

Today's lecture: methods for developing simplified, coarse-grained models that can be used to span larger length and longer time scales than those reachable by fully atomistic simulations

Multiscale simulations

For nearly all systems of interest to us, the most transferrable and fundamental description of matter is one that invokes quantum mechanics. At the highest level of accuracy, this amounts to solving Schrodinger's equation for all of the subatomic particles in a system. Computationally, many approximations need to be made in order to use ab initio methods, and even these techniques are limited to small numbers of atoms.

There are both practical and philosophical reasons for performing simulations on simpler systems that do not entail a full solution of the quantum-mechanical equations. Practical reasons stem from the need to treat larger systems and run simulations for longer times than those that ab initio methods can achieve. Philosophically, we are encouraged to use simpler models because the driving forces behind the phenomena of interest often naturally manifest at larger length and time scales such that we do not need fine-grained detail in order to understand them.

Ultimately our view of the world is **multiscale** in nature: there are different models and ways of looking at systems' behaviors as we traverse different length and time scales. At larger scales, we are concerned with the effective interactions and driving forces that **emerge** from the detailed description of the system.

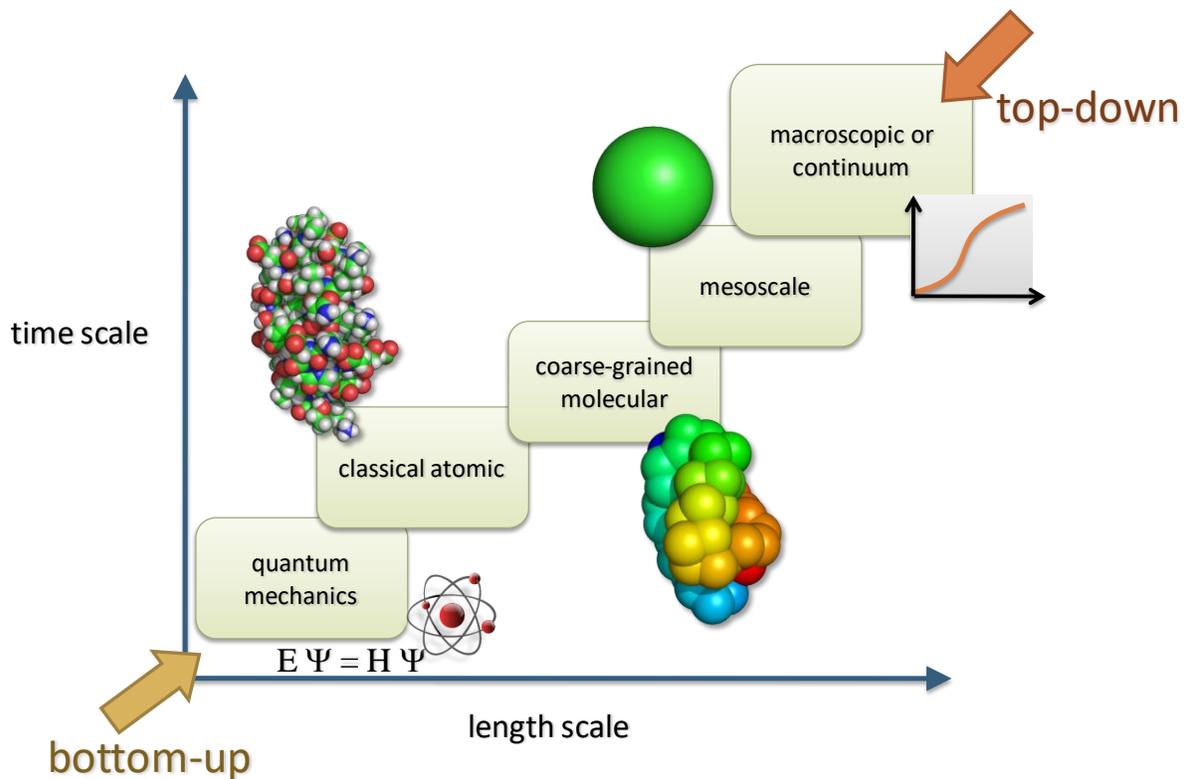
The term **multiscale modeling** is widely used to describe a hierarchy of simulation approaches to treating systems across different scales. For a given scale of interest, one picks a simulation method capable of simulating systems at the length and time scales of interest. A common way of viewing this approach is a multiscale diagram, below.

