

Bonded Potentials of Coarse-Grained Polymer Models

Qiang Wang^{*}

Department of Chemical and Biological Engineering, Colorado State University

1370 Campus Delivery, Fort Collins, CO 80523-1370

Abstract

^{*} E-mail: q.wang@colostate.edu

Polymers are long-chain molecules consisting of many repeating units called monomers. Given the large number (typically thousands to millions) of monomers, each containing several to hundreds of atoms, on each chain, full atomistic simulations of many-chain systems such as polymer melts used in experiments are, in most cases, not feasible at present due to their formidable computational requirements. Developing coarse-grained (CG) models at different length scales and linking these scales have therefore been proposed to **solve real-world polymer problems that could not be solved in any other way**.¹⁻⁶ In most CG polymer models, however, bonded interactions are heuristically taken to be analogous to those in atomistic models, which consist of bond stretch (two-body), angle bend (three-body), and sometimes torsion (four-body) potentials, but among consecutive CG segments (superatoms), instead of atoms, on the same chain.^{1-4, 7} These bonded potentials are used to reproduce only the *local* intrachain structures. Also, they and non-bonded CG potentials are often obtained sequentially, sometimes even independently. In our recent paper⁸, however, we showed that bonded CG potentials (taken to be isotropic pair potentials), in the most commonly used structure-based coarse graining (which reproduces some distribution functions of the original system)^{9, 10}, need to reproduce the intrachain segment pair correlations functions (PCFs) at all length scales, and that bonded and non-bonded CG potentials are coupled and need to be obtained simultaneously. Here we show how to achieve these in various ways, focusing on the case where each CG segment represents the center-of-mass of a group of consecutive monomers on the same chain.

As in our previous work⁸, we consider the original system of a melt of n homopolymer chains each of N_m monomers in volume V at thermodynamic temperature T . Its canonical-ensemble configuration integral is given by $Z_m = \int d\mathbf{r} \exp(-\beta H_m^b(\mathbf{r}) - \beta H_m^{nb}(\mathbf{r}))$, where $\beta \equiv 1/k_B T$ with k_B being the Boltzmann constant, $\mathbf{r} \equiv \{\mathbf{r}_{k,t}\}$ with $\mathbf{r}_{k,t}$ being the spatial position of the t^{th} monomer on

the k^{th} chain, H_m^b is the Hamiltonian due to chain connectivity, and H_m^{nb} is due to non-bonded interactions in the system. For the CG system, we similarly have

$$Z = \int d\mathbf{R} \exp\left(-\beta H^b(\mathbf{R}) - \beta H^{nb}(\mathbf{R})\right), \text{ where } \mathbf{R} \equiv \{\mathbf{R}_{k,\alpha}\} \text{ with } \mathbf{R}_{k,\alpha} = \frac{1}{l} \sum_{t=(\alpha-1)l+1}^{\alpha l} \mathbf{r}_{k,t} \text{ being the spatial}$$

position of α^{th} segment representing the center-of-mass of l consecutive monomers on the k^{th} chain;

this defines the mapping operator \mathbf{M} that gives $\mathbf{R} = \mathbf{M}(\mathbf{r})$. In the first approach, we take

$$H^b = \sum_{k=1}^n \sum_{\alpha=1}^{N-1} \sum_{\gamma=\alpha+1}^N v_{\alpha,\gamma}^b(|\mathbf{R}_{k,\alpha} - \mathbf{R}_{k,\gamma}|; \boldsymbol{\lambda}_{\alpha,\gamma}^b), \text{ where } N \equiv N_m/l \text{ is the number of CG segments on each}$$

chain, $v_{\alpha,\gamma}^b(r; \boldsymbol{\lambda}_{\alpha,\gamma}^b)$ denotes the isotropic bonded CG pair potential acting between segments α and

