13.9	-0.006711	-0.003152
117	14	-0.006403	-0.003002
118	14.1	-0.00611	-0.002859
119	14.2	-0.005831	-0.002724
120	14.3	-0.005564	-0.0026

121	14.4	-0.00531	-0.002481
122	14.5	-0.005068	-0.002375
123	14.6	-0.004835	-0.00227
124	14.7	-0.004613	-0.002168
125	14.8	-0.004401	-0.002079
126	14.9	-0.004198	-0.001988
127	15	-0.004003	-0.001901
128	15.1	-0.003817	-0.001824
129	15.2	-0.003638	-0.001749
130	15.3	-0.003467	-0.00168
131	15.4	-0.003303	-0.001612
132	15.5	-0.003144	-0.001549
133	15.6	-0.002993	-0.001488
134	15.7	-0.002847	-0.001431
135	15.8	-0.002706	-0.001377
136	15.9	-0.002571	-0.001323
137	16	-0.002441	-0.001277
138	16.1	-0.002316	-0.001231
139	16.2	-0.002195	-0.001187
140	16.3	-0.002078	-0.001143
141	16.4	-0.001966	-0.001104
142	16.5	-0.001858	-0.001066
143	16.6	-0.001753	-0.00103
144	16.7	-0.001652	-0.000994
145	16.8	-0.001554	-0.000961

146	16.9	-0.001459	-0.000932
147	17	-0.001368	-0.000902
148	17.1	-0.001279	-0.000874
149	17.2	-0.001193	-0.000847
150	17.3	-0.001109	-0.000819
151	17.4	-0.001029	-0.000796
152	17.5	-0.00095	-0.000772
153	17.6	-0.000874	-0.00075
154	17.7	-0.0008	-0.000729
155	17.8	-0.000728	-0.000707
156	17.9	-0.000659	-0.000688
157	18	-0.000591	-0.000669
158	18.1	-0.000525	-0.000652
159	18.2	-0.000461	-0.000633
160	18.3	-0.000398	-0.000618
161	18.4	-0.000337	-0.000602
162	18.5	-0.000277	-0.000587
163	18.6	-0.000219	-0.000574
164	18.7	-0.000163	-0.000561
165	18.8	-0.000107	-0.000548
166	18.9	-0.000053	-0.000536
167	19	0	-0.000524
MORSE_mpdtape3			
N	166		

1	0.2	675160789.9	438079554.8
2	1.6	61849413.18	438079554.8
3	1.7	31122968.98	176449329.2
4	2	2343397.257	15414482.28
5	2.1	1201930.322	7414856.421
6	2.2	646683.4013	3690082
7	2.4	177548.3863	1001268.151
8	2.5	100343.7884	542823.8081
9	2.6	58107.22159	301907.5273
10	2.7	34435.40506	171528.8033
11	2.8	20874.85412	99682.21543
12	2.9	12943.57785	58943.30994
13	3	8214.125517	35645.7367
14	3.1	5331.49429	22006.88783
15	3.3	2276.581208	8542.242989
16	3.4	1562.374083	5741.899516
17	3.5	1088.620443	3733.173284
18	3.6	777.234125	2494.553074
19	3.7	567.048909	1709.151246
20	4	221.073276	597.352977
21	4.5	43.208217	114.10726
22	4.6	33.02347	89.587667
23	4.7	24.983999	71.201763
24	4.8	19.273933	42.999546

25	4.9	15.180085	38.877411
26	5	11.193269	40.858921
27	5.1	7.788157	27.243322
28	5.2	5.377641	20.966998
29	5.3	3.635056	13.884699
30	5.4	2.431917	10.178075
31	5.5	1.570046	7.059355
32	5.6	0.958071	5.180146
33	5.7	0.511653	3.7482
34	5.8	0.177617	2.932531
35	5.9	-0.097592	2.571643
36	6	-0.321872	1.91396
37	6.1	-0.506665	1.781901
38	6.2	-0.662633	1.33745
39	6.3	-0.785581	1.121525
40	6.4	-0.884401	0.854861
41	6.5	-0.955296	0.56304
42	6.6	-0.999947	0.329992
43	6.7	-1.023091	0.132886
44	6.8	-1.029058	-0.013545
45	6.9	-1.022343	-0.120755
46	7	-1.007188	-0.182348
47	7.1	-0.986474	-0.231938
48	7.2	-0.962299	-0.25156
49	7.3	-0.935948	-0.275462

50	7.4	-0.90757	-0.29209
51	7.5	-0.878063	-0.298048
52	7.6	-0.847794	-0.307344
53	7.7	-0.817143	-0.305679
54	7.8	-0.786787	-0.301441
55	7.9	-0.756867	-0.296963
56	8	-0.72782	-0.283967
57	8.1	-0.70004	-0.271632
58	8.2	-0.673484	-0.259486
59	8.3	-0.647498	-0.260238
60	8.4	-0.621163	-0.266455
61	8.5	-0.594355	-0.269717
62	8.6	-0.567172	-0.27393
63	8.7	-0.539081	-0.2879
64	8.8	-0.509923	-0.295252
65	8.9	-0.480247	-0.298282
66	9	-0.450331	-0.300036
67	9.1	-0.42033	-0.299978
68	9.2	-0.390622	-0.294172
69	9.3	-0.361669	-0.284906
70	9.4	-0.33375	-0.273455
71	9.5	-0.307132	-0.258906
72	9.6	-0.282058	-0.242585
73	9.7	-0.258641	-0.225749
74	9.8	-0.236895	-0.20918

75	9.9	-0.216843	-0.191863
76	10	-0.198444	-0.176118
77	10.1	-0.181522	-0.162323
78	10.2	-0.16596	-0.148911
79	10.3	-0.151733	-0.135622
80	10.4	-0.138753	-0.123976
81	10.5	-0.126874	-0.113605
82	10.6	-0.115994	-0.103994
83	10.7	-0.106048	-0.094926
84	10.8	-0.096945	-0.087147
85	10.9	-0.088593	-0.079893
86	11	-0.08091	-0.073754
87	11.1	-0.073899	-0.066478
88	11.2	-0.067484	-0.061806
89	11.3	-0.061566	-0.056563
90	11.4	-0.056171	-0.051337
91	11.5	-0.051227	-0.047543
92	11.6	-0.046681	-0.043368
93	11.7	-0.042501	-0.040237
94	11.8	-0.038644	-0.036904
95	11.9	-0.03509	-0.034178
96	12	-0.031823	-0.031163
97	12.1	-0.028847	-0.028365
98	12.2	-0.026116	-0.026259
99	12.3	-0.023598	-0.024097

100	12.4	-0.021292	-0.02202
101	12.5	-0.019175	-0.020312
102	12.6	-0.017225	-0.018691
103	12.7	-0.015444	-0.01694
104	12.8	-0.013815	-0.015629
105	12.9	-0.012319	-0.014293
106	13	-0.010959	-0.012899
107	13.1	-0.009724	-0.011806
108	13.2	-0.008594	-0.010805
109	13.3	-0.007563	-0.009814
110	13.4	-0.006627	-0.0089
111	13.5	-0.005777	-0.008099
112	13.6	-0.005003	-0.007386
113	13.7	-0.004302	-0.006636
114	13.8	-0.003674	-0.005923
115	13.9	-0.003115	-0.005256
116	14	-0.002616	-0.004722
117	14.1	-0.002172	-0.004148
118	14.2	-0.001781	-0.003683
119	14.3	-0.001435	-0.003226
120	14.4	-0.001135	-0.002786
121	14.5	-0.000877	-0.002361
122	14.6	-0.000658	-0.002028
123	14.7	-0.000474	-0.001649
124	14.8	-0.000324	-0.001345

125	14.9	-0.000205	-0.001037
126	15	-0.000115	-0.000759
127	15.1	-0.000053	-0.000497
128	15.2	-0.000016	-0.000244
129	15.3	-0.000002	-0.000035
130	15.4	0	0
131	15.5	0	0
132	15.6	0	0
133	15.7	0	0
134	15.8	0	0
135	15.9	0	0
136	16	0	0
137	16.1	0	0
138	16.2	0	0
139	16.3	0	0
140	16.4	0	0
141	16.5	0	0
142	16.6	0	0
143	16.7	0	0
144	16.8	0	0
145	16.9	0	0
146	17	0	0
147	17.1	0	0
148	17.2	0	0
149	17.3	0	0