When one moves to larger scales, a **coarse-grained** model is required to enable simulations. One challenge and ongoing area of research is linking simulations of detailed models with simulations of coarse-grained ones, such that the behavior of the system can be resolved at multiple scales and the interactions at different scales inform each other. There are two ways that this information can be propagated:

- **bottom-up** – fundamental physical principles at the more detailed scale are used to parametrize a model at a coarse-grained scale. One example is when atomistic simulations

are used to calculate fluid transport coefficients that can then be used in continuum simulations of the fluid transport equations (e.g., fluid dynamics, diffusion.)

- **top-down** – the behavior at larger scales is used to inform the interactions at more detailed scales. An example might be that bulk, continuum electrostatic calculations on a large system could be used to parameterize the electric field in atomistic simulations. Or, input from experiments—such as phase behavior—could guide the development of molecular models.



Coarse-grained (CG) models

In this lecture, we will focus only on one aspect of the multiscale technique: the bottom-up approach in which information at smaller scales is used to inform models at larger scales. The large-scale models are generally termed **coarse-grained**. We will focus on CG models that are pseudo-molecular in nature (e.g., not continuum) and derived from atomistic classical models.

The primary challenge is to develop a CG model that is significantly easier to simulate but that reproduces the same physical behavior as a reference all-atom one. There are three choices that we can make when designing a CG model:

- CG models typically entail **pseudoatom sites** that are designed to represent combined groups of multiple atoms. For example, a bead-spring polymer model might involve one pseudoatom per monomer. We must choose how many pseudoatom sites there are and how they connect to the detailed model.
- We must define an **effective energy function** U_{CG} for the CG model. This energy function defines the interactions between the pseudoatoms and is chosen so as to reproduce the same thermodynamic (static) properties of the reference detailed system.
- If we are interested in dynamic properties, we must find **effective dynamical equations** for describing the time-evolution of the CG system. When we remove some degrees of freedom during coarse-graining, we remove the time scales associated with them; to properly evolve the CG system, we need to build those time scales back into the dynamical equations to get the dynamics of the CG system (approximately) correct.

The last point may seem the most non-obvious, and indeed there are limited methods for addressing it in a rigorous but practical way. However, we can get a very simple idea of the need for CG dynamical equations by considering the following example: we have a system of solutes immersed in solvent and we coarse-grain away the atomic solvent into an effective continuum one. In this case, we remove the time scales associated with the solvent degrees of freedom and this amounts for the need to build the effects of solvent viscosity back into the dynamical equations. We might then treat the solutes as interacting with Langevin-type dynamics to approximate the solvent effects on dynamics. That is, we capture the effect of the random “collisions” of solvent particles with solutes.

In what follows, we will discuss considerations for choosing the pseudoatoms and U_{CG} in more detail. Three goals are of importance:

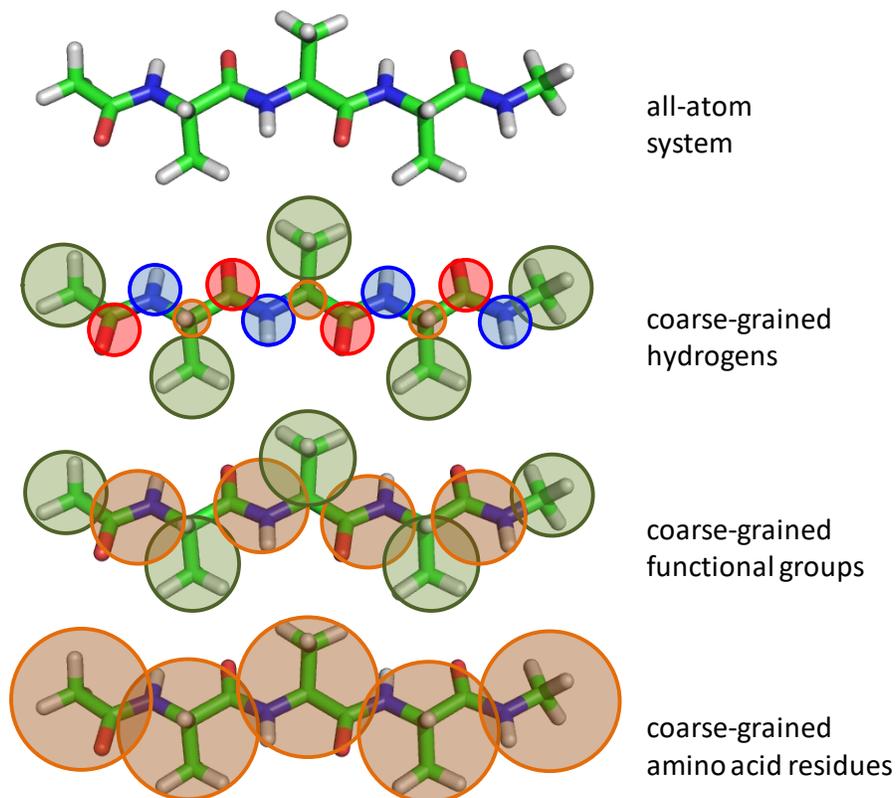
- to develop a cheaper, easier-to-simulate CG model
- to develop a CG model that has a **smoother underlying energy landscape** that is faster to equilibrate and for which we can take large time steps
- to maintain (approximately) correct physical behavior