γ , and $\boldsymbol{\lambda}_{\alpha,\gamma}^b$ denotes a set of parameters used to parameterize $v_{\alpha,\gamma}^b(r)$;

$$H^{nb} = \frac{1}{2} \sum_{k=1}^n \sum_{\alpha=1}^N \sum_{j=1}^n \sum_{\gamma=1}^N v^{nb}(|\mathbf{R}_{k,\alpha} - \mathbf{R}_{j,\gamma}|; \boldsymbol{\lambda}^{nb}) - \frac{nN}{2} v^{nb}(0; \boldsymbol{\lambda}^{nb}), \text{ where } v^{nb}(r; \boldsymbol{\lambda}^{nb}) \text{ denotes the isotropic}$$

non-bonded CG pair potential, $\boldsymbol{\lambda}^{nb}$ denotes a set of parameters used to parameterize $v^{nb}(r)$, and the

last term deducts the self-interaction included in the previous term. Minimizing the relative entropy

$$(RE)^{11} \quad S \equiv -\int d\mathbf{r} \mathcal{P}_m(\mathbf{r}) \ln \mathcal{P}(\mathbf{M}(\mathbf{r})) = \beta \langle H^b \rangle_m + \beta \langle H^{nb} \rangle_m + \ln \int d\mathbf{R} \exp\left(-\beta H^b(\mathbf{R}) - \beta H^{nb}(\mathbf{R})\right)$$

$$\text{with respect to } \boldsymbol{\lambda} \equiv \{\boldsymbol{\lambda}_{\alpha,\gamma}^b, \boldsymbol{\lambda}^{nb}\} \text{ then gives } \left\langle \frac{\partial \beta H^b}{\partial \boldsymbol{\lambda}_{\alpha,\gamma}^b} \right\rangle_m = \left\langle \frac{\partial \beta H^b}{\partial \boldsymbol{\lambda}_{\alpha,\gamma}^b} \right\rangle_{\text{CG}} \text{ and } \left\langle \frac{\partial \beta H^{nb}}{\partial \boldsymbol{\lambda}^{nb}} \right\rangle_m = \left\langle \frac{\partial \beta H^{nb}}{\partial \boldsymbol{\lambda}^{nb}} \right\rangle_{\text{CG}},$$

where $\mathcal{P}_m(\mathbf{r}) \equiv \exp\left(-\beta H_m^b(\mathbf{r}) - \beta H_m^{nb}(\mathbf{r})\right)/Z_m$ and $\langle \rangle_m$ are the configuration probability and

ensemble average, respectively, in the original system, and

$\mathcal{P}(\mathbf{R}) \equiv \exp\left(-\beta H^b(\mathbf{R}) - \beta H^{nb}(\mathbf{R})\right)/Z$ and $\langle \rangle_{\text{CG}}$ are those in the CG system. With

$$\langle \beta H^b \rangle_m = 4\pi n \int_0^\infty dr r^2 \sum_{\alpha=1}^{N-1} \sum_{\gamma=\alpha+1}^N \beta v_{\alpha,\gamma}^b(r; \boldsymbol{\lambda}_{\alpha,\gamma}^b) \omega_{\alpha,\gamma}^{ss}(r) \quad \text{and}$$

$$\langle \beta H^{nb} \rangle_m = 2\pi n N^2 \int_0^\infty dr r^2 \left[\omega^{ss}(r) + \rho_c (h^{ss}(r) + 1) \right] \beta v^{nb}(r; \boldsymbol{\lambda}^{nb}) - \frac{nN}{2} \beta v^{nb}(0; \boldsymbol{\lambda}^{nb}), \text{ where } \omega_{\alpha,\gamma}^{ss}(r)$$

is the normalized (i.e., $\int_0^\infty dr 4\pi r^2 \omega_{\alpha,\gamma}^{ss}(r) = 1$) intrachain PCF between segments α and γ in the

original system, $\omega^{ss}(r) \equiv \frac{1}{N^2} \sum_{\alpha=1}^N \sum_{\gamma=1}^N \omega_{\alpha,\gamma}^{ss}(r)$, $\rho_c \equiv n/V$ is the chain number density, and $h^{ss}(r)$ is the

interchain total PCF between two segments (which is just the interchain segment radial distribution

function minus 1) in the original system, and similar results for the CG system, where $\omega_{\alpha,\gamma}(r; \boldsymbol{\lambda})$,