150	17.4	0	0
151	17.5	0	0
152	17.6	0	0
153	17.7	0	0
154	17.8	0	0
155	17.9	0	0
156	18	0	0
157	18.1	0	0
158	18.2	0	0
159	18.3	0	0
160	18.4	0	0
161	18.5	0	0
162	18.6	0	0
163	18.7	0	0
164	18.8	0	0
165	18.9	0	0
166	19	0	0
MORSE_o			
N	190		
1	1	42959.86973	363952.0675
2	1.1	17800.10488	139243.2294
3	1.2	7948.485791	57789.15246
4	1.3	3775.605755	25668.44825

5	1.4	1888.47713	12074.12425
6	1.5	986.556139	5964.295573
7	1.6	534.703572	3072.755757
8	1.7	298.980917	1641.697349
9	1.8	171.63319	905.257203
10	1.9	100.717244	513.0617
11	2	60.174365	297.795885
12	2.1	36.462193	176.447558
13	2.2	22.319398	106.408343
14	2.3	13.742419	65.131231
15	2.4	8.468237	40.352414
16	2.5	5.188836	25.235606
17	2.6	3.132893	15.883262
18	2.7	1.837355	10.02749
19	2.8	1.019755	6.324514
20	2.9	0.505307	3.964435
21	3	0.184495	2.451803
22	3.1	-0.01206	1.47931
23	3.2	-0.128726	0.854009
24	3.3	-0.194096	0.45338
25	3.4	-0.226708	0.198874
26	3.5	-0.238637	0.039702
27	3.6	-0.237761	-0.057223
28	3.7	-0.229221	-0.113582
29	3.8	-0.216359	-0.143645

30	3.9	-0.201335	-0.156838
31	4	-0.185524	-0.159392
32	4.1	-0.169783	-0.155425
33	4.2	-0.154629	-0.147646
34	4.3	-0.140356	-0.137829
35	4.4	-0.127108	-0.127119
36	4.5	-0.11494	-0.116244
37	4.6	-0.103845	-0.105653
38	4.7	-0.093782	-0.095608
39	4.8	-0.084689	-0.086248
40	4.9	-0.076495	-0.077636
41	5	-0.069124	-0.069782
42	5.1	-0.062502	-0.062667
43	5.2	-0.056556	-0.05625
44	5.3	-0.051219	-0.050482
45	5.4	-0.04643	-0.045311
46	5.5	-0.04213	-0.040684
47	5.6	-0.038268	-0.036547
48	5.7	-0.034798	-0.032851
49	5.8	-0.031678	-0.029551
50	5.9	-0.028871	-0.026603
51	6	-0.026342	-0.023972
52	6.1	-0.024062	-0.021619
53	6.2	-0.022005	-0.019518
54	6.3	-0.020148	-0.017638

55	6.4	-0.018468	-0.015955
56	6.5	-0.016948	-0.014448
57	6.6	-0.015571	-0.013096
58	6.7	-0.014322	-0.011884
59	6.8	-0.013188	-0.010796
60	6.9	-0.012157	-0.009817
61	7	-0.011219	-0.008937
62	7.1	-0.010365	-0.008144
63	7.2	-0.009587	-0.007428
64	7.3	-0.008876	-0.006784
65	7.4	-0.008227	-0.006201
66	7.5	-0.007633	-0.005674
67	7.6	-0.00709	-0.005197
68	7.7	-0.006591	-0.004764
69	7.8	-0.006135	-0.004372
70	7.9	-0.005715	-0.004017
71	8	-0.00533	-0.003694
72	8.1	-0.004975	-0.003399
73	8.2	-0.004648	-0.003132
74	8.3	-0.004348	-0.002888
75	8.4	-0.00407	-0.002665
76	8.5	-0.003814	-0.002462
77	8.6	-0.003577	-0.002276
78	8.7	-0.003358	-0.002106
79	8.8	-0.003155	-0.00195

80	8.9	-0.002967	-0.001808
81	9	-0.002792	-0.001678
82	9.1	-0.002631	-0.001557
83	9.2	-0.00248	-0.001448
84	9.3	-0.002341	-0.001346
85	9.4	-0.002211	-0.001253
86	9.5	-0.00209	-0.001167
87	9.6	-0.001977	-0.001088
88	9.7	-0.001872	-0.001015
89	9.8	-0.001774	-0.000948
90	9.9	-0.001682	-0.000886
91	10	-0.001596	-0.000829
92	10.1	-0.001516	-0.000776
93	10.2	-0.001441	-0.000727
94	10.3	-0.001371	-0.00068
95	10.4	-0.001305	-0.000638
96	10.5	-0.001243	-0.0006
97	10.6	-0.001185	-0.000563
98	10.7	-0.00113	-0.000529
99	10.8	-0.001079	-0.000498
100	10.9	-0.00103	-0.000469
101	11	-0.000985	-0.000442
102	11.1	-0.000942	-0.000417
103	11.2	-0.000901	-0.000394
104	11.3	-0.000863	-0.000372

105	11.4	-0.000827	-0.000351
106	11.5	-0.000793	-0.000332
107	11.6	-0.00076	-0.000314
108	11.7	-0.00073	-0.000297
109	11.8	-0.000701	-0.000282
110	11.9	-0.000674	-0.000267
111	12	-0.000648	-0.000253
112	12.1	-0.000623	-0.000241
113	12.2	-0.000599	-0.000229
114	12.3	-0.000577	-0.000217
115	12.4	-0.000556	-0.000207
116	12.5	-0.000536	-0.000198
117	12.6	-0.000516	-0.000188
118	12.7	-0.000498	-0.000179
119	12.8	-0.00048	-0.000171
120	12.9	-0.000464	-0.000163
121	13	-0.000448	-0.000157
122	13.1	-0.000432	-0.00015
123	13.2	-0.000418	-0.000144
124	13.3	-0.000404	-0.000137
125	13.4	-0.00039	-0.000132
126	13.5	-0.000377	-0.000126
127	13.6	-0.000365	-0.000121
128	13.7	-0.000353	-0.000117
129	13.8	-0.000341	-0.000112

130	13.9	-0.00033	-0.000108
131	14	-0.00032	-0.000104
132	14.1	-0.00031	-0.0001
133	14.2	-0.0003	-0.000096
134	14.3	-0.00029	-0.000094
135	14.4	-0.000281	-0.00009
136	14.5	-0.000272	-0.000087
137	14.6	-0.000264	-0.000084
138	14.7	-0.000255	-0.000082
139	14.8	-0.000247	-0.000079
140	14.9	-0.000239	-0.000077
141	15	-0.000232	-0.000074
142	15.1	-0.000224	-0.000073
143	15.2	-0.000217	-0.00007
144	15.3	-0.00021	-0.000068
145	15.4	-0.000204	-0.000067
146	15.5	-0.000197	-0.000065
147	15.6	-0.000191	-0.000064
148	15.7	-0.000184	-0.000062
149	15.8	-0.000178	-0.00006
150	15.9	-0.000172	-0.000058
151	16	-0.000167	-0.000057
152	16.1	-0.000161	-0.000056
153	16.2	-0.000155	-0.000055
154	16.3	-0.00015	-0.000054

155	16.4	-0.000145	-0.000053
156	16.5	-0.000139	-0.000052
157	16.6	-0.000134	-0.000051
158	16.7	-0.000129	-0.000049
159	16.8	-0.000124	-0.000049
160	16.9	-0.000119	-0.000048
161	17	-0.000115	-0.000048
162	17.1	-0.00011	-0.000047
163	17.2	-0.000105	-0.000046
164	17.3	-0.000101	-0.000045
165	17.4	-0.000096	-0.000044
166	17.5	-0.000092	-0.000044
167	17.6	-0.000088	-0.000043
168	17.7	-0.000083	-0.000042
169	17.8	-0.000079	-0.000042
170	17.9	-0.000075	-0.000041
171	18	-0.000071	-0.000041
172	18.1	-0.000067	-0.00004
173	18.2	-0.000063	-0.00004
174	18.3	-0.000059	-0.000039
175	18.4	-0.000055	-0.000039
176	18.5	-0.000051	-0.000039
177	18.6	-0.000047	-0.000038
178	18.7	-0.000043	-0.000038
179	18.8	-0.00004	-0.000037