It is important to keep in mind that we are **not guaranteed to find a CG model that adequately reproduces the properties of a detailed system**. It may simply be that the physics in the detailed system is highly sensitive to small-scale phenomena.

Pseudoatoms

Typically pseudoatoms are defined as groups of atoms of common chemistries, like methyl or carbonyl groups. Alternatively, they can contain many functional units, and such is a common strategy in polymer modeling.

There are many ways to develop pseudoatoms of varying resolution. Consider polyalanine:



The development of a pseudoatom description of a system requires the definition of a **mapping function M** . This mapping function takes as input a set of atomic coordinates in the detailed system and maps this to a unique pseudoatom configuration in the CG system. For simplicity of notation, define

$$\mathbf{R} \equiv \mathbf{r}_{CG}^N$$

$$\mathbf{r} \equiv \mathbf{r}_{AA}^n$$

where the subscript 'A' denotes the reference all-atom system. Then, the mapping function has the form:

$$\mathbf{R} = \mathbf{M}(\mathbf{r})$$

Notice that there may be more than one atomistic configuration mapping to any one coarse-grained configuration.

Typically, pseudoatoms sites are defined as the center-of-mass coordinates of groups of atoms in the all-atom representation:

$$\mathbf{R}_I = \left(\sum_{j \in \text{atoms for } I} m_j \mathbf{r}_j \right) \left(\sum_{j \in \text{atoms for } I} m_j \right)^{-1}$$

This kind of expression gives rise to CG coordinates that can always be expressed as a linear combination of atomic coordinates. In this case, the mapping function becomes a **mapping matrix**:

$$\mathbf{R} = \mathbf{M}\mathbf{r}$$

Here, \mathbf{r} and \mathbf{R} are length $3n$ and $3N$ vectors, respectively. The matrix \mathbf{M} has dimensions $(3N, 3n)$.

Usually the choice of mapping is motivated by chemical intuition and the desired level of coarsening. There are, however, very recent efforts to develop theoretical or algorithmic strategies for automating and optimizing the mapping design.

Methods for finding U_{CG}

Although we are not guaranteed to be able to find a coarse-grained representation of an all-atom system that reproduces the same physical behavior, we can posit some form of the CG model (in terms of pseudo atoms) and then search for an energy function that is optimal.

There are several methods for approaching this problem that are used in the literature. All of these require that a reference simulation of the atomistic system first be performed so that U_{CG} can be optimized to it. Why would one develop a CG model if they can simulate an atomistic system? Typically, the CG models are developed by parameterization to a very small representative atomistic system. Then, they are used to perform simulations with much larger system sizes than what would be attainable atomistically.

Fundamentals for optimal U_{CG} in the absence of constraints

How do we define “optimal” in terms of the effective interactions between pseudoatom sites? In a thermodynamic sense, there is a unique way to do this if we allow U_{CG} to take any functional form—that is, we do not specify the kinds of interactions in U_{CG} or demand that they are pairwise additive.