$\omega(r; \boldsymbol{\lambda})$ and $h(r; \boldsymbol{\lambda})$ are defined, we finally obtain

$$\begin{cases} \int_0^\infty dr r^2 \left[\omega_{\alpha,\gamma}^{ss}(r) - \omega_{\alpha,\gamma}(r; \boldsymbol{\lambda}) \right] \frac{\partial \beta v_{\alpha,\gamma}^b(r; \boldsymbol{\lambda}_{\alpha,\gamma}^b)}{\partial \boldsymbol{\lambda}_{\alpha,\gamma}^b} = 0 & (\text{for } 1 \leq \alpha < \gamma \leq N) \\ \int_0^\infty dr r^2 \left[\omega^{ss}(r) - \omega(r; \boldsymbol{\lambda}) + \rho_c (h^{ss}(r) - h(r; \boldsymbol{\lambda})) \right] \frac{\partial \beta v^{nb}(r; \boldsymbol{\lambda}^{nb})}{\partial \boldsymbol{\lambda}^{nb}} = 0 \end{cases} \quad (1)$$

When $v_{\alpha,\gamma}^b(r)$ and $v^{nb}(r)$ are unconstrained (e.g., having an infinite number of parameters $\boldsymbol{\lambda}_{\alpha,\gamma}^b$ and

$\boldsymbol{\lambda}^{nb}$, respectively), the above RE-based coarse graining becomes structure-based coarse graining¹²

and gives $\omega_{\alpha,\gamma}(r; \boldsymbol{\lambda}) = \omega_{\alpha,\gamma}^{ss}(r)$ and $h(r; \boldsymbol{\lambda}) = h^{ss}(r)$ at each r . This clearly shows that bonded and

non-bonded CG potentials are coupled and thus need to be determined simultaneously. In our

recently proposed systematic and simulation-free strategy for structure-based coarse graining of

homopolymer melts, $\beta v^{nb}(r)$ is obtained *under the conditions of* $\omega(r; \boldsymbol{\lambda}) = \omega^{ss}(r)$ and

$h(r; \boldsymbol{\lambda}) = h^{ss}(r)$;¹³ with this $\beta v^{nb}(r)$, we can then adjust $v_{\alpha,\gamma}^b(r; \boldsymbol{\lambda}_{\alpha,\gamma}^b)$ such that $\omega_{\alpha,\gamma}(r; \boldsymbol{\lambda}) = \omega_{\alpha,\gamma}^{ss}(r)$

for $1 \leq \alpha < \gamma \leq N$ (i.e., to achieve structure-based coarse graining). In this case, however, $(N^2 - 1)/4$ (for

odd N) or $N^2/4$ (for even N) isotropic bonded CG pair potentials need to be obtained simultaneously,

which is computationally challenging for large N .

In the second approach, we take $H^b = \sum_{k=1}^n \sum_{i \in \{i^b\}} \sum_{\alpha=1}^{N-i} v_i^b(|\mathbf{R}_{k,\alpha+i} - \mathbf{R}_{k,\alpha}|; \boldsymbol{\lambda}_i^b)$, where $\{i^b\}$ denotes

a set of m different integers within $[1, N-1]$, $v_i^b(r; \boldsymbol{\lambda}_i^b)$ denotes the isotropic bonded CG pair potential acting between two segments separated by $i-1$ CG segments along the chain, and $\boldsymbol{\lambda}_i^b$ denotes a set of parameters used to parameterize $v_i^b(r)$. Following the above derivation, we finally obtain

$$\int_0^\infty dr r^2 [\omega_i^{ss}(r) - \omega_i(r; \boldsymbol{\lambda})] \frac{\partial \beta v_i^b(r; \boldsymbol{\lambda}_i^b)}{\partial \boldsymbol{\lambda}_i^b} = 0 \quad (\text{for } i \in \{i^b\}) \quad (2)$$

along with the second equation in Eq. (1), where $\omega_i^{ss}(r) \equiv \frac{1}{N-i} \sum_{\alpha=1}^{N-i} \omega_{\alpha, \alpha+i}^{ss}(r)$, $\omega_i(r; \boldsymbol{\lambda})$ is similarly defined in the CG system, and $\boldsymbol{\lambda} \equiv \{\boldsymbol{\lambda}_i^b, \boldsymbol{\lambda}^{nb}\}$. In the case of structure-based coarse graining, only m ($< N$) isotropic bonded CG pair potentials need to be adjusted simultaneously to ensure $\omega_i(r; \boldsymbol{\lambda}) = \omega_i^{ss}(r)$ for $i \in \{i^b\}$, which is computationally much easier for large N than the first approach. We further note that, although $m=N-1$ is needed for the CG model to exactly reproduce all $\omega_i^{ss}(r)$, $m \ll N$ can be used for large N to reproduce *all* $\omega_i^{ss}(r)$ *within good accuracy*, which makes our second approach very attractive. This is demonstrated below.

