Lecture 7: September 14

Lecturer: Prof. Eric Vigoda *Scribes: Yury Park (babyseal33@gatech.edu)*

Topics for the next couple weeks:

- Introduction to Conductance
- Spectral Gap

7.1 Coloring

We will continue our discussion of coupling and mixing time from last week, this time using the (slightly more complicated) example of coloring.

Given k different colors and an undirected graph $G = (V, E)$ whose maximum-degree vertex has degree $\triangle$, a valid coloring is defined as the assignment of a color to each vertex such that no adjacent vertices have the same colors. Formally, we can define a coloring as:

$$\sigma : V \mapsto \{1, ..., k\}, \text{ where } \forall (v, w) \in E, \sigma(v) \neq \sigma(w).$$

If we define Ω as the set of all valid k-colorings, figuring out $|\Omega|$ is a $\#P$-complete problem. We can, however, design a Markov Chain and a coupling scheme to get a rapid mixing time, thus allowing us to sample u.a.r. from Ω. We will show that rapid mixing can be obtained if $k > 2\triangle$. We will do this in two stages: (1) use identity coupling to show rapid mixing time when $k > 3\triangle$, then (2) use path coupling to accomplish the same when $k > 2\triangle$.

7.2 Markov Chain

As always, we will start by designing an ergodic Markov Chain whose distribution is uniform over Ω. The transition matrix P will adhere to the following scheme:

From any valid state $X_t \in \Omega$,

1. Choose $v \in V$ u.a.r. and a color $c \in \{1, ..., k\}$ u.a.r.
2. For all $w \in V \setminus v$, set $X_{t+1}(w) = X_t(w)$.
3. $X_{t+1}(v) = \begin{cases} c & \text{if } c \notin X_t(N(v)), \text{ where } N(v) = \text{ set of all neighboring vertices to } v \\ X_t(v) & \text{otherwise.} \end{cases}$

The chain is aperiodic (since $P(\sigma, \sigma) > 0$, as outlined above) and irreducible when $k \geq \triangle + 2$. (To see why we cannot establish irreducibility when $k = \triangle + 1$, consider a complete graph such as K_3 with $k = 3$, and notice that it is impossible to transition from one state to another.) When $k \geq \triangle + 2$, we can show irreducibility in the following way. Let $X, Y \in \Omega$ be any two valid states (i.e. two valid colorings). We can transition (eventually) from X to Y by arbitrarily ordering the vertices $v_1, ..., v_n$, and then sequentially re-coloring the vertices in X one by one to match the corresponding vertices in Y. If there is any conflict along the way for a vertex v_i in X such that coloring v_i to match the corresponding vertex in Y would cause X's coloring to be invalid, then it must be because v_i is somehow "blocked" by at least one neighboring vertex v_j in X where $j > i$. But since $k \geq \triangle + 2$, it is always possible to temporarily re-color v_j with some other valid color so as to allow v_i to be colored as needed, and then continue on to v_{i+1}. (Note that it is never necessary to re-color some other neighboring vertex v_l where $l < i$ to resolve the conflict for v_i, because the coloring for $v_1, ..., v_i$ in X is valid by virtue of the fact that Y is a valid coloring.)

Since the chain is aperiodic and irreducible, it is ergodic. Additionally, for any pair of states $(i, j) \in \Omega$, $P(i, j) = P(j, i) = \frac{1}{nk}$ if it is possible to transition from i to j in a single time step, and 0 otherwise. Thus, π is unique and has a uniform distribution. We will now proceed with identity coupling to show that when $k > 3\triangle$, $T_{mix} = O(n \log n)$.

7.3 Identity Coupling

Let X_t, Y_t be a pair of colorings at time t. Choose, u.a.r, the same vertex v and same color c for both X_t and Y_t, and for each state re-color v with c only if doing so would be legal. It should be clear that this coupling scheme will sometimes result in a "bad" outcome, i.e. one in which the number of disagreeing vertices increases after a time step. Let's make this more concrete by defining the concept of "agreeing" and "disagreeing" vertices at time t:

$$A_t = \{v : X_t(v) = Y_t(v)\}$$
$$D_t = \{v : X_t(v) \neq Y_t(v)\}$$

Let us now compute the probability of such a "bad" outcome. The only way this outcome can occur is if $X_t(v) = Y_t(v)$ and the color c is legal for exactly one of the two states. There are a total of $|A_t|$ vertices that agree at time t, and assuming that $v \in A_t$, there are at most $2\delta_t(v)$ colors that would cause the "bad" outcome, where $\delta_t(v) = |D_t \cap N(v)|$. In other words, the number of "bad" colors is bounded from above by twice the number of disagreeing neighbors of v at time t. As an illustrative example, Fig. 7.1 below shows that there are a total of 4 disagreeing neighbors of vertex v between the two states, and $2(4) = 8$ bad colors to choose from. It should be clear that in this example, the number of bad colors cannot exceed 8 no matter how the neighboring vertices were colored.
Therefore, in general the total number of "bad" outcomes from one time step to the next is $\sum_{v \in A_t} 2\delta_t(v)$, and thus

$$Pr(|D_{t+1}| = |D_t| + 1) = \frac{1}{nk} \sum_{v \in A_t} 2\delta_t(v)$$

where the normalizing factor $\frac{1}{nk}$ reflects the fact that we have a total of n vertices and k colors to choose from.

Let us now compute the probability of a "good" outcome, where $|D_{t+1}| = |D_t| - 1$. A good outcome occurs when $v \in D_t$ and c is legal for both states (i.e. no neighboring vertex in either X_t or Y_t contains color c). The number of such good outcomes given $v \in D_t$ is bounded from below by: $\sum_{v \in D_t} (k - 2\triangle + a_t(v))$, where $a_t(v) = |A_t \cap N(v)|$. To convince yourself this lower bound is correct, consider Fig. 7.2 and note that there are a total of $k - 2(4) + 2$ valid colors to choose from (as Y, B, M, R, O, and C are the only "bad" colors),

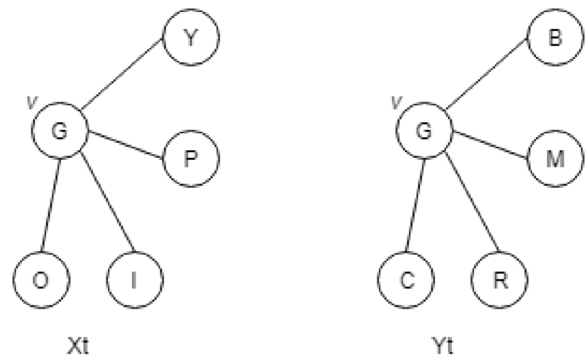

Figure 7.1: *An example of X_t and Y_t, with vertex v as shown and the capital letters indicating the color associated with each vertex.*

and that this lower bound would remain the same if we were to change the color of a neighboring vertex from its current color M to R, such that in each state two neighboring vertices now have R as their color, and the number of "bad" colors has shrunk from 6 to 5 (Y, B, R, O, and C).

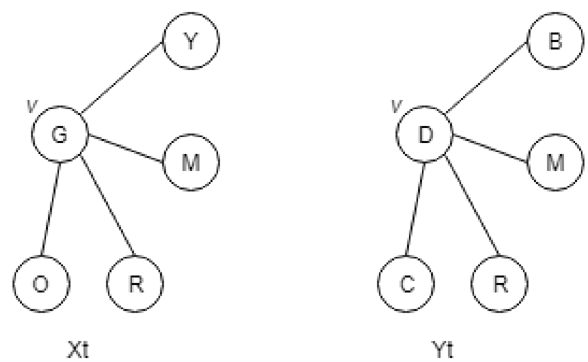

Figure 7.2: *Another example of X_t and Y_t, with slightly different coloring.*

Hence, we can compute the probability of a "good" outcome:

$$Pr(|D_{t+1}| = |D_t| - 1) \geq \frac{1}{nk} \sum_{v \in D_t} k - 2\triangle + a_t(v)$$