Consider the canonical ensemble. If we perform a simulation using U_{CG} we would like the configurational probabilities for \mathbf{R} in the CG ensemble to be equal to those that we would see in the atomistic ensemble. In other words, we would like

$$\rho_{CG}(\mathbf{R}) = \rho_{AA}(\mathbf{R})$$

Notice that in the atomistic ensemble, the probability of a coarse-grained configuration is the sum of the probabilities of all of the atomistic configurations that map to it via the mapping function,

$$\rho_{AA}(\mathbf{R}) = \int \rho_{AA}(\mathbf{r}) \delta[\mathbf{M}(\mathbf{r}) - \mathbf{R}] d\mathbf{r}$$

Here, the delta function selects out all of those configurations in the atomistic configuration space \mathbf{r} that map to the same coarse-grained configuration \mathbf{R} . We also have that the CG and atomistic configurational probabilities follow

$$\rho_{CG}(\mathbf{R}) = \frac{e^{-\beta U_{CG}(\mathbf{R})}}{Z_{CG}} \quad \rho_{AA}(\mathbf{r}) = \frac{e^{-\beta U_{AA}(\mathbf{r})}}{Z_{AA}}$$

Combining all of these expressions into our equality for the two probability distributions,

$$\frac{e^{-\beta U_{CG}(\mathbf{R})}}{Z_{CG}} = \int \frac{e^{-\beta U_{AA}(\mathbf{r})}}{Z_{AA}} \delta[\mathbf{M}(\mathbf{r}) - \mathbf{R}] d\mathbf{r}$$

We now take the logarithm of this expression. Because the probability distributions are insensitive to additive shifts in the energy functions, we will rewrite by absorbing the partition functions into a constant of normalization:

$$U_{CG}(\mathbf{R}) = -k_B T \ln \int e^{-\beta U_{AA}(\mathbf{r})} \delta[\mathbf{M}(\mathbf{r}) - \mathbf{R}] d\mathbf{r} + \text{const}$$

Notice that the term on the RHS just gives the part of the free energy corresponding to given values of \mathbf{R} . In other words, it is the **multidimensional potential of mean force** or the **free energy surface** along the reduced degrees of freedom \mathbf{R} :

$$U_{CG}(\mathbf{R}) = F_{AA}(\mathbf{R}) + \text{const}$$

In other words, the CG energy function should ultimately reproduce the underlying free energy surface along the CG degrees of freedom. If it does, all thermodynamic averages in the CG system should be equal to their values in the atomistic system since the configurational distributions will be the same.

Keep in mind, however, that there is no guarantee that $F_{AA}(\mathbf{R})$ will have a particular functional form. In particular, there is no a priori reason that it should be a pairwise additive sum of interactions between the pseudoatom sites. In actuality, it is usually not well-modeled as a pairwise sum of interactions due to the fact that the degrees of freedom that are coarse-grained away have now been included in the effective remaining interactions.

On the other hand, when we develop CG model energy functions we typically specify pairwise terms. This approximation keeps our simulations computationally efficient, but its ultimate effect is to cause $U_{CG}(\mathbf{R})$ to deviate from the correct underlying $F_{AA}(\mathbf{R})$.

Iterative Boltzmann inversion

The efforts of McGreevy & Pusztai [1988] and Lyubartsev & Laaksonen [1995] resulted in a general method for finding an optimal U_{CG} as a sum of pairwise potentials:

$$U_{CG}(\mathbf{R}) = \sum_{i < j} u_{CG}(R_{ij})$$

Their work is based on the so-called **uniqueness theorem** due to Henderson [1974]. This theorem says that, for a given pair radial distribution function $g(r)$, there is a unique underlying pair potential $u(r)$ that will produce it. The **iterative Boltzmann inversion** approach couples this observation with iterative simulations to determine effective CG potentials.

The general approach is the following:

1. Perform a reference simulation using the detailed, all-atom simulation.
2. Determine the design of the CG model and its associated pseudoatoms. For each pair of pseudoatom types, compute $g_{AA}(R)$ from the all-atom simulation. This requires converting instantaneous configurations in trajectory frames in the atomistic simulation to coarse-grained pseudoatom models.
3. Estimate the pairwise potential between each pair of pseudoatom types using a direct Boltzmann inversion:

$$u_{CG}(R) = -k_B T \ln g_{AA}(R)$$

Since the computed $g_A(R)$ will be discrete, the potential will be finely-discretized on the distance interval, or a spline approximation.