While both approaches apply to any original model system for which $\omega_{\alpha, \gamma}^{ss}(r)$ is known (usually via simulations, thus with some statistical uncertainties), hereafter we take the ideal conformations of continuous Gaussian chains (CGCs) as an example, for which we have the exact result¹³ of $\omega_{\alpha, \alpha+i}^{ss}(r) = \omega_i^{ss}(r) = (a_i/\pi)^{3/2} \exp(-a_i r^2)$ with $a_i \equiv 3N/2(i-1/3)$ and r in units of the ideal-chain root-mean-square end-to-end distance in the original system; this is a good approximation for long-chain polymer melts under the Flory's ideality hypothesis¹⁴. We first set

$\beta v^{nb}(r)=0$; the results for this case are denoted by a superscript “0”. One can show that

$\beta v_{\alpha,\gamma}^{b,0}(r) = \lambda_{\alpha,\gamma}^{b,0} r^2$ in our first approach for structure-based coarse graining. In particular, we find the analytical solution of $\lambda_{1,2}^{b,0} = \lambda_{2,3}^{b,0} = 9$ and $\lambda_{1,3}^{b,0} = -9/5$ for $N=3$, and that of $\lambda_{1,2}^{b,0} = \lambda_{3,4}^{b,0} = 171/14$, $\lambda_{2,3}^{b,0} = 225/14$, $\lambda_{1,3}^{b,0} = \lambda_{2,4}^{b,0} = -45/14$, and $\lambda_{1,4}^{b,0} = 9/14$ for $N=4$. The difference between $\lambda_{1,2}^{b,0} = \lambda_{3,4}^{b,0}$ and $\lambda_{2,3}^{b,0}$ for $N=4$ (which is due to the chain-end effects of the CG model) further indicates that $\beta v_i^{b,0}(r)$ in our second approach cannot all be harmonic in structure-based coarse graining for $N \geq 4$.

We therefore take $\beta v_i^{b,0}(r) = \lambda_i^{b,0} r^2$ and perform the RE-based coarse graining, where Eq. (2) becomes

$$\langle R_{e,i}^2 \rangle_{\text{CG}} = \langle R_{e,i}^2 \rangle_m \quad (\text{for } i \in \{i^b\}) \quad (3)$$

with $\langle R_{e,i}^2 \rangle_{\text{CG}} = \int_0^\infty dr r^2 P_i(r; \lambda)$ being the ensemble average of mean-square internal distance between two segments separated by $i-1$ CG segments along the chain in the CG system, $P_i(r) \equiv 4\pi r^2 \omega_i(r)$, and $\langle R_{e,i}^2 \rangle_m$ similarly defined in the original system. For $N=4$ (and $m=N-1$),

we find the analytical solution of $\lambda_1^{b,0} = 9(8 + \sqrt{7})/8$, $\lambda_2^{b,0} = 9(\sqrt{7} - 4)/8$, and $\lambda_3^{b,0} = 9(5 - 2\sqrt{7})/8$, which gives exactly the same $P_i(r)$, except for $i=1$, as the above analytical

solution obtained in our first approach for structure-based coarse graining (which gives

$P_i(r) = P_i^{ss}(r) \equiv 4\pi r^2 \omega_i^{ss}(r)$ for all i); Fig. 1 further shows that

$$P_1(r) = \frac{2}{3} \times \frac{216}{29} (5 + 4\sqrt{7})^{3/2} \sqrt{\frac{2}{29\pi}} r^2 \exp\left[-\frac{18}{29} (5 + 4\sqrt{7}) r^2\right] + \frac{1}{3} \times \frac{81}{(7 - 2\sqrt{7})^{3/2}} \sqrt{\frac{6}{\pi}} r^2 \exp\left[-\frac{9}{14} (7 + 2\sqrt{7}) r^2\right]$$

obtained in the second approach is nearly indistinguishable from $P_1^{ss}(r)$.