180	18.9	-0.000036	-0.000037
181	19	-0.000032	-0.000037
182	19.1	-0.000028	-0.000037
183	19.2	-0.000025	-0.000036
184	19.3	-0.000021	-0.000036
185	19.4	-0.000017	-0.000036
186	19.5	-0.000014	-0.000035
187	19.6	-0.00001	-0.000035
188	19.7	-0.000007	-0.000035
189	19.8	-0.000003	-0.000034
190	19.9	0	-0.000035
MORSE_obpada3			
N	176		
1	0.2	553894556	363933850
2	1.6	44387166.08	363933850
3	1.7	18861249.23	146584487.1
4	1.8	8422420.607	62192085.39
5	1.9	3930855.951	27639207.72
6	2	1908642.128	12805068.75
7	2.1	960411.0742	6159552.331
8	2.2	499154.5139	3065578.875
9	2.3	267183.2327	1573846.749
10	2.4	146925.5753	831306.3979

11	2.5	82824.91163	450706.8757
12	2.6	47773.79163	250315.5243
13	2.7	28150.09859	142158.3366
14	2.8	16918.7816	82468.00307
15	2.9	10353.37406	48840.14773
16	3	6443.971664	29347.90021
17	3.1	4077.700607	17977.52095
18	3.2	2619.906214	11178.3669
19	3.3	1708.073486	7058.28767
20	3.4	1129.026817	4522.645711
21	3.5	755.832278	2941.245056
22	3.6	512.105725	1933.286011
23	3.7	351.008947	1288.649555
24	3.8	243.033881	870.851751
25	3.9	169.986476	590.096353
26	4	120.164375	406.345669
27	4.1	85.706613	282.809577
28	4.2	61.635014	198.622399
29	4.3	44.664029	140.797305
30	4.4	32.556301	101.357243
31	4.5	23.791338	73.942021
32	4.6	17.45525	52.779747
33	4.7	12.906787	38.189514
34	4.8	9.599766	27.950899
35	4.9	7.175807	20.528281

36	5	5.374789	15.492088
37	5.1	4.019263	11.618422
38	5.2	2.995709	8.85267
39	5.3	2.217446	6.712585
40	5.4	1.62332	5.16993
41	5.5	1.166196	3.972558
42	5.6	0.818359	2.984178
43	5.7	0.559045	2.202096
44	5.8	0.36886	1.601598
45	5.9	0.230212	1.171367
46	6	0.128385	0.865174
47	6.1	0.052171	0.659113
48	6.2	-0.006688	0.518061
49	6.3	-0.053355	0.415272
50	6.4	-0.090454	0.326725
51	6.5	-0.119157	0.24733
52	6.6	-0.141425	0.198021
53	6.7	-0.15821	0.137675
54	6.8	-0.169888	0.095902
55	6.9	-0.177424	0.054803
56	7	-0.180759	0.011906
57	7.1	-0.179945	-0.028184
58	7.2	-0.175561	-0.059497
59	7.3	-0.168266	-0.086412
60	7.4	-0.15886	-0.101696

61	7.5	-0.148291	-0.109682
62	7.6	-0.137267	-0.110809
63	7.7	-0.12631	-0.108337
64	7.8	-0.115703	-0.103791
65	7.9	-0.105643	-0.097414
66	8	-0.096228	-0.090877
67	8.1	-0.087496	-0.083767
68	8.2	-0.079452	-0.077116
69	8.3	-0.07208	-0.070328
70	8.4	-0.065364	-0.063994
71	8.5	-0.059233	-0.058614
72	8.6	-0.053643	-0.053192
73	8.7	-0.048538	-0.048903
74	8.8	-0.043874	-0.044385
75	8.9	-0.039653	-0.040036
76	9	-0.035818	-0.036655
77	9.1	-0.032317	-0.03336
78	9.2	-0.029141	-0.030159
79	9.3	-0.026261	-0.02746
80	9.4	-0.023636	-0.025033
81	9.5	-0.021252	-0.022643
82	9.6	-0.01909	-0.020605
83	9.7	-0.017124	-0.018718
84	9.8	-0.015333	-0.017083
85	9.9	-0.013707	-0.015455

86	10	-0.012235	-0.013968
87	10.1	-0.010905	-0.01265
88	10.2	-0.009694	-0.011563
89	10.3	-0.008594	-0.01043
90	10.4	-0.0076	-0.009464
91	10.5	-0.006698	-0.008569
92	10.6	-0.005882	-0.00774
93	10.7	-0.005147	-0.00696
94	10.8	-0.004487	-0.00625
95	10.9	-0.003893	-0.005623
96	11	-0.003357	-0.005092
97	11.1	-0.002877	-0.004508
98	11.2	-0.002451	-0.004013
99	11.3	-0.002071	-0.003585
100	11.4	-0.001734	-0.003166
101	11.5	-0.001435	-0.002802
102	11.6	-0.001174	-0.002433
103	11.7	-0.000946	-0.002111
104	11.8	-0.00075	-0.001817
105	11.9	-0.000581	-0.001553
106	12	-0.000438	-0.001308
107	12.1	-0.000319	-0.001075
108	12.2	-0.000222	-0.000868
109	12.3	-0.000144	-0.000689
110	12.4	-0.000084	-0.000512

111	12.5	-0.000042	-0.000327
112	12.6	-0.000016	-0.000203
113	12.7	-0.000003	-0.000055
114	12.8	0	0
115	12.9	0	0
116	13	0	0
117	13.1	0	0
118	13.2	0	0
119	13.3	0	0
120	13.4	0	0
121	13.5	0	0
122	13.6	0	0
123	13.7	0	0
124	13.8	0	0
125	13.9	0	0
126	14	0	0
127	14.1	0	0
128	14.2	0	0
129	14.3	0	0
130	14.4	0	0
131	14.5	0	0
132	14.6	0	0
133	14.7	0	0
134	14.8	0	0
135	14.9	0	0

136	15	0	0
137	15.1	0	0
138	15.2	0	0
139	15.3	0	0
140	15.4	0	0
141	15.5	0	0
142	15.6	0	0
143	15.7	0	0
144	15.8	0	0
145	15.9	0	0
146	16	0	0
147	16.1	0	0
148	16.2	0	0
149	16.3	0	0
150	16.4	0	0
151	16.5	0	0
152	16.6	0	0
153	16.7	0	0
154	16.8	0	0
155	16.9	0	0
156	17	0	0
157	17.1	0	0
158	17.2	0	0
159	17.3	0	0
160	17.4	0	0

161	17.5	0	0
162	17.6	0	0
163	17.7	0	0
164	17.8	0	0
165	17.9	0	0
166	18	0	0
167	18.1	0	0
168	18.2	0	0
169	18.3	0	0
170	18.4	0	0
171	18.5	0	0
172	18.6	0	0
173	18.7	0	0
174	18.8	0	0
175	18.9	0	0
176	19	0	0
MORSE_ompd			
N	171		
1	0.2	169167.6117	82145.62768
2	2.1	13090.91915	82145.62768
3	2.2	6907.934014	41514.07497
4	2.3	3740.898406	21826.63719
5	2.4	2076.117901	11468.9729

6	2.5	1188.489229	6283.600546
7	2.6	697.738208	3531.419873
8	2.7	419.658291	2030.178471
9	2.8	258.531579	1192.355772
10	2.9	162.914734	719.98112
11	3	104.694422	444.42513
12	3.1	68.300339	283.456535
13	3.2	44.793859	186.673058
14	3.3	29.458544	120.033237
15	3.4	19.556174	78.014163
16	3.5	13.086237	51.384588
17	3.6	8.807604	34.188055
18	3.7	5.947706	23.009906
19	3.8	4.01487	15.646816
20	3.9	2.695152	10.747545
21	4	1.784546	7.464578
22	4.1	1.148363	5.259081
23	4.2	0.696884	3.770505
24	4.3	0.371297	2.741235
25	4.4	0.133036	2.023981
26	4.5	-0.042839	1.493519
27	4.6	-0.172453	1.098758
28	4.7	-0.266896	0.790104
29	4.8	-0.333687	0.545708
30	4.9	-0.377681	0.334188