4. Perform a simulation in the canonical ensemble for the CG system using the computed pairwise $u_{CG}(R)$ potentials. Measure the pair distribution $g_{CG}(R)$ for the pseudoatoms in the CG ensemble.
5. Update the pairwise potential function using

$$u_{CG}(R) \leftarrow u_{CG}(R) - k_B T \ln \left(\frac{g_{AA}(R)}{g_{CG}(R)} \right)$$

6. Go back to step 4 and repeat a CG simulation. Continue until all $u_{CG}(R)$ converge.

Notice that the update equation in step 5 has the following behavior:

- **If $g_{AA}(R) = g_{CG}(R)$** then the CG simulation reproduces the same pairwise structure as the all-atom one. The update then leaves the CG energy function the same. In other words, one has found the unique underlying pair potential that reproduces the all-atom pairwise structure.
- **Where $g_{AA}(R) > g_{CG}(R)$,** the current CG energy function *underestimates* the correlation between pseudoatoms at r and the update equation will lower the pairwise CG energy at that distance, making the interaction more favorable there.
- **Where $g_{AA}(R) < g_{CG}(R)$,** the current CG energy function *overestimates* the correlation between pseudoatoms at r and the update equation will increase the pairwise CG energy at that distance.

Thus, this procedure systematically couples errors in the CG pairwise structure with the determination of the CG energy function.

Typically the final $u_{CG}(R)$ deviates substantially from the initial estimate determined from the atomistic function $u_{CG}(R) = -k_B T \ln g_{AA}(R)$. This is because the PMF between two pseudoatoms cannot be summed to reproduce the total free energy surface:

$$F_A(\mathbf{R}) \neq \sum_{i < j} -k_B T \ln g_{AA}(R_{ij})$$

That is, the multidimensional free energy surface cannot be decomposed into a sum of pairwise PMFs. By iterating the procedure to find $u_{CG}(R)$ that reproduces the pairwise structure, one is in effect finding a better pairwise decomposition of the total free energy surface.

The inverse MC approach has been frequently applied to CG simulations of polymeric and surfactant systems. Typically, one determines an effective interaction between CG polymer/surfactant “beads” using small systems and this approach, and then scales the system size up to perform

much larger systems (more polymers and surfactants) using the newly-determined CG energy function.

Force matching

The **force matching** approach is an alternative to the inverse MC approach that was initially developed by Ercolessi and Adams [1995] as a way to extract classical potentials from ab initio calculations. Later, Izvekov and Voth [2005] generalized this method to deriving pseudoatom CG models from atomistic classical models. Like the inverse MC approach, one performs a reference all-atom simulation and then iterates to find an optimal CG model that “matches” it in some optimal sense.

The idea of the force matching strategy is to match the average force on pseudoatoms in the CG system to that expected from the all-atom system. In equation form,

$$\begin{aligned} \langle \mathbf{F}_I \rangle_{AA} &= \mathbf{F}_{I,CG} && \text{for all pseudoatom sites } i = 1, \dots, N_{CG} \\ &= - \frac{\partial U_{CG}(\mathbf{R})}{\partial \mathbf{R}_i} \end{aligned}$$

Both sets of forces here correspond to those on the pseudoatoms for the same pseudoatom configuration. In the all-atom system, the average forces on each pseudoatom are constructed using the mapping function and propagating the individual atom forces to the pseudoatom sites via usual force-mechanical expressions. In the CG system, the CG forces are evaluated using the CG energy function for the pseudoatom configuration to which the atomistic configuration maps.

This procedure is applied to the entire reference trajectory in the all-atom system. In order to match the forces, one defines a least-squares objective function:

$$Y = \sum_{k=1}^n \sum_{i=1}^{N_{CG}} \left(\mathbf{F}_{i,AA}(\mathbf{R}_{A,k}) - \mathbf{F}_{i,CG}(\mathbf{R}_{A,k}) \right)^2$$

The first sum here is over all trajectory configurations from an all-atom ensemble. The second is over all pseudoatom sites. To determine optimal parameters in the for U_{CG} , one minimizes Y with respect to variations in all U_{CG} free parameters. For linear parameters, the minimization can be formulated as a matrix inversion problem.

Typically, a form for the CG energy function for pseudoatom sites is chosen as

$$u_{ij}(R) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{R} + f_{ij}(R)$$

Here, $f_{ij}(R)$ is a **cubic spline** function for atom types i and j . The spline function contains **knot values**, which are the values of the function f_{ij} at predetermined discrete distances in R . Then, the spline continuously interpolates (to third polynomial order) between these values at intermediate distances. A continuous function is required in order to evaluate the forces.