To obtain $\langle R_{e,i}^2 \rangle_{\text{CG}}$ for larger N , we use single-chain Monte Carlo (SCMC) simulations

with hopping¹⁴ and pivot¹⁵ trial moves; the latter greatly reduce the sample correlation length (estimated by the correlation-function method¹⁶), especially for large N . We use $\sim 10^5$ statistically uncorrelated samples collected after equilibration to calculate $\langle R_{e,i}^2 \rangle_{\text{CG}}$ and its statistical uncertainty σ_i (taken to be three times the standard deviation of the collected, statistically uncorrelated samples).

Most importantly, instead of solving Eq. (3), we minimize $\sum_{i=1}^{N-1} \left(\langle R_{e,i}^2 \rangle_{\text{CG}} - \langle R_{e,i}^2 \rangle_m \right)^2 / \sigma_i^2$

with respect to the m parameters $\lambda_i^{b,0}$ using the Levenberg-Marquardt method¹⁷, where we also employ the reweighting technique¹⁷ to re-use samples collected in the previous SCMC simulation.

While this non-linear minimization is equivalent to solving Eq. (3) when $m=N-1$, it allows us to use $m \ll N$ for large N to reproduce *all* $\langle R_{e,i}^2 \rangle_{\text{CG}}$ within the simulation accuracy; that is, we can

choose the smallest m such that $\left| \langle R_{e,i}^2 \rangle_{\text{CG}} - \langle R_{e,i}^2 \rangle_m \right| < \sigma_i$ is satisfied for all i ($=1, \dots, N-1$) in the

last SCMC simulation with the minimized parameters, when the iteration of alternating non-linear minimization and SCMC simulation converges. Finally, as we use histograms for the distance

between segment pairs in SCMC simulations to obtain $P_{i,j} \equiv \int_{(j-1)\Delta r}^{j\Delta r} dr P_i(r)$, where $\Delta r=0.1$ is the

bin size of the histograms, we calculate its maximum absolute deviation ΔP_i from

$P_{i,j}^{ss} \equiv \int_{(j-1)\Delta r}^{j\Delta r} dr P_i^{ss}(r)$ over all j ($=1, 2, \dots$) in the last SCMC simulation.

Table 1 lists the minimized $\lambda_i^{b,0}$ for various N at convergence, where the uncertainties are taken to be the square root of the diagonal elements of the covariance matrix calculated at the end

of the minimization[□], and an entry of 0 means that we set $\lambda_i^{b,0} = 0$; we also set $\lambda_i^{b,0} = 0$ for all other i -values not shown in the table.

Table 1: Minimized $\lambda_i^{b,0}$ for various N at convergence; see main text for details.

		i				
		1	2	3	4	$N-1$
N	10	35.6 ± 0.4	-7.33 ± 0.27	1.05 ± 0.11	0	-0.081 ± 0.024
	20	76.5 ± 0.9	-17.6 ± 0.9	3.33 ± 0.62	-0.38 ± 0.20	-0.014 ± 0.014
	40	159 ± 3	-39.0 ± 2.1	8.3 ± 1.3	-1.10 ± 0.32	0
	80	324 ± 3	-80.0 ± 2.9	16.4 ± 1.7	-1.95 ± 0.42	0
	160	654 ± 5	-164 ± 5	34.1 ± 2.5	-4.08 ± 0.61	0

Now we include $\beta v^{nb}(r)$ obtained from our structure-based coarse graining¹³. In this case, many-chain molecular simulations (MCMS), instead of SCMC simulations, can be used to obtain $\langle R_{e,i}^2 \rangle_{\text{CG}}$ and P_{ij} , and our second approach described above can still be used to obtain $\beta v_i^{b,\text{MCMS}}(r)$ (still harmonic?) with the smallest m such that all $\langle R_{e,i}^2 \rangle_m$ and $P_{i,j}^{ss}$ are reproduced within the simulation accuracy. With the idea underlying the well-developed self-consistent polymer reference interaction site model (SC-PRISM) theory¹⁸, however, we can avoid MCMS and still use the above SCMC results. In the SC-PRISM theory, the interchain interaction is approximated by an effective intrachain solvation potential $w(r)$ acting between all segment pairs; combined with our systematic and simulation-free strategy for structure-based coarse graining¹³, it is given by¹⁹⁻