31	5	-0.402308	0.158352
32	5.1	-0.411429	0.024051
33	5.2	-0.409005	-0.072527
34	5.3	-0.398613	-0.135311
35	5.4	-0.3832	-0.172949
36	5.5	-0.364954	-0.191975
37	5.6	-0.345419	-0.198713
38	5.7	-0.325614	-0.197402
39	5.8	-0.306191	-0.191057
40	5.9	-0.287545	-0.18186
41	6	-0.269895	-0.171141
42	6.1	-0.253348	-0.159785
43	6.2	-0.23794	-0.148387
44	6.3	-0.223652	-0.137376
45	6.4	-0.21044	-0.126855
46	6.5	-0.198248	-0.116996
47	6.6	-0.187005	-0.107861
48	6.7	-0.176642	-0.099398
49	6.8	-0.167091	-0.091626
50	6.9	-0.158284	-0.084507
51	7	-0.150157	-0.078038
52	7.1	-0.142649	-0.072121
53	7.2	-0.135706	-0.066741
54	7.3	-0.129277	-0.061836
55	7.4	-0.123317	-0.057373

56	7.5	-0.117782	-0.053308
57	7.6	-0.112636	-0.049612
58	7.7	-0.107844	-0.046232
59	7.8	-0.103375	-0.043155
60	7.9	-0.0992	-0.040351
61	8	-0.095293	-0.037791
62	8.1	-0.091631	-0.035445
63	8.2	-0.088194	-0.033298
64	8.3	-0.084962	-0.031325
65	8.4	-0.08192	-0.02952
66	8.5	-0.079051	-0.027862
67	8.6	-0.076341	-0.026333
68	8.7	-0.073778	-0.024928
69	8.8	-0.071351	-0.023628
70	8.9	-0.069048	-0.022429
71	9	-0.06686	-0.021319
72	9.1	-0.06478	-0.020292
73	9.2	-0.062798	-0.01934
74	9.3	-0.060908	-0.018457
75	9.4	-0.059104	-0.017633
76	9.5	-0.057379	-0.01687
77	9.6	-0.055727	-0.016155
78	9.7	-0.054145	-0.015489
79	9.8	-0.052627	-0.014867
80	9.9	-0.05117	-0.014285

81	10	-0.049768	-0.013741
82	10.1	-0.04842	-0.01323
83	10.2	-0.047121	-0.01275
84	10.3	-0.045868	-0.012299
85	10.4	-0.04466	-0.011874
86	10.5	-0.043492	-0.011474
87	10.6	-0.042364	-0.011097
88	10.7	-0.041272	-0.010741
89	10.8	-0.040215	-0.010405
90	10.9	-0.03919	-0.010086
91	11	-0.038197	-0.009784
92	11.1	-0.037232	-0.009497
93	11.2	-0.036296	-0.009226
94	11.3	-0.035387	-0.008968
95	11.4	-0.034502	-0.008722
96	11.5	-0.033642	-0.008489
97	11.6	-0.032804	-0.008265
98	11.7	-0.031988	-0.008053
99	11.8	-0.031193	-0.00785
100	11.9	-0.030418	-0.007656
101	12	-0.029661	-0.00747
102	12.1	-0.028923	-0.007293
103	12.2	-0.028202	-0.007123
104	12.3	-0.027498	-0.00696
105	12.4	-0.02681	-0.006804

106	12.5	-0.026137	-0.006654
107	12.6	-0.025479	-0.00651
108	12.7	-0.024835	-0.006371
109	12.8	-0.024204	-0.006238
110	12.9	-0.023587	-0.006109
111	13	-0.022982	-0.005985
112	13.1	-0.02239	-0.005866
113	13.2	-0.021809	-0.005751
114	13.3	-0.021239	-0.00564
115	13.4	-0.020681	-0.005533
116	13.5	-0.020133	-0.005429
117	13.6	-0.019595	-0.005329
118	13.7	-0.019067	-0.005232
119	13.8	-0.018548	-0.005138
120	13.9	-0.018039	-0.005048
121	14	-0.017538	-0.004959
122	14.1	-0.017047	-0.004873
123	14.2	-0.016564	-0.004791
124	14.3	-0.016088	-0.004711
125	14.4	-0.015621	-0.004633
126	14.5	-0.015162	-0.004557
127	14.6	-0.01471	-0.004484
128	14.7	-0.014265	-0.004412
129	14.8	-0.013827	-0.004342
130	14.9	-0.013396	-0.004275

131	15	-0.012972	-0.004208
132	15.1	-0.012555	-0.004144
133	15.2	-0.012143	-0.004082
134	15.3	-0.011738	-0.004022
135	15.4	-0.011339	-0.003962
136	15.5	-0.010946	-0.003904
137	15.6	-0.010558	-0.003847
138	15.7	-0.010176	-0.003793
139	15.8	-0.009799	-0.00374
140	15.9	-0.009428	-0.003687
141	16	-0.009062	-0.003635
142	16.1	-0.008701	-0.003586
143	16.2	-0.008345	-0.003537
144	16.3	-0.007993	-0.00349
145	16.4	-0.007647	-0.003442
146	16.5	-0.007305	-0.003397
147	16.6	-0.006967	-0.003353
148	16.7	-0.006634	-0.003309
149	16.8	-0.006305	-0.003267
150	16.9	-0.005981	-0.003225
151	17	-0.00566	-0.003185
152	17.1	-0.005344	-0.003145
153	17.2	-0.005031	-0.003105
154	17.3	-0.004723	-0.003066
155	17.4	-0.004418	-0.003029

156	17.5	-0.004117	-0.002993
157	17.6	-0.003819	-0.002957
158	17.7	-0.003526	-0.002922
159	17.8	-0.003235	-0.002887
160	17.9	-0.002948	-0.002853
161	18	-0.002665	-0.002819
162	18.1	-0.002384	-0.002786
163	18.2	-0.002107	-0.002754
164	18.3	-0.001833	-0.002724
165	18.4	-0.001562	-0.002693
166	18.5	-0.001295	-0.002663
167	18.6	-0.00103	-0.002632
168	18.7	-0.000768	-0.002603
169	18.8	-0.000509	-0.002575
170	18.9	-0.000253	-0.002546
171	19	0	-0.002519
MORSE_otape3			
N	171		
1	0.2	2380984711	1659271395
2	1.5	223931896.7	1659271395
3	1.6	109457708.1	630212376.5
4	1.9	7746520.787	47862205.75
5	2	4244705.55	22174099

6	2.3	509794.8909	2725305.392
7	2.4	301555.8829	1439474.768
8	2.7	48707.36051	246182.0483
9	2.8	29260.46005	142755.9611
10	2.9	17903.74853	84378.26918
11	3	11145.4186	50788.3295
12	3.1	7051.634776	31087.34692
13	3.2	4530.433305	19336.68249
14	3.3	2953.201762	12207.94838
15	3.4	1952.377823	7808.530388
16	3.5	1308.469198	5069.642112
17	3.6	887.527551	3349.190836
18	3.7	608.875674	2223.846696
19	3.8	422.848074	1496.705309
20	3.9	296.938749	1021.48118
21	4	210.269905	711.895713
22	4.1	149.844652	496.609332
23	4.2	107.729657	345.690584
24	4.3	78.052917	247.844213
25	4.4	56.894381	175.3265
26	4.5	41.724536	128.070397
27	4.6	30.631966	93.781018
28	4.7	22.533861	68.181074
29	4.8	16.61034	50.289348
30	4.9	12.285932	36.198803

31	5	9.136602	26.787793
32	5.1	6.806526	19.81373
33	5.2	5.080624	14.704323
34	5.3	3.796369	10.980767
35	5.4	2.843487	8.076873
36	5.5	2.134042	6.112037
37	5.6	1.599858	4.571628
38	5.7	1.199479	3.435969
39	5.8	0.898688	2.579841
40	5.9	0.672255	1.948815
41	6	0.498479	1.526711
42	6.1	0.361385	1.215166
43	6.2	0.251977	0.973003
44	6.3	0.163425	0.798023
45	6.4	0.09011	0.668279
46	6.5	0.030037	0.533181
47	6.6	-0.01941	0.455764
48	6.7	-0.061801	0.39206
49	6.8	-0.098433	0.340567
50	6.9	-0.129721	0.285204
51	7	-0.155166	0.223694
52	7.1	-0.174842	0.169824
53	7.2	-0.188608	0.105489
54	7.3	-0.196192	0.046189
55	7.4	-0.197875	-0.012521