Importantly, the spline is linear in all of the knot values. The potential is also linear in all of the partial charge combinations $q_i q_j$ for pairs of pseudo atoms sites. The force matching approach determines the optimal values of all of the q_i and knot values by minimizing Y using the matrix inversion formalism. The net charge on the system $\sum_i q_i$ is also constrained to equal that of the atomistic system during this process.

Relative entropy optimization

One drawback of the above two approaches is that they do not provide a statistically interpretable metric for the quality of a CG model. That is, there is no way to compare the relative quality of several CG models with different numbers of pseudoatom sites each. Moreover, there are specific limitations: iterative Boltzmann inversion yields only finely discretized pairwise potentials, and the force-matching approach is limited to continuous potentials.

Shell and coworkers introduced a new framework for the optimization of CG models that is based upon a presumed universal metric for CG model quality called the **relative entropy**. This quantity stems from information theory and has the following form for discrete systems:

$$S_{\text{rel}} = \sum_m \wp_{AA}(m) \ln \left[\frac{\wp_{AA}(m)}{\wp_{CG}(m)} \right] + S_{\text{map}}$$

Here, the summation proceeds over all configurations (microstates) m in the atomistic system and the two sets of probabilities correspond to the configurational probabilities in either the atomistic or CG ensemble—that is, the probabilities one would expect if a simulation were performed using U_{AA} or U_{CG} , respectively. The term S_{map} simply measures the degeneracy of the mapping:

$$S_{\text{map}} = \ln(\text{avg \# atomistic configs mapping to a single CG config})$$

Notice that S_{map} is not a function of the coarse-grained energy function, but is completely determined by the mapping operator.

For continuous systems, one can write:

$$S_{\text{rel}} = \int \wp_{AA}(\mathbf{r}) \ln \left[\frac{\wp_{AA}(\mathbf{r})}{\wp_{CG}(\mathbf{M}(\mathbf{r}))} \right] d\mathbf{r} + S_{\text{map}}$$

The relative entropy can be derived on the basis of maximum likelihood arguments. We will not consider the derivation in further detail, although we will note that it has fundamental connections to inverse problems in statistical mechanics.

The relative entropy is an equilibrium overlap metric, and it has several properties:

- Higher values indicate less overlap of the CG model with the all-atom one.
- At model optimality, S_{rel} is at a minimum and it is zero in the case that the model perfectly reproduces the reference AA system.
- S_{rel} is always positive or zero, a mathematical consequence of its construction from two probability distributions and its connection to a log-likelihood.

Canonical ensemble

In the canonical ensemble, we can substitute

$$\wp_{CG}(\mathbf{R}) = \frac{e^{-\beta U_{CG}(\mathbf{R})}}{Z_{CG}} \quad \wp_{AA}(\mathbf{r}) = \frac{e^{-\beta U_{AA}(\mathbf{r})}}{Z_{AA}}$$

to find that

$$S_{\text{rel}} = \beta \langle U_{CG} - U_{AA} \rangle_{AA} - \beta (A_{CG} - A_{AA}) + S_{\text{map}}$$

The first term on the RHS is an average performed in the all-atom ensemble, i.e., using an all atom trajectory.

The relative entropy approach says that we should find the optimal value of any parameters in our CG energy function by minimizing S_{rel} , with respect to some reference all-atom simulation trajectory. Consider the optimization of a single parameter λ , which could be a Lennard-Jones parameter ϵ or σ , or a partial charge, for example. At a minimum, we demand that

$$\begin{aligned} 0 &= \frac{\partial S_{\text{rel}}}{\partial \lambda} \\ &= \beta \frac{\partial \langle U_{CG} \rangle_{AA}}{\partial \lambda} - \beta \frac{\partial A_{CG}}{\partial \lambda} \\ &= \beta \left\langle \frac{\partial U_{CG}}{\partial \lambda} \right\rangle_{AA} - \beta \left\langle \frac{\partial U_{CG}}{\partial \lambda} \right\rangle_{CG} \end{aligned}$$

The last term on the last line stems from the result we found in our lecture on thermodynamic integration: the derivative of the free energy with respect to any energy function parameter simply returns the average derivative of the energy function with respect to that parameter.

Notice that this expression entails two averages of the same derivative of the CG energy function: one in the CG ensemble and one in the atomistic ensemble. At optimality, the S_{rel} approach demands that these two averages be equal.