21

$$\beta \hat{w} = -\kappa \rho_c (\hat{h}^{ss})^2 / N^2 (\hat{w}^{ss})^2 (\hat{w}^{ss} + \rho_c \hat{h}^{ss}), \quad (4)$$

where the short-hand notation $\hat{f} \equiv (4\pi/q) \int_0^\infty dr f(r) r \sin(qr)$ denotes the 3D Fourier transform of a radial function $f(r)$ with q being the wavenumber, and $\kappa \in [0.9, 1.1]$ was introduced by Mendez et

al.²¹ to obtain the correct behavior of $\langle R_{e,N-1}^2 \rangle_{CG}$. We then have $v_i^{b,0}(r) = v_i^{b,SCMC}(r) + v^{nb}(r) + w(r)$

for $i=1,\dots,N-1$ in order for the SCMC simulation with $\{\beta v_i^{b,SCMC}(r)\}$, $\beta v^{nb}(r)$, and $\beta w(r)$ to give approximately the same results as the MCMS simulation with $\{\beta v_i^{b,MCMS}(r)\}$ and $\beta v^{nb}(r)$ (i.e., reproducing all $\langle R_{e,i}^2 \rangle_m$ and $P_{i,j}^{ss}$). While $v_i^{b,SCMC}(r) \neq v_i^{b,MCMS}(r)$, their difference should be small if the above approximation of the SC-PRISM theory is good.

Perform MCMS simulation with $\{\beta v_i^{b,SCMC}(r)\}$ and $\beta v^{nb}(r)$ to see if it can reproduce all

$\langle R_{e,i}^2 \rangle_m$ and $P_{i,j}^{ss}$? Start with $N=3$ or 4.

References:

1. F. Muller-Plathe, *Chemphyschem* **3** (9), 754-769 (2002).
2. F. Muller-Plathe, *Soft Mater.* **1** (1), 1-31 (2003).
3. C. Peter and K. Kremer, *Soft Matter* **5** (22), 4357-4366 (2009).
4. C. Peter and K. Kremer, *Faraday Discuss.* **144**, 9-24 (2010).
5. R. Potestio, C. Peter and K. Kremer, *Entropy-Switz* **16** (8), 4199-4245 (2014).
6. M. Praprotnik, L. Delle Site and K. Kremer, *Annu. Rev. Phys. Chem.* **59**, 545-571 (2008).
7. E. Brini, E. A. Algaer, P. Ganguly, C. L. Li, F. Rodriguez-Ropero and N. F. A. van der Vegt, *Soft Matter* **9** (7), 2108-2119 (2013).
8. D. L. Yang and Q. Wang, *Soft Matter* **11** (36), 7109-7118 (2015).
9. A. P. Lyubartsev and A. Laaksonen, *Phys. Rev. E* **52** (4), 3730-3737 (1995).
10. A. K. Soper, *Chem. Phys.* **202** (2-3), 295-306 (1996).
11. M. S. Shell, *J. Chem. Phys.* **129** (14), 144108 (2008).
12. A. Chaimovich and M. S. Shell, *J. Chem. Phys.* **134** (9), 094112 (2011).
13. D. L. Yang and Q. Wang, *J. Chem. Phys.* **142** (5), 054905 (2015).
14. Q. Wang and Y. H. Yin, *J. Chem. Phys.* **130** (10), 104903 (2009).
15. J. Zong and Q. Wang, *J. Chem. Phys.* **139** (12), 124907 (2013).
16. G. Jacucci and A. Rahman, *Nuovo Cimento D* **4** (4), 341-356 (1984).
17. A. M. Ferrenberg and R. H. Swendsen, *Phys Rev Lett* **61** (23), 2635-2638 (1988).
18. D. R. Heine, G. S. Grest and J. G. Curro, *Adv. Polym. Sci.* **173**, 209-249 (2005).
19. D. Chandler, Y. Singh and D. M. Richardson, *J. Chem. Phys.* **81** (4), 1975-1982 (1984).
20. A. L. Nichols, D. Chandler, Y. Singh and D. M. Richardson, *J. Chem. Phys.* **81** (11), 5109-5115 (1984).
21. S. Mendez, J. G. Curro, M. Putz, D. Bedrov and G. D. Smith, *J. Chem. Phys.* **115** (12), 5669-5678 (2001).