Statistics C206B Lecture 5 Notes

Daniel Raban

February 1, 2022

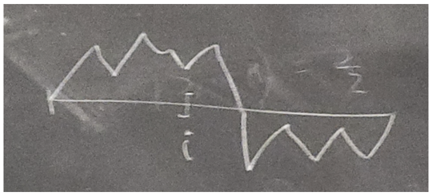
1 Techniques for Bounding Mixing Times of Particle Systems

1.1 Monotone chains

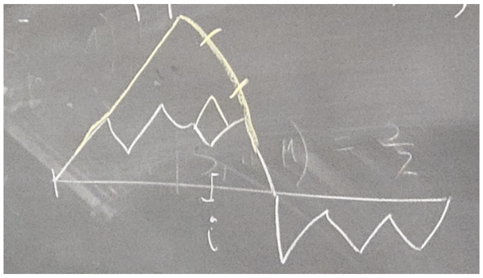
Previously, we were reviewing general Markov chain mixing time theory. Our goal now is to understand mixing times of the exclusion process in 1 dimension. For now, we will discuss monotone chains.

Roughly, a **monotone chain** is a Markov chain where there is a partial order on the state space that the dynamics preserves. If $X \leq Y$ in the state space, then there is a coupling of X_1 and Y_1 such that $X \leq Y$ a.s. If we want to characterize these distributional conditions in terms of expectations, we will need to check the expectations with respect to all increasing functions.

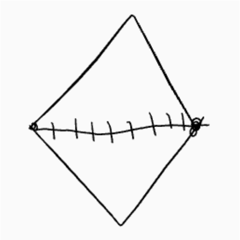
Example 1.1. The exclusion process is a monotone chain. Think about this in terms of the associated height functions. Recall that every particle configuration σ corresponds to a height function h:



In this case, we can define monotonicity by $h_1 \leq h_2$ if this domination occurs pointwise. Pick the coupling where we update the same site for both chains. If there is a large gap, we can see that there is no change in domination:



So the only issue is a case like this:

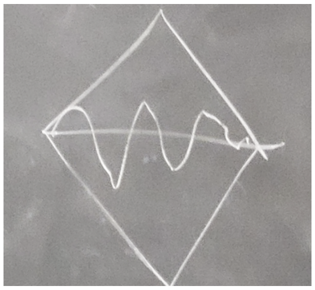


In this case, we can couple them so that the monotonicity is preserved. This is because the probability of each outcome at this site is 1/2 for each of the chains. This requires laziness of the chain.

Example 1.2. The Ising model is also a monotone chain. Suppose we have configurations $\sigma_1, \sigma_2 \in \{\pm 1\}^V$. We can let $\sigma_1 \leq \sigma_2$ if $\sigma_1(v) \leq \sigma_2(v)$ for all $v \in V$. For the coupling, pick the same site to update for both chains. By the monotonicity, the probability that the dominant chain makes the site $+$ is $\geq$ the same probability for the less dominant chain. Then, we can use the same coupling trick from before with Bernoulli random variables to couple them appropriately so that if the less dominant chain labels b as $+$, then so does the dominant chain.

Often, monotone chains have a unique maximum element and a unique minimum element.

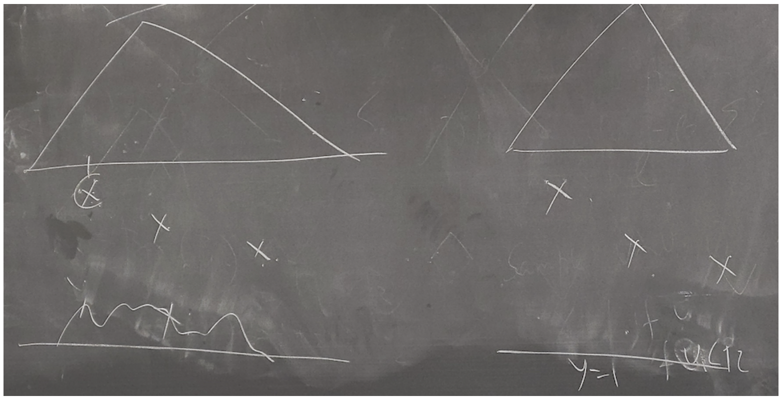
Example 1.3. Here are the maximum and minimum states for the exclusion process:



If we can show a coupling where starting from the maximum and the minimum state, the chains meet, then this tells us that the coupling starting from any state will coalesce by being sandwiched between the max and min. In particular, the Markov chain admits a grand coupling of the evolution started from all possible states. Under this coupling, the coalescence time for the max and the min can be used to upper bound the mixing time.

