Lecture 36 - Analysis of Riffle Shuffle II

Scribe: Sam Greenberg

As stated before, our shuffling algorithm gives each card an additional random bit, eventually ordering the cards. More formally, we create a matrix of n rows. At each step, we add a column of 0s and 1s. We stop when each card has a distinct label.

Lemma 36.1. *This is a strong stopping time.*

Proof. Consider what happens after one step. All the cards labeled 0 are on the top, and all the cards labeled 1 are on the bottom.

After two steps, the order goes $(0,0),(1,0),(0,1),(1,1)$

Once labels are distinct, the order is uniquely determined. Since the deck is symmetric, any permutation is equally likely, so any ordering is equally likely. $\square$

For a mixing time bound, remember that

$$||p^t - \pi||_{TV} = ||p^{\star t} - \pi|| \leq \mathsf{P}[T > t]$$

Now, $T > t$ if, after t steps of the inverse shuffle, two cards have the same labelling. There are 2^t t-bit strings and n cards. Hence

$$\begin{aligned}\mathsf{P}[T \leq t] &= \mathsf{P}[\textit{all distinct labels}] \\ &= \prod_{i=1}^{n-1} 1 - \frac{i}{2^t}\end{aligned}$$

That means

$$\begin{aligned}||p^{\star t} - \pi||_{TV} &= \mathsf{P}[T > t] \\ &= 1 - \prod_{i=1}^{n-1} 1 - \frac{i}{2^t} \\ &= e^{\sum \log(1 - \frac{i}{2^t})} \\ &\approx e^{-\sum \frac{i}{2^t}} \\ &= e^{-\frac{\frac{n(n-1)}{2}}{2^t}}\end{aligned}$$

If $t = 2\log(\frac{2}{c})$, then

$$\begin{aligned}||p^t - \pi||_{TV} &\leq \frac{\frac{n(n-1)}{2}}{2^t} \\ &= \frac{n(n-1)}{2\left(\frac{n}{c}\right)^2} \\ &\approx \frac{c^2}{2}\end{aligned}$$

This means it takes 11 shuffles to bring a 52 card deck to distance $\frac{1}{2}$.

Bayer and Diaconis found the exacty total variation distance of the riffle shuffle. If $t = \log(\frac{3}{2}\log(n) + c$, then

$$||\cdot||_{TV} = 1 - 2\Phi\left(\frac{-2^{-c}}{4\sqrt{3}}\right) + O\left(\frac{1}{n^{\frac{1}{4}}}\right)$$

where $\Phi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{\frac{-t^2}{2}}\,dt$.

This calculation gives total variation distance 0.924 after 5 shuffles, but 0.614 after 6, and 0.334 after 7. It drops down further after that. This was the published result that seven shuffles suffices for a 52 card deck.

We now turn to random walks on the truncated hypercube.

Let $\Omega = \{x \in [0,1]^n : \sum_{i=1}^{n} a_i x_i \leq b\}$ for fixed a and b. We use the lazy max-degree walk to sample from Ω:

1. Choose the direction $i \in 1, \ldots, n$ uniformly at random.
2. Half of the time, do nothing.
3. The other half of the time, change coordinate i if doing so keeps us in Ω.

We might try canonical paths, but we have to be careful we stay in Ω. Next lecture we'll talk about bounding away the distance from the border, while still keeping conductance low.