56	7.5	-0.194251	-0.05995
57	7.6	-0.186601	-0.09305
58	7.7	-0.176315	-0.11267
59	7.8	-0.164607	-0.12149
60	7.9	-0.152326	-0.12413
61	8	-0.140062	-0.121154
62	8.1	-0.128187	-0.116349
63	8.2	-0.116962	-0.108151
64	8.3	-0.106471	-0.101681
65	8.4	-0.096682	-0.094095
66	8.5	-0.087641	-0.086723
67	8.6	-0.079353	-0.079032
68	8.7	-0.071748	-0.073067
69	8.8	-0.064762	-0.066659
70	8.9	-0.058402	-0.060536
71	9	-0.052607	-0.055362
72	9.1	-0.04734	-0.049977
73	9.2	-0.042564	-0.045538
74	9.3	-0.038236	-0.041036
75	9.4	-0.034302	-0.037636
76	9.5	-0.030728	-0.033851
77	9.6	-0.027499	-0.030719
78	9.7	-0.024578	-0.02771
79	9.8	-0.02193	-0.025241
80	9.9	-0.019522	-0.02292

81	10	-0.017347	-0.02059
82	10.1	-0.015378	-0.018785
83	10.2	-0.013589	-0.017002
84	10.3	-0.011977	-0.015221
85	10.4	-0.010529	-0.013753
86	10.5	-0.009221	-0.012398
87	10.6	-0.008047	-0.011091
88	10.7	-0.006992	-0.010013
89	10.8	-0.006042	-0.008972
90	10.9	-0.005193	-0.008014
91	11	-0.004435	-0.007141
92	11.1	-0.003762	-0.006332
93	11.2	-0.003165	-0.005596
94	11.3	-0.002637	-0.004979
95	11.4	-0.002171	-0.004329
96	11.5	-0.001766	-0.003765
97	11.6	-0.001413	-0.003296
98	11.7	-0.001109	-0.002802
99	11.8	-0.00085	-0.002366
100	11.9	-0.000634	-0.001951
101	12	-0.000456	-0.001623
102	12.1	-0.00031	-0.001287
103	12.2	-0.000195	-0.001018
104	12.3	-0.000108	-0.000713
105	12.4	-0.000049	-0.000466

106	12.5	-0.000014	-0.000237
107	12.6	-0.000001	-0.000023
108	12.7	0	0
109	12.8	0	0
110	12.9	0	0
111	13	0	0
112	13.1	0	0
113	13.2	0	0
114	13.3	0	0
115	13.4	0	0
116	13.5	0	0
117	13.6	0	0
118	13.7	0	0
119	13.8	0	0
120	13.9	0	0
121	14	0	0
122	14.1	0	0
123	14.2	0	0
124	14.3	0	0
125	14.4	0	0
126	14.5	0	0
127	14.6	0	0
128	14.7	0	0
129	14.8	0	0
130	14.9	0	0

131	15	0	0
132	15.1	0	0
133	15.2	0	0
134	15.3	0	0
135	15.4	0	0
136	15.5	0	0
137	15.6	0	0
138	15.7	0	0
139	15.8	0	0
140	15.9	0	0
141	16	0	0
142	16.1	0	0
143	16.2	0	0
144	16.3	0	0
145	16.4	0	0
146	16.5	0	0
147	16.6	0	0
148	16.7	0	0
149	16.8	0	0
150	16.9	0	0
151	17	0	0
152	17.1	0	0
153	17.2	0	0
154	17.3	0	0
155	17.4	0	0

156	17.5	0	0
157	17.6	0	0
158	17.7	0	0
159	17.8	0	0
160	17.9	0	0
161	18	0	0
162	18.1	0	0
163	18.2	0	0
164	18.3	0	0
165	18.4	0	0
166	18.5	0	0
167	18.6	0	0
168	18.7	0	0
169	18.8	0	0
170	18.9	0	0
171	19	0	0
MORSE_tape3			
N	168		
1	0.2	21207222294	14360573521
2	1.5	2538476717	14360573521
3	1.7	882731062.9	2196883016
4	2.5	1276760.096	6752741.032
5	2.6	751630.6694	3749847.492

6	2.8	253217.5703	1234283.5
7	2.9	155040.8622	729250.663
8	3	96643.89637	438688.6527
9	3.1	61278.50807	268619.1132
10	3.2	39493.30665	167084.9152
11	3.3	25858.95066	105602.2046
12	3.4	17189.96343	67777.53999
13	3.5	11609.05407	43840.64725
14	3.6	7971.97725	28900.88915
15	3.7	5554.107451	19456.50682
16	3.8	3921.071471	13204.21279
17	3.9	2807.184206	9073.532496
18	4	2035.045887	6369.233882
19	4.1	1491.278252	4506.118823
20	4.2	1102.122701	3276.992198
21	4.3	821.038257	2344.69667
22	4.4	616.89203	1738.227879
23	4.5	468.179401	1236.0247
24	4.6	359.049845	946.566429
25	4.7	278.461567	665.199128
26	4.8	221.297334	478.085523
27	4.9	162.935213	689.156912
28	5	114.582591	277.89552
29	5.1	89.331096	227.134387
30	5.2	69.318698	173.113561

31	5.3	53.979552	133.669371
32	5.4	41.973188	106.457897
33	5.5	32.690069	79.204496
34	5.6	24.630305	81.99077
35	5.7	17.631821	57.978911
36	5.8	12.728827	40.080978
37	5.9	9.170543	31.084689
38	6	6.398989	24.346392
39	6.1	4.285614	17.921121
40	6.2	2.703281	13.72553
41	6.3	1.495579	10.428505
42	6.4	0.580855	7.865987
43	6.5	-0.105276	5.856635
44	6.6	-0.639809	4.83402
45	6.7	-1.080698	3.983768
46	6.8	-1.454325	3.488771
47	6.9	-1.754667	2.518057
48	7	-1.986602	2.120642
49	7.1	-2.164771	1.442744
50	7.2	-2.288838	1.038598
51	7.3	-2.37416	0.667837
52	7.4	-2.42461	0.341176
53	7.5	-2.450089	0.1684
54	7.6	-2.457425	-0.021684
55	7.7	-2.448982	-0.147176

56	7.8	-2.430386	-0.224749
57	7.9	-2.403345	-0.316064
58	8	-2.369454	-0.361759
59	8.1	-2.329508	-0.43715
60	8.2	-2.28219	-0.509226
61	8.3	-2.229455	-0.545471
62	8.4	-2.173374	-0.576143
63	8.5	-2.113623	-0.618871
64	8.6	-2.050548	-0.642633
65	8.7	-1.986071	-0.646919
66	8.8	-1.920604	-0.662416
67	8.9	-1.853891	-0.671847
68	9	-1.785482	-0.696331
69	9.1	-1.715551	-0.70229
70	9.2	-1.645171	-0.705297
71	9.3	-1.574968	-0.698772
72	9.4	-1.505101	-0.69856
73	9.5	-1.435546	-0.692537
74	9.6	-1.366513	-0.688134
75	9.7	-1.298126	-0.679613
76	9.8	-1.231344	-0.656024
77	9.9	-1.166443	-0.641986
78	10	-1.103638	-0.61411
79	10.1	-1.043547	-0.587719
80	10.2	-0.985874	-0.565741