1.2 Censoring inequality

Suppose that in the exclusion process, we fix a sequence of sites to update but keep the updates at each site random. This still gives a distribution on the possible height functions at each time step.



If we repeat the sequence of sites to update but suppress some of the moves, how can we compare these distributions? The unsuppressed distribution will be closer in distribution to equilibrium.

Here is the most basic version of the censoring inequality:

Theorem 1.1 (Censoring inequality)**.** *Consider a "monotone Markov chain,"*[1] *and consider μ and ν to be the distributions starting from the maximal state with updates at a deterministic sequence and ν obtained after removing a deterministic subsequence. Then*

$$\|\mu - \pi\|_{\mathrm{TV}} \leq \|\nu - \pi\|_{\mathrm{TV}}.$$

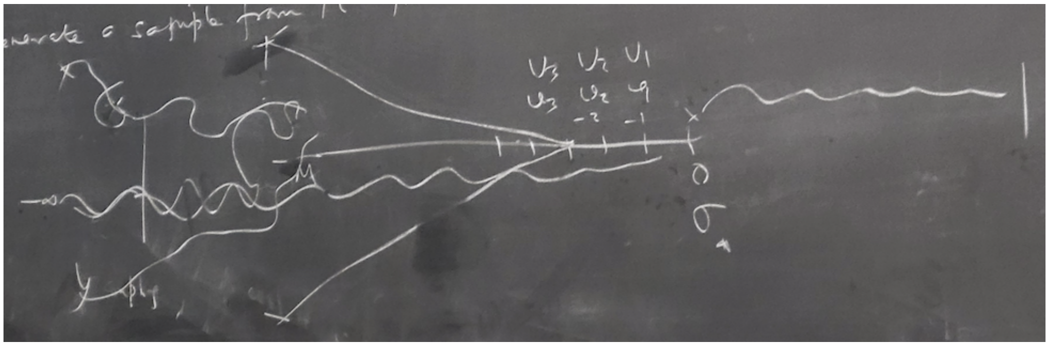
1.3 Coupling from the past

The goal here is to generate a sample from π for, say, the Ising model. If you run the chain for a long time, you can sample from approximately the equilibrium measure. However, what if we want to exactly sample? The key idea is that if the chain started at time $-\infty$, then at time 0, the chain will have run for infinite time and should be at π.

---

[1] We talk about them as if they have an underlying lattice, which is not true of all monotone chains. Nevertheless, this property can be expressed more abstractly for other types of chains.

We set up a grand coupling, starting from every state in the state space, started at some time in the distant past. If we look far enough into the past, then at time 0, the coupling will have coalesced and we can sample the distribution at time $t = 0$ to get the stationary distribution.



This does not require the chain to be reversible. We are running the chain forwards in time from a large negative time. What makes this different from starting the chain at 0 and waiting for it to coalesce? The key difference is that the time we start the chain at is random, but the time we sample from is deterministic.

1.4 Interchange process on the line

Suppose we have a line of length n and k particles. We can deal with a related process to the exclusion process. The state space for the exclusion process is $\{x \in \{0,1\}^{[n]} : \sum_i x_i = k\}$. The **interchange process** is the lift of the exclusion process, where we label each site with one of the numbers $1, \ldots, n$ in a permutation. This has state space S_n. We can treat this as a lift of the exclusion process by labeling k of the numbers as particles.

In general, if a Markov chain is a projection of another, the spectral gap of the projection at least as big as the spectral gap of the lift. Thus, $\gamma^{\text{IP}} \leq \gamma^{\text{EP}}$. Since the random walk is the exclusion process with $k = 1$, $\gamma^{\text{IP}} \leq \gamma^{\text{RW}}$. Aldous' spectral gap conjecture (which is now a theorem) is that $\gamma^{\text{IP}} = \gamma^{\text{RW}}$ for any graph. Since $\gamma^{\text{EP}} \leq \gamma^{\text{RW}}$, this gives us a tool to study γ^{EP} and therefore bound the mixing time.