We are now ready to compute an upper bound on the expected value of $|D_{t+1}|$:

$$\begin{aligned}
E[|D_{t+1}] &\leq |D_t| + \frac{1}{nk} \sum_{v \in A_t} 2\delta_t(v) - \frac{1}{nk} \sum_{v \in D_t} (k - 2\triangle + a_t(v)) \\
&\leq |D_t| + \frac{1}{nk} \left[\sum_{v \in A_t} 2\delta_t(v) - \sum_{v \in D_t} 2a_t(v) - \sum_{v \in D_t} (k - 3\triangle) \right] && \text{since } 2\triangle - a_t(v) \leq 3\triangle - 2a_t(v) \\
&\leq |D_t| + \frac{1}{nk}(0 - |D_t|) && \text{since } \sum_{v \in A_t} 2\delta_t(v) = \sum_{v \in D_t} 2a_t(v), \text{ and } k - 3\triangle \geq 1 \\
&= |D_t| \left(1 - \frac{1}{nk} \right).
\end{aligned}$$

Now recall that $Pr(X_t \neq Y_t) \leq E[|D_t|]$, and that by induction $E[|D_t|] \leq |D_0|\big(1 - \frac{1}{nk}\big)^t$. Therefore,

$$Pr(X_t \neq Y_t) \leq E[|D_t|] \leq |D_0|\left(1 - \frac{1}{nk}\right)^t \leq ne^{\frac{-t}{nk}} \leq \frac{1}{4} \text{ for } t \geq nk \text{ log } (4n).$$

We have thus achieved a coupling time T_{couple} with high confidence $\geq \frac{3}{4}$, and by extension shown that T_{mix} is bounded asymptotically by $O(nk \text{ log } n)$ in the event where $k > 3\triangle$. As a reminder, what we have shown is related to mixing time simply by assuming that Y_0 is sampled from the stationary distribution π (which is allowed since Y_0 can be any arbitrary starting state), and we are assured with a high probability that $X_t = Y_t$, meaning the total variation distance between X_t, Y_t is $\leq \frac{1}{4}$.

Next, we will improve upon our result by letting $k > 2\triangle$ and showing via a different coupling scheme the same upper bound for T_{mix}.

7.4 Path Coupling

We will assume the same Markov Chain as before. Recall that $|D_t|$=the total number of disagreeing vertices between states X_t, Y_t at time t. We can equivalently think of this as the Hamming Distance $H(X_t, Y_t)$. It turns out that if we can devise an improved coupling scheme for the situation where $H(X_t, Y_t) = 1$, then we can use this to derive a better coupling for any arbitrary pair of states.

Let us consider a pair of states X_t, Y_t where $H(X_t, Y_t) = 1$ and z is the sole disagreeing vertex. As before, we stipulate that a vertex v and color c are each chosen uniformly at random. We will analyze three possible cases:

1. $v \notin z \cup N(z)$. In this case, we know $v \in A_t$ and that $v \in A_{t+1}$ since the chosen color c will be either legal for both states or illegal for both states. Therefore, $H(X_{t+1}, Y_{t+1}) = 1$. This is neither a "bad" nor a "good" outcome.

2. $v = z$. In this case a "bad" outcome isn't possible, and there are at least $k - \triangle$ "good" color choices which would cause $X_{t+1} = Y_{t+1}$.

3. $v \in N(z)$. In this case a "good" outcome isn't possible, and there are at most 2 "bad" color choices which would cause $H(X_{t+1}, Y_{t+1}) = H(X_t, Y_t) + 1$.

In general, then, the probability of "good" and "bad" outcomes provided that $H(X_t, Y_t) = 1$ is as follows:

$$Pr(|D_{t+1}| = 0 \mid H(X_t, Y_t) = 1) \geq \frac{k - \triangle}{nk}$$
$$Pr(|D_{t+1}| = 2 \mid H(X_t, Y_t) = 1) \leq \frac{2\triangle}{nk}$$

Thus, if we make no changes to our previous identity coupling scheme and maintain the constraint of $k > 3\triangle$, we can verify the conditional expectation $E[|D_{t+1}| \mid |D_t| = 1] \leq |D_t| - \frac{k-\triangle}{nk} + \frac{2\triangle}{nk} = 1 - \frac{1}{nk}(k - 3\triangle) \leq 1 - \frac{1}{nk}$. But now we can let $k > 2\triangle$ by improving our coupling for case 3 above: Assume that $v \in N(z)$ and the two "bad" colors are $\{R, B\}$. If the randomly chosen color $c \in \{R, B\}$, then let $X_{t+1}(v) = c$ and $Y_{t+1}(v) \in \{R, B\} \setminus \{c\}$ provided the coloring is legal, otherwise let $X_{t+1}(v) = X_t(v)$ and $Y_{t+1}(v) = Y_t(v)$. Notice that with this new coloring scheme, we still preserve the marginal randomness for the individual states (when looked at isolation, X_t and Y_t are each a faithful copy of the original Markov Chain that adheres to the transition matrix P), but now there is only one "bad" color choice when $v \in N(z)$. This means that the conditional expected value is now:

$$E[|D_{t+1}| \mid |D_t|=1] \leq |D_t| - \frac{k-\triangle}{nk} + \frac{\triangle}{nk} = 1 - \frac{1}{nk}(k-2\triangle) \leq 1 - \frac{1}{nk}$$

where $k > 2\triangle$.

The good news is that we can use the above result to devise a coupling for other pairs of states with Hamming Distances greater than 1. Suppose we have a pair of states $(X, Y) \in \Omega^2$ where $H(X, Y) = l, l > 1$. We can define a sequence $W_0, W_1, ..., W_l \in \Omega$ where:

(a) $\forall i, H(W_{i-1}, W_i) = 1$.

(b) $W_0 = X, W_l = Y$.

Essentially, we have constructed a "shortest path" from X to Y where all of the intermediate states W_i are such that adjacent states are apart by one Hamming Distance. Recall the coupling scheme we previously devised and notice that it can simply be thought of as a function: given a set of colors to be assigned to state X_t, we can figure out the set of colors to be assigned to state Y_t. Now it is clear that for every $i \leq l$, there is a coupling for (W_{i-1}, W_i). We can thus compose couplings (similar to how we can compose functions) along the path $W_0, W_1, ..., W_l$, as follows:

- Map $(W_0 = X, W_1)$ to $(W'_0 = X', W'_1)$ according to the coupling, where $W_0 \to W'_0$ is a random transition, and the same is true for $W_1 \to W'_1$.
- For each $i \geq 1$, map (W_i, W_{i+1}) to (W'_i, W'_{i+1}) in accordance to the coupling, conditional on W'_i already having been chosen.

Continue this process all the way to $W_l = Y$ and $W'_l = Y'$. Then we have constructed a coupling for $(X, Y) \to (X', Y')$ via composition. It remains to be proven how good this coupling is in terms of the expected change in Hamming Distance:

$$\begin{aligned}
E[H(X', Y')] &\leq E\left[\sum_{i=1}^{l} H(W'_{i-1}, W'_i)\right] \\
&= \sum_{i=1}^{l} E[H(W'_{i-1}, W'_i)] \\
&\leq \sum_{i=1}^{l} \left(1 - \frac{1}{nk}\right) && \text{for } k > 2\triangle \\
&= l\left(1 - \frac{1}{nk}\right) \\
&= H(X, Y)\left(1 - \frac{1}{nk}\right)
\end{aligned}$$

In other words, we have just shown that with the above path coupling, $E[|D_{t+1}|] = |D_t|\left(1 - \frac{1}{nk}\right)$, and by extension $E[|D_t|] \leq n\left(1 - \frac{1}{nk}\right)^t$, meaning that once again we have

$$Pr(X_t \neq Y_t) \leq n\left(1 - \frac{1}{nk}\right)^t \leq \frac{1}{4} \text{ for } t \geq nk \ln(4n), k > 2\triangle.$$

We finish by stating the Path Coupling Theorem (Bubley, Dyer '97):

Theorem 7.1 *For a finite ergodic Markov chain* $\in \Omega$, *let* $S \subseteq \Omega \times \Omega$ *such that* (Ω, S) *is connected. For* $(X, Y) \in \Omega \times \Omega$, *let* $dist(X, Y)$ = *length of shortest path between* X *and* Y *in* (Ω, S). *If there exists* $\beta < 1$ *such that* $\forall (X_t, Y_t) \in S$, *there exists a coupling* $(X_t, Y_t) \rightarrow (X_{t+1}, Y_{t+1})$ *where* $E[dist(X_{t+1}, Y_{t+1})] \leq \beta dist(X_t, Y_t)$, *then* $T_{mix}(\epsilon) \leq \frac{log(D_{max}/\epsilon)}{1-\beta}$, *where* $D_{max} = \max_{(X,Y) \in \Omega^2} (dist(X, Y))$.