81	10.3	-0.931064	-0.530466
82	10.4	-0.879714	-0.49652
83	10.5	-0.831737	-0.463023
84	10.6	-0.786486	-0.442
85	10.7	-0.743836	-0.410994
86	10.8	-0.703715	-0.391438
87	10.9	-0.665505	-0.37276
88	11	-0.628955	-0.358234
89	11.1	-0.594011	-0.340647
90	11.2	-0.560469	-0.330206
91	11.3	-0.528181	-0.315541
92	11.4	-0.497113	-0.305821
93	11.5	-0.466932	-0.297792
94	11.6	-0.437701	-0.286841
95	11.7	-0.409668	-0.27381
96	11.8	-0.383067	-0.258213
97	11.9	-0.357778	-0.247561
98	12	-0.333977	-0.228477
99	12.1	-0.311835	-0.21435
100	12.2	-0.291119	-0.199984
101	12.3	-0.271918	-0.184023
102	12.4	-0.254196	-0.170414
103	12.5	-0.23772	-0.15912
104	12.6	-0.222463	-0.146014
105	12.7	-0.208357	-0.136105

106	12.8	-0.195181	-0.12742
107	12.9	-0.182911	-0.117978
108	13	-0.171513	-0.109976
109	13.1	-0.160963	-0.101034
110	13.2	-0.151169	-0.094842
111	13.3	-0.141987	-0.088784
112	13.4	-0.133399	-0.082975
113	13.5	-0.125385	-0.077316
114	13.6	-0.117882	-0.072738
115	13.7	-0.110882	-0.06726
116	13.8	-0.104376	-0.062872
117	13.9	-0.098282	-0.058992
118	14	-0.092553	-0.055597
119	14.1	-0.087163	-0.052208
120	14.2	-0.082118	-0.048682
121	14.3	-0.077394	-0.045799
122	14.4	-0.072948	-0.043116
123	14.5	-0.068764	-0.040576
124	14.6	-0.064819	-0.038323
125	14.7	-0.061095	-0.036158
126	14.8	-0.057585	-0.034038
127	14.9	-0.054277	-0.032124
128	15	-0.051154	-0.030333
129	15.1	-0.048201	-0.028736
130	15.2	-0.045413	-0.027026

131	15.3	-0.042787	-0.025486
132	15.4	-0.040302	-0.024214
133	15.5	-0.037943	-0.022974
134	15.6	-0.035709	-0.021699
135	15.7	-0.033593	-0.020619
136	15.8	-0.031589	-0.019467
137	15.9	-0.029686	-0.018585
138	16	-0.027877	-0.01759
139	16.1	-0.026161	-0.016745
140	16.2	-0.02453	-0.015864
141	16.3	-0.022982	-0.015106
142	16.4	-0.02151	-0.014324
143	16.5	-0.020112	-0.013636
144	16.6	-0.018781	-0.012995
145	16.7	-0.017514	-0.012349
146	16.8	-0.016309	-0.01175
147	16.9	-0.01516	-0.011216
148	17	-0.014063	-0.010722
149	17.1	-0.013019	-0.010163
150	17.2	-0.012025	-0.009712
151	17.3	-0.011075	-0.009295
152	17.4	-0.010168	-0.008849
153	17.5	-0.009302	-0.008467
154	17.6	-0.008474	-0.008093
155	17.7	-0.007682	-0.007742

156	17.8	-0.006925	-0.007395
157	17.9	-0.006202	-0.00708
158	18	-0.005509	-0.006774
159	18.1	-0.004845	-0.006497
160	18.2	-0.00421	-0.006203
161	18.3	-0.003603	-0.005934
162	18.4	-0.003022	-0.005694
163	18.5	-0.002465	-0.005452
164	18.6	-0.00193	-0.005244
165	18.7	-0.001417	-0.005019
166	18.8	-0.000925	-0.004823
167	18.9	-0.000453	-0.004617
168	19	0	-0.004437

Appendix C

Definition of Polymer Features

The definition of all features (i.e., descriptors) of polyimides are shown in Table C.1.

Table C.1: Definition of all features.

Constitutional descriptors	
Constitutional descriptors are OD-descriptors, independent from molecular connectivity and conformations, Atom and bond counts, molecular weight sum of atomic properties, ets.	
nDB	number of double bonds
PJI2	2D Petitjean shape index
Topological descriptors	
Topological descriptors are molecular descriptors obtained from molecular graph (usually H-depleted), i.e. 2D-descriptors conformationally independent.	
TNN	sum of topological distances between N..N
TOO	sum of topological distances between O..O
TOS	sum of topological distances between O..S
Walk and path counts	
Walk and path counts are molecular descriptors obtained from the molecular graph, counting paths, walks and self-returning walks of different lengths.	
SRW09	self-returning walk count of order 9
Information indices	

Information indices are molecular descriptors calculated as information content of molecules, based on the calculation of equivalence classes from the molecular graph. Among them, the indices of neighborhood symmetry take into account also neighbor degree and edge multiplicity.	
IC2	Information Content index (neighborhood symmetry of 3-order)
IC3	Information Content index (neighborhood symmetry of 3-order)
BIC3	Bond Information Content index (neighborhood symmetry of 3-order)
2D autocorrelations	
Molecular descriptors calculated from molecular graph by summing the products of atom weights of the terminal atoms of all the paths of the considered path length (the lag). 2D autocorrelations by Moreau-Broto (ATS), Moran (MATS) and Geary (GATS) algorithms are calculated from lag 1 to lag 8 for 4 different weighting schemes.	
MATS1m	Moran autocorrelation of lag 1 weighted by mass
MATS4m	Moran autocorrelation of lag 4 weighted by mass
MATS8m	Moran autocorrelation of lag 8 weighted by mass

MATS4v	Moran autocorrelation of lag 4 weighted by van der Waals volume
MATS6v	Moran autocorrelation of lag 6 weighted by van der Waals volume
MATS8v	Moran autocorrelation of lag 8 weighted by van der Waals volume
MATS4e	Moran autocorrelation of lag 4 weighted by Sanderson electronegativity
MATS4p	Moran autocorrelation of lag 4 weighted by polarizability
MATS6p	Moran autocorrelation of lag 6 weighted by polarizability
MATS7p	Moran autocorrelation of lag 7 weighted by polarizability
GATS4m	Geary autocorrelation of lag 4 weighted by mass
GATS5m	Geary autocorrelation of lag 5 weighted by mass
GATS8m	Geary autocorrelation of lag 8 weighted by mass
GATS3e	Geary autocorrelation of lag 3 weighted by Sanderson electronegativity
GATS4e	Geary autocorrelation of lag 4 weighted by Sanderson electronegativity

GATS8e	Geary autocorrelation of lag 8 weighted by Sanderson electronegativity
GATS4p	Geary autocorrelation of lag 4 weighted by polarizability
GATS5p	Geary autocorrelation of lag 5 weighted by polarizability
BCUT descriptors	
BCUT descriptors are molecular descriptors obtained from the positive and negative eigenvalues of the adjacency matrix, weighting the diagonal elements with atom weights.	
BELm4	lowest eigenvalue n. 4 of Burden matrix / weighted by atomic masses
BELv5	lowest eigenvalue n. 5 of Burden matrix / weighted by atomic van der Waals volumes
BELp5	lowest eigenvalue n. 5 of Burden matrix / weighted by atomic polarizabilities
Topological charge indices	
First 10 eigenvalues (absolute values) obtained from a corrected adjacency matrix, i.e. diagonal elements correspond to	
GGI8	topological charge index of order 8

GGI9	topological charge index of order 9
GGI10	topological charge index of order 10
JGI1	mean topological charge index of order 1
JGI2	mean topological charge index of order 2
JGI3	mean topological charge index of order 3
JGI4	mean topological charge index of order 4
JGI7	mean topological charge index of order 7
JGI8	mean topological charge index of order 8
JGI9	mean topological charge index of order 9
JGI10	mean topological charge index of order 10
Eigenvalue-based indices	
Topological descriptors calculated by the eigenvalues of a square (usually symmetric) matrix representing a molecular graph.	
VEA1	eigenvector coefficient sum from adjacency matrix
Geometrical descriptors	
Geometrical descriptors are different kinds of conformationally dependent descriptors based on the molecular geometry. Reliable values are obtained if reliable conformations were previously calculated.	
SPAM	average span R