Functional optimization in the canonical ensemble

Consider the case in which U_{CG} can take any form, and is not constrained to be pairwise additive or have any other functional dependence. In this case, the optimal U_{CG} is the one that functionally minimizes the relative entropy:

$$\frac{\delta S_{\text{rel}}[U_{CG}(\mathbf{R})]}{\delta U_{CG}(\mathbf{R}')} = 0$$

In an approximate interpretation, this expression says the following: the relative entropy is a functional of the function $U_{CG}(\mathbf{R})$. Find the variation in the relative entropy with respect to changing the value of the function U_{CG} at a specific set of coarse-grained coordinates \mathbf{R}' . Set this variation equal to zero at a minimum. The result will apply to all sets \mathbf{R}' .

To perform the functional derivative, we express S_{rel} in terms of configurational integrals:

$$S_{\text{rel}} = \beta \int \wp_{AA}(\mathbf{r}) [U_{CG}(\mathbf{M}(\mathbf{r})) - U_{AA}(\mathbf{r})] d\mathbf{r} + \ln \int e^{-\beta U_{CG}(\mathbf{R})} d\mathbf{R} - \ln \int e^{-\beta U_{AA}(\mathbf{r})} d\mathbf{r} + S_{\text{map}}$$

When we take the derivative, the only integrals with terms that are dependent on U_{CG} are the first and second on the RHS. Note that the mapping entropy in particular is independent of U_{CG} (e.g., it only depends on the choice of pseudoatoms and the mapping function). It can then be shown that the functional derivative of the relative entropy gives [Chaimovich and Shell, 2011],

$$\frac{\delta S_{\text{rel}}[U_{CG}]}{\delta U_{CG}} = \beta \int \wp_{AA}(\mathbf{r}) \delta[\mathbf{M}(\mathbf{r}) - \mathbf{R}'] d\mathbf{r} - \wp_{CG}(\mathbf{R}')$$

Therefore we have that,

$$\wp_{AA}(\mathbf{R}') = \wp_{CG}(\mathbf{R}')$$

This is our key result. It says that relative entropy minimization for an unconstrained U_{CG} rigorously returns configurational probabilities in the CG ensemble equal to those that would be expected in the all-atom ensemble. This identity also means that the optimal U_{CG} (for which S_{rel} is at a minimum) corresponds to the case in which it reproduces the free energy surface in the all-atom ensemble along the reduced degrees of freedom \mathbf{R} :

$$U_{CG}(\mathbf{R}) = -k_B T \ln \int e^{-\beta U_{AA}(\mathbf{r})} \delta[\mathbf{M}(\mathbf{r}) - \mathbf{R}] d\mathbf{r} + \text{const}$$

This proof therefore shows that the relative entropy fundamentally demands equality of U_{CG} with the underlying all-atom free energy surface.

In this case, we considered a fully unconstrained U_{CG} . However, if we constrain its form, perhaps by demanding that it be a sum of pairwise interactions, then the equality above will not hold and the relative entropy will attain a higher value than its minimum possible. In any case, S_{rel} minimization always strives to find a U_{CG} of the given functional form that reproduces, as best as possible, the reference probabilities $\wp_{AA}(\mathbf{R})$.

Numerical implementation

One can optimize CG models using a Newton-Raphson approach for minimizing S_{rel} :

$$\begin{aligned} \lambda^{k+1} &= \lambda^k - \chi \frac{\left(\frac{\partial S_{rel}}{\partial \lambda}\right)}{\left(\frac{\partial^2 S_{rel}}{\partial \lambda^2}\right)} \\ &= \lambda^k - \chi \left[\frac{\left\langle \frac{\partial U}{\partial \lambda} \right\rangle_{AA} - \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_{CG}}{\left\langle \frac{\partial^2 U}{\partial \lambda^2} \right\rangle_{AA} - \left\langle \frac{\partial^2 U}{\partial \lambda^2} \right\rangle_{CG} + \left\langle \left(\frac{\partial U}{\partial \lambda} \right)^2 \right\rangle_{CG} - \beta \left(\left\langle \frac{\partial U}{\partial \lambda} \right\rangle_{CG} \right)^2} \right] \end{aligned}$$

Here, k is an iteration index for optimizing successive values of the parameter λ and χ is a mixing parameter between 0 and 1 that ensures stability. This equation implies an iteration over different values of the parameter that converges to an optimal value at a relative entropy minimum. At each iteration,

- averages in the all-atom ensemble $\langle \dots \rangle_{AA}$ are computed using the pre-saved trajectory from a single, reference all-atom simulation
- averages in the CG ensemble $\langle \dots \rangle_{CG}$ are computed using a short simulation in the canonical ensemble using the current value of the CG energy function, i.e., using U_{CG} for the current values of the parameters.

