Coupling Part 2

Guy Oren

December 29, 2016
Outline

Reminders
  Markov Chain Coupling
  Mixing Time
  Bounding TV By Coupling

Warm Up Example - Winning Streak

Random Coloring

Random Colorings - Revisited
Markov Chain Coupling
Markov Chain Coupling - Stickyness

\[ \text{if } X_s = Y_s, \text{ then } X_t = Y_t \text{ for } t \geq s \]
Mixing Time

\[ d(t) = \max_{x \in \Omega} \| P^t(x, \cdot) - \pi \|_{TV} \]  

(2)
Mixing Time

\[ d(t) = \max_{x \in \Omega} \| P^t(x, \cdot) - \pi \|_{TV} \]  \hspace{1cm} (2)

\[ \bar{d}(t) = \max_{x, y \in \Omega} \| P^t(x, \cdot) - P^t(y, \cdot) \|_{TV} \]  \hspace{1cm} (3)

\[ t_{mix}(\epsilon) = \min \{ t \geq 0 \mid d(t) \leq \epsilon \} \]  \hspace{1cm} (5)
Mixing Time

\[ d(t) = \max_{x \in \Omega} \| P^t(x, \cdot) - \pi \|_{TV} \]  \hspace{1cm} (2)

\[ \tilde{d}(t) = \max_{x, y \in \Omega} \| P^t(x, \cdot) - P^t(y, \cdot) \|_{TV} \]  \hspace{1cm} (3)

\[ d(t) \leq \tilde{d}(t) \]  \hspace{1cm} (4)
Mixing Time

\[ d(t) = \max_{x \in \Omega} \| P^t(x, \cdot) - \pi \|_{TV} \quad (2) \]

\[ \bar{d}(t) = \max_{x, y \in \Omega} \| P^t(x, \cdot) - P^t(y, \cdot) \|_{TV} \quad (3) \]

\[ d(t) \leq \bar{d}(t) \quad (4) \]

\[ t_{mix}(\epsilon) = \min \{ t \geq 0 \mid d(t) \leq \epsilon \} \quad (5) \]
Mixing Time

\[ d(t) = \max_{x \in \Omega} \| P^t(x, \cdot) - \pi \|_{TV} \]  
\[ \tilde{d}(t) = \max_{x, y \in \Omega} \| P^t(x, \cdot) - P^t(y, \cdot) \|_{TV} \]  
\[ d(t) \leq \tilde{d}(t) \]  
\[ t_{mix}(\epsilon) = \min\{ t \geq 0 \mid d(t) \leq \epsilon \} \]  
\[ t_{mix} = t_{mix}(\frac{1}{4}) \]
**Theorem**
Let \( \{(X_t, Y_t)\} \) be a coupling satisfying stickyness property for which \( X_0 = x \) and \( Y_0 = y \). Let \( \tau_{\text{couple}} \) be the first time the chains meet:

\[
\tau_{\text{couple}} = \min\{ t \mid X_t = Y_t \}
\]  \hfill (7)

Then:

\[
\left\| P^t(x, \cdot) - P^t(y, \cdot) \right\|_{TV} \leq P_{x,y}\{\tau_{\text{couple}} > t}\}
\]  \hfill (8)
Corollary

Suppose that for each pair of states $x, y \in \Omega$ there is a coupling $(X_t, Y_t)$ with $X_0 = x$ and $Y_0 = y$. For each such coupling, let $\tau_{\text{couple}}$ be the first time the chains meet. Then:

$$d(t) \leq \max_{x, y \in \Omega} \left( P_{x, y} \{ \tau_{\text{couple}} > t \} \right).$$  \hspace{1cm} (9)
Winning Streak - Revisited

time $t$:  1 0 1 0 0 1 1 1 1 0 0 0 0 0

time $t+1$:  1 0 1 0 0 1 1 1 1 0 0 0 0 0

time $t+2$:  1 0 1 0 0 1 1 1 1 0 0 0 0 0
Winning Streak - Revisited

\[ P(i, 0) = \frac{1}{2} \quad \text{for } 0 \leq i \leq n \]
\[ P(i, i + 1) = \frac{1}{2} \quad \text{for } 0 \leq i < n \]
\[ P(n, n) = \frac{1}{2} \] (10)
Winning Streak - Revisited

And for the reverse chain:

\[
\begin{align*}
\hat{P}(0, i) &= \pi(i) & \text{for } 0 \leq i \leq n \\
\hat{P}(i, i - 1) &= 1 & \text{for } 0 \leq i < n \\
\hat{P}(n, n) &= \frac{1}{2}
\end{align*}
\]
We also saw that the stationary distribution $\pi$ for both chains is:

$$
\pi(i) = \begin{cases} 
\frac{1}{2^{i+1}} & \text{if } i = \{0, \ldots, n-1\} \\
\frac{1}{2^n} & \text{if } i = n
\end{cases}
$$

(12)

What about mixing time?
Winning Streak - Revisited

We also saw that the stationary distribution $\pi$ for both chains is:

$$\pi(i) = \begin{cases} 
\frac{1}{2^{i+1}} & \text{if } i = \{0, ..., n - 1\} \\
\frac{1}{2^n} & \text{if } i = n
\end{cases}$$