MEcc	molecular eccentricity
FDI	folding degree index
PJI3	3D Petitjean shape index
LBw	length-to-breadth ratio by WHIM
HOMA	Harmonic Oscillator Model of Aromaticity index
RCI	3D Petitjean shape index
AROM	aromaticity index
DISPv	displacement value / weighted by van der Waals volume
GNN	sum of geometrical distances between N..N
GNF	sum of geometrical distances between N..F
GOS	sum of geometrical distances between O..S
RDF Descriptors	
RDF Descriptors are Molecular descriptors obtained by radial basis functions centered on different interatomic distances (from 0.5A to 15.5A).	
RDF110u	Radial Distribution Function - 110 / unweighted
RDF135u	Radial Distribution Function - 135 / unweighted
RDF140u	Radial Distribution Function - 140 / unweighted
RDF145u	Radial Distribution Function - 145 / unweighted
RDF045m	Radial Distribution Function - 045 / weighted by mass
RDF055m	Radial Distribution Function - 055 / weighted by mass
RDF075m	Radial Distribution Function - 075 / weighted by mass
RDF080m	Radial Distribution Function - 080 / weighted by mass

RDF085m	Radial Distribution Function - 085 / weighted by mass
RDF090m	Radial Distribution Function - 090 / weighted by mass
RDF095m	Radial Distribution Function - 095 / weighted by mass
RDF100m	Radial Distribution Function - 100 / weighted by mass
RDF105m	Radial Distribution Function - 105 / weighted by mass
RDF125m	Radial Distribution Function - 125 / weighted by mass
RDF130m	Radial Distribution Function - 130 / weighted by mass
RDF135m	Radial Distribution Function - 135 / weighted by mass
RDF145m	Radial Distribution Function - 145 / weighted by mass
RDF150m	Radial Distribution Function - 150 / weighted by mass
RDF065v	Radial Distribution Function - 065 / weighted by van der Waals volume
RDF135v	Radial Distribution Function - 135 / weighted by van der Waals volume
RDF150v	Radial Distribution Function - 150 / weighted by van der Waals volume
RDF155v	Radial Distribution Function - 155 / weighted by van der Waals volume
RDF090e	Radial Distribution Function - 090 / weighted by Sanderson electronegativity
RDF120e	Radial Distribution Function - 120 / weighted by Sanderson electronegativity
RDF145e	Radial Distribution Function - 145 / weighted by Sanderson electronegativity

RDF135p	Radial Distribution Function - 135 / weighted by polarizability
RDF150p	Radial Distribution Function - 150 / weighted by polarizability
RDF155p	Radial Distribution Function - 155 / weighted by polarizability
3D-MoRSE descriptors	
3D-MoRSE is a very flexible 3D structure encoding framework for chemoinformatics and QSAR purposes due to the range of scattering parameter values and variety of weighting schemes used.	
Mor02u	signal 02 / unweighted
Mor04u	signal 04 / unweighted
Mor05u	signal 05 / unweighted
Mor08u	signal 08 / unweighted
Mor09u	signal 09 / unweighted
Mor12u	signal 12 / unweighted
Mor13u	signal 13 / unweighted
Mor16u	signal 16 / unweighted
Mor17u	signal 17 / unweighted
Mor21u	signal 21 / unweighted
Mor22u	signal 22 / unweighted
Mor24u	signal 24 / unweighted

Mor25u	signal 25 / unweighted
Mor28u	signal 28 / unweighted
Mor03m	signal 03 / weighted by mass
Mor11m	signal 11 / weighted by mass
Mor16m	signal 16 / weighted by mass
Mor23m	signal 23 / weighted by mass
Mor25m	signal 25 / weighted by mass
Mor26m	signal 26 / weighted by mass
Mor28m	signal 28 / weighted by mass
Mor29m	signal 29 / weighted by mass
Mor03v	signal 03 / weighted by van der Waals volume
Mor06v	signal 06 / weighted by van der Waals volume
Mor17v	signal 17 / weighted by van der Waals volume
Mor21v	signal 21 / weighted by van der Waals volume
Mor22v	signal 22 / weighted by van der Waals volume
Mor24v	signal 24 / weighted by van der Waals volume
Mor28v	signal 28 / weighted by van der Waals volume
Mor29v	signal 29 / weighted by van der Waals volume
Mor30v	signal 30 / weighted by van der Waals volume
Mor02e	signal 02 / weighted by van der Waals volume
Mor03e	signal 03 / weighted by van der Waals volume
Mor08e	signal 08 / weighted by van der Waals volume
Mor11e	signal 11 / weighted by van der Waals volume
Mor14e	signal 14 / weighted by van der Waals volume

Mor15e	signal 15 / weighted by van der Waals volume
Mor17e	signal 17 / weighted by van der Waals volume
Mor18e	signal 18 / weighted by van der Waals volume
Mor21e	signal 21 / weighted by van der Waals volume
Mor25e	signal 25 / weighted by van der Waals volume
Mor30e	signal 30 / weighted by van der Waals volume
Mor03p	signal 03 / weighted by polarizability
Mor04p	signal 04 / weighted by polarizability
Mor10p	signal 10 / weighted by polarizability
Mor17p	signal 17 / weighted by polarizability
Mor24p	signal 24 / weighted by polarizability
Mor29p	signal 29 / weighted by polarizability
WHIM descriptors	
WHIM descriptors are molecular descriptors obtained as statistical indices of the atoms projected onto the 3 principal components obtained from weighted covariance matrices of the atomic coordinates.	
G1u	1st component symmetry directional WHIM index / unweighted
G3u	3rd component symmetry directional WHIM index / unweighted
E1u	1st component accessibility directional WHIM index / unweighted
E2u	2ed component accessibility directional WHIM index / unweighted
E3u	3rd component accessibility directional WHIM index / unweighted

L3m	3rd component size directional WHIM index / weighted by atomic masses
G1m	1st component symmetry directional WHIM index / weighted by atomic masses
G2m	2ed component symmetry directional WHIM index / weighted by atomic masses
E1m	1st component accessibility directional WHIM index / weighted by atomic masse
E1v	1st component accessibility directional WHIM index / weighted by atomic van der Waals volumes
P2e	2nd component shape directional WHIM index / weighted by atomic Sanderson electronegativities
G2e	2st component symmetry directional WHIM index / weighted by atomic Sanderson electronegativities
E1e	1st component accessibility directional WHIM index / weighted by atomic Sanderson electronegativities
E2e	2nd component accessibility directional WHIM index / weighted by atomic Sanderson electronegativities
G2p	2st component symmetry directional WHIM index / weighted by atomic polarizabilities
G3p	3st component symmetry directional WHIM index / weighted by atomic polarizabilities
E1p	1st component accessibility directional WHIM index / weighted by atomic polarizabilities

E2p	2ed component accessibility directional WHIM index / weighted by atomic polarizabilities
L3s	3rd component size directional WHIM index / weighted by atomic electrotopological states
E1s	1st component accessibility directional WHIM index / weighted by atomic electrotopological states
Gu	G total symmetry index / unweighted
Du	D total accessibility index / unweighted
Dm	D total accessibility index / weighted by atomic masses
Ds	D total accessibility index / weighted by atomic electrotopological states
GETAWAY descriptors	
GETAWAY descriptors are descriptors calculated from the leverage matrix obtained by the centered atomic coordinates (molecular influence matrix, MIM). The first four descriptors are calculated as information content and connectivity indices. HATS and H descriptors are 3D-autocorrelation descriptors obtained from MIM; R and R+ descriptors are analogously obtained from the leverage/geometry matrix.	
ISH	standardized information content on the leverage equality
H7u	H autocorrelation of lag 7 / unweighted
H8u	H autocorrelation of lag 8 / unweighted
HATS7u	leverage-weighted autocorrelation of lag 7 / unweighted

H8m	H autocorrelation of lag 8 / weighted by atomic masses
HATS8m	leverage-weighted autocorrelation of lag 8 / weighted by atomic masses
H8e	H autocorrelation of lag 8 / weighted by atomic Sanderson electronegativities
R2u+	R maximal autocorrelation of lag 2 / unweighted
R6u+	R maximal autocorrelation of lag 6 / unweighted
R7u+	R maximal autocorrelation of lag 7 / unweighted
R5m+	R maximal autocorrelation of lag 5 / weighted by atomic masses
R7m+	R maximal autocorrelation of lag 7 / weighted by atomic masses
R8m+	R maximal autocorrelation of lag 8 / weighted by atomic masses
R1v+	R maximal autocorrelation of lag 1 / weighted by atomic van der Waals volumes
R2v+	R maximal autocorrelation of lag 2 / weighted by atomic van der Waals volumes
R3v+	R maximal autocorrelation of lag 3 / weighted by atomic van der Waals volumes
R6v+	R maximal autocorrelation of lag 6 / weighted by atomic van der Waals volumes
R5e+	R maximal autocorrelation of lag 5 / weighted by atomic Sanderson electronegativities

