Good afternoon, thank you for coming. I guess we are all OK with the analytical bonded potentials acting among consecutive backbone atoms on a chain used in atomistic models. In most coarse-grained models, however, similar bonded potentials are heuristically used to reproduce, to some extent, only the local intrachain structures, i.e., the distributions of the bond length, the bond angle, and sometimes the torsion angle, not among the atoms but among the coarse-grained segments each representing a group of consecutive monomers on the same chain, and such distributions are often obtained from simulations of the atomistic (or say the original) system. Also, the bonded and non-bonded coarse-grained potentials are usually determined sequentially, for example, in the iterative Boltzmann inversion, which requires some tedious iterations among the various types of potentials to obtain more accurate results. Sometimes people are lazy, so they just borrow the bonded coarse-grained potentials from another system and claim that they do not matter that much. So my question here is: “Can we do better?”

1 min

Consider the relative-entropy-based coarse graining of homopolymer melts for example. Let the small-**r** be a configuration of the original system, and capital-**R** the corresponding configuration of the coarse-grained system. The relative entropy *S* quantifying the information loss due to coarse graining, can be defined as this, where *m* and  are the configuration-space probabilities in the original and coarse-grained systems, respectively, *Hm* is the total Hamiltonian of the original system and includes both its bonded and non-bonded potentials, *Hb* and *Hnb* are such potentials of the coarse-grained system, and *****b* and *****nb* are their parameters to be determined. Minimizing *S* with respect to these parameters then gives two coupled equations, where the angle brackets denote the ensemble-average in the original and coarse-grained systems, respectively. We therefore see that the bonded and non-bonded potentials are coupled and must be determined simultaneously.

For simplicity, we use isotropic pair potentials *vb* and *vnb* in a system of *n* chains at chain number density *c*, where each chain is coarse-grained as *N* segment. In particular, as the first approach we use different *vb* for different segment pairs *s* and *t*. The ensemble-average of the bonded potentials can therefore be written like this, where *ms,t* is the normalized intrachain correlation function for the pair. Similarly, here is the ensemble-average of the non-bonded potential, where **(*r*) is the intrachain correlation function averaged over all segment pairs, *h*(*r*) is the interchain segment total pair correlation function, and this last term is due to the segment self-interaction. These are for the original system as denoted by “*m*”, and we also have the corresponding results for the coarse-grained system without “*m*”. The two coupled equations then become these, where **** denotes ***all*** the coarse-grained potential parameters.

When we have an infinite number of such parameters, or say when the forms of the coarse-grained pair potentials are unlimited, the relative-entropy-based coarse graining becomes the structure-based coarse graining, which gives these two equations at all *r*. Note that in our previous work of structure-based coarse graining of homopolymer melts using the polymer integral-equation theories, we obtained *vnb* under these two conditions; so we can further obtain *vbs,t* from this equation. In other words, these bonded potentials reproduce the intrachain structures at all length scales, and are in fact determined together with the non-bonded potential as they should be. The only downside here is that there are too many of them, on the order of *N*2; so this approach is computational challenging when *N* is large.

As the second approach, we use different *vb* for segment pairs separated by different number of segments, *i*1, along a chain, so *i*1 corresponds to two adjacent segments. Note that we take only *m* values of *i*, denoted by this subset {*ib*}. In this ensemble-average of the bonded potentials, *mi* is the intrachain correlation function averaged over all segment pairs separated by *i*1 segments along a chain in the original system. Following our first approach, we then obtain these two equations given by the relative-entropy-based coarse graining. Again, for the structure-based coarse graining, these two equations must be satisfied at all *r*. And with our previous *vnb*, we can further obtain *vbi* from this equation, the total number of which is less than *N*. Even better, although we need *N*1 of such bonded potentials for the coarse-grained model to exactly reproduce all *mi*, a much smaller number of them can just be used in the case of large *N* to reproduce all *mi* within good accuracy.

4 min

To demonstrate this, we take the ideal conformations of continuous Gaussian chains as an example, for which the exact results of *m* are known; this is also a good approximation for long-chain polymer melts under the Flory’s ideality hypothesis. Let’s first consider the case where the non-bonded potential is 0, denoted by the superscript “0” here. One can show that the bonded potentials in our first approach must be harmonic in this case. In particular, for *N*4, we have the **-values shown here. That this one is different from these two is due to the chain-end effect and indicates that, in our second approach, the bonded potentials cannot all be harmonic. So for the structure-based coarse graining using our second approach, we have to solve these three equations at all *r* numerically to obtain the bonded potentials.

Let’s do something different: Instead of the structure-based coarse graining, we go back to the relative-entropy based coarse graining and take the bonded potentials in our second approach to be harmonic, which then gives this equation with *Re,i* denoting the internal distance between two segments separated by *i*1 segments along a chain. The corresponding **-values are shown here, which give exactly the same **2 and **3 as in the original system, and **1 nearly indistinguishable from that in the original system (or say, given by the structure-based coarse graining).

For larger *N*, we use single-chain Monte Carlo simulations to obtain the mean-square internal distances in the coarse-grained system. Most importantly, instead of solving these *m* equations, we minimize, with respect to *m* **-parameters, the summation of the squared deviations in the mean-square internal distances between the original and coarse-grained systems, each weighted by the estimated statistical uncertainty *i* of the mean-square internal distance from our simulations; note that here the summation is over all *i*-values, not just those in this subset.

While this non-linear minimization is equivalent to solving these *m* equations when *m**N*1, it allows us to use *m*-values much smaller than *N* in the case of large *N* to reproduce *all* the mean-square internal distances within the simulation accuracy; that is, we can choose the smallest subset {*ib*} such that this is satisfied for all *i*. Here show the so-obtained coefficients of the harmonic potentials, where an entry of “0” means that bonded potential is set to 0, so are other bonded potentials not shown here. We see that, for small *N* (less than about 20), a bonded potential is needed between the two end segments; while for large *N*, bonded potentials are needed only for segment pairs separated by up to three segments along a chain. These results are for the case where the non-bonded potential is 0.

4 min (10 min)

When it’s not 0 or say taken from our previous work that I showed you before, we can still use this idea, where the mean-square internal distances can be obtained from many-chain simulations. We can, however, avoid such simulations and take advantage of these single-chain Monte Carlo results, if we apply the idea underlying the well-developed self-consistent polymer integral-equation theories, where the interchain interaction is approximated by an intrachain solvation potential *w* acting between all segment pairs; here the hat denotes the 3D Fourier transform of a radial function with *q* being the wavenumber. We therefore have this result, since all the potentials here act on all segment pairs.

To summarize, we have shown that the bonded coarse-grained potentials in the structure-based coarse graining need to reproduce the intrachain segment pair correlation functions at ***all*** length scales, and that the bonded and non-bonded potentials need to be determined simultaneously; we have proposed several ways to achieve this.