Though this approach implies iterative simulations of the CG model, this system is by design fast to simulate.

Absolute relative entropies

Importantly, we can use the value of the relative entropy as a universal metric for the quality of a coarse-grained model, regardless of the number of pseudoatom sites or the functional form of U_{CG} . To compute the value of S_{rel} we can reformulate the free energy difference in the above expression using standard free energy perturbation (FEP) techniques,

$$-\beta(A_{CG} - A_{AA}) + S_{\text{map}} = \ln \langle e^{-\beta(U_{CG} - U_{AA})} \rangle_{AA}$$

Here, the inclusion of S_{map} stems from extending the FEP relation to systems with different numbers of degrees of freedom. We can now write,

$$\begin{aligned} S_{\text{rel}} &= \beta \langle U_{CG} - U_{AA} \rangle_{AA} + \ln \langle e^{-\beta(U_{CG} - U_{AA})} \rangle_{AA} \\ &= \ln \langle e^{-\beta(U_{CG} - U_{AA}) + \beta \langle U_{CG} - U_{AA} \rangle_{AA}} \rangle_{AA} \end{aligned}$$

It is convenient to define a dimensionless energy difference

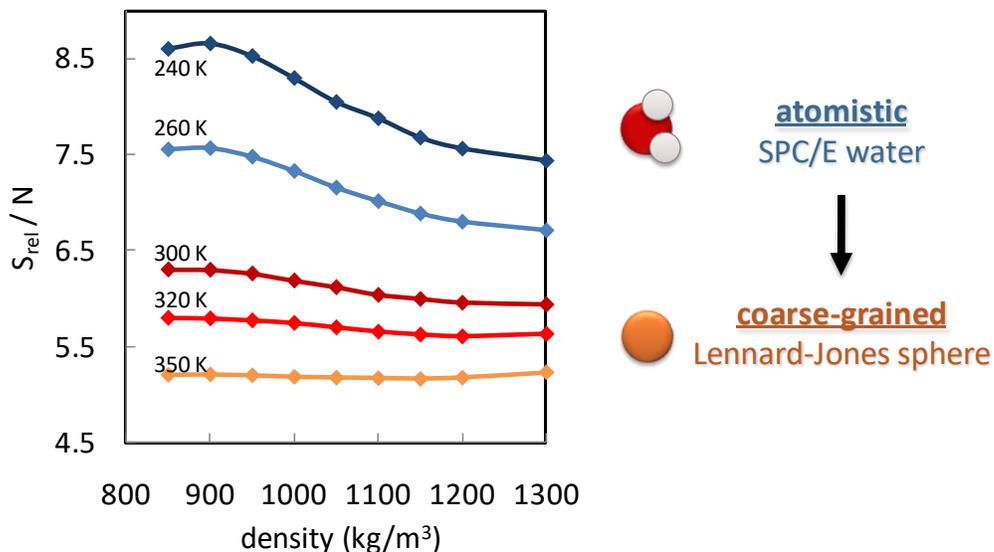
$$\Delta \equiv \beta(U_{AA} - U_{CG})$$

Then,

$$S_{\text{rel}} = \ln \langle e^{\Delta - \langle \Delta \rangle_{AA}} \rangle_{AA}$$

This equation shows that the relative entropy measures fluctuations in differences in the energies between the CG and atomistic ensemble. Importantly, it enables the absolute calculation of S_{rel} . Using an atomic trajectory, one computes the value of A for every configuration and then performs an average for this equation.

How well does this approach work? The following graph from [Shell, JCP, 2008] shows the value of the relative entropy when a three-site, all-atom model of liquid water is coarse-grained into single-site Lennard-Jones center:



This graph shows the value of the relative entropy at different state points: for each, the minimum value of S_{rel} has been found by adjusting the Lennard-Jones ϵ and σ parameters. Notice

that the value of the relative entropy increases sharply at low temperatures and low densities. This indicates that the Lennard-Jones model is increasingly inadequate as a model for liquid water at these conditions. This behavior is consistent with the fact that liquid water can form open, tetrahedrally-coordinated structures at these conditions due to hydrogen bonding interactions, which lends it very different behavior than of simple liquids like the Lennard-Jones system.