R7e+	R maximal autocorrelation of lag 7 / weighted by atomic Sanderson electronegativities
R8p	R autocorrelation of lag 8 / weighted by atomic polarizabilities
R1p+	R maximal autocorrelation of lag 1 / weighted by atomic polarizabilities
R2p+	R maximal autocorrelation of lag 2 / weighted by atomic polarizabilities
R23+	R maximal autocorrelation of lag 3 / weighted by atomic polarizabilities
R5p+	R maximal autocorrelation of lag 5 / weighted by atomic polarizabilities
R6p+	R maximal autocorrelation of lag 6 / weighted by atomic polarizabilities
R8p+	R maximal autocorrelation of lag 8 / weighted by atomic polarizabilities
RTp+	R maximal index / weighted by atomic polarizabilities
Functional group counts	
Molecular descriptors based on the counting of chemical functional groups.	
nCs	number of total secondary C(sp3)
nCconj	number of exo-conjugated C(sp2)
nArCO	number of ketones (aromatic)
nHBonds	number of intramolecular H-bonds (with N,O,F)

Atom-centred fragments	
Atom-centred fragments are molecular descriptors based on the counting of 120 atom-cantered fragments, as defined by Ghose-Crippen. Some fragments are undefined by the authors.	
C-039	
O-058	
Molecular properties	
Molecular properties calculated from models together with some empirical descriptors.	
LAI	Lipinski Alert index
GVWAI-50	Ghose-Viswanadhan-Wendoloski Alert index at 50% (drug-like index)
Inflammat-50	Ghose-Viswanadhan-Wendoloski antiinflammatory-like index at 50%
Depressant-80	Ghose-Viswanadhan-Wendoloski antidepressant at 80% (drug-like index)
Psychotic-80	Ghose-Viswanadhan-Wendoloski antipsychotic at 80% (drug-like index)
Psychotic-50	Ghose-Viswanadhan-Wendoloski antipsychotic at 50% (drug-like index)
Neoplastic-50	Ghose-Viswanadhan-Wendoloski antineoplastic at 50% (drug-like index)

Appendix D

Predictive Model of Glass Transition Temperature of Polyimides from Machine Learning

Table D.1 shows all features and their coefficients in the predictive model of T_g of polyimides trained with 151 data points from Ref. [142] using the training method based on "statistical splitting" and bagging.

Table D.1: Linear coefficients of all features in the predictive model of T_g.

Feature	Coefficient
Mor04u	13.046288
JGI2	9.477253
Mor11m	3.543599
GVWAI-50	3.024074
Mor17p	1.755194
Mor22v	1.721365
Mor03v	1.662579
MEcc	1.508814
Mor17e	1.23228
MATS6v	1.117188
R7e+	1.037976
Mor17u	1.027713
Neoplastic-50	0.974793
nHBonds	0.782506
Mor24p	0.760354
RDF120e	0.749853
Mor03p	0.729866
JGI9	0.709341
Mor17v	0.675846
Mor03m	0.626338
R7u+	0.570961
E2u	0.570157
nDB	0.545641

Feature	Coefficient
IC3	-5.176466
MATS4m	-3.955315
VEA1	-3.682337
Mor26m	-3.398099
R1p+	-3.056214
E1v	-2.842052
Mor02e	-2.532373
R1v+	-2.327931
RDF130m	-2.202375
PJI2	-1.651294
JGI3	-1.553203
GGI10	-1.425599
Mor29m	-1.285398
Mor02u	-1.02832
Mor25e	-0.947649
ISH	-0.93868
MATS4e	-0.932648
RDF075m	-0.828565
MATS4v	-0.806
RDF100m	-0.764006
MATS4p	-0.638701
Mor25u	-0.63294
RDF085m	-0.627684

R6u+	0.502103	E1p	-0.620146
GGI9	0.49145	Mor25m	-0.603496
O-058	0.484375	R2v+	-0.543559
MATS6p	0.478762	RCI	-0.471169
JGI4	0.464782	JGI10	-0.450297
R5m+	0.449641	BIC3	-0.433845
Mor21u	0.418468	E3s	-0.433083
SPAM	0.390473	L3s	-0.400674
RDF145u	0.377485	Dm	-0.344892
R2u+	0.347814	Mor06v	-0.305683
E2e	0.324195	R3v+	-0.299862
Mor08u	0.264495	Mor29v	-0.297368
Mor22e	0.23801	Mor10p	-0.297121
Mor19u	0.235263	RTp+	-0.294698
PW5	0.218135	R5u+	-0.288912
RDF045m	0.17446	JGI1	-0.286692
Mor22u	0.165338	RDF070m	-0.285107
Mor04p	0.165232	RTv+	-0.277181
TOS	0.164877	EEig12x	-0.267931
Mor23m	0.152134	Mor10v	-0.236662
JGI8	0.137205	RDF090e	-0.226201
Mor16m	0.13109	Mor13u	-0.21849
H7u	0.123767	BELv5	-0.218087
MATS1m	0.114605	Mor10u	-0.205804
H6m	0.106284	RDF055m	-0.202087

nCconj	0.105095	R8e+	-0.182816
GATS5m	0.092177	R6v+	-0.168503
GATS5p	0.089884	GATS8e	-0.142111
SRW09	0.088207	HOMA	-0.127141
Mor14e	0.072911	BELp5	-0.123678
R5e+	0.071676	RDF095m	-0.122296
G3p	0.071423	R1m+	-0.121098
RDF120m	0.071132	RDF050v	-0.113444
Mor24u	0.05863	L2p	-0.111242
MATS8m	0.057825	LAI	-0.106511
JGI7	0.050859	R6p+	-0.105849
nCs	0.050296	H8u	-0.098917
Mor28m	0.04959	RDF090m	-0.096145
Mor31m	0.048422	Hypnotic-50	-0.092558
RDF115u	0.046823	R2p+	-0.083373
RDF110u	0.044665	RDF155v	-0.079466
RDF145e	0.042256	Mor30e	-0.071162
C-025	0.039671	Mor02v	-0.064702
R6e+	0.037562	BELm5	-0.063507
G3v	0.037239	BELp8	-0.059291
MATS8p	0.036776	Ku	-0.059057
P2s	0.035681	DISPp	-0.055435
RDF125m	0.034567	MATS8e	-0.050401
Mor12u	0.033794	L2v	-0.050032
Mor28u	0.033505	R8u	-0.048706

G2s	0.031251	AROM	-0.044368
Mor18m	0.030071	R3m+	-0.0398
Mor24v	0.026671	Psychotic-80	-0.038048
Psychotic-50	0.024805	RTm+	-0.034295
GNF	0.023216	DISPe	-0.033192
Mor03e	0.021333	Mor30m	-0.032476
Mor05u	0.021315	GATS8p	-0.032221
TNN	0.020443	SPH	-0.031749
Mor08e	0.020156	Mor28v	-0.030508
RDF155m	0.019378	RDF070e	-0.025199
R7m+	0.015938	MATS8v	-0.022534
FDI	0.015222	E1m	-0.019097
RDF150m	0.015163	PJI3	-0.015838
R8m+	0.01308	R4m+	-0.015492
Mor16e	0.012181	P1u	-0.015364
Mor08p	0.010846	Mor12m	-0.014549
RDF140v	0.01021	G2m	-0.011805
TNF	0.006238	RDF055v	-0.010748
C-039	0.004455	Mor21m	-0.010587
R8p+	0.004017	RDF155p	-0.0087
Mor22p	0.00348	H8m	-0.006323
RDF105m	0.00193	Mor09u	-0.004866
DISPv	0.001492	RDF140u	-0.003518
C-002	0.000152	GOF	-0.001866
		G2v	-0.001213

		GATS8v	-0.000789
		TOF	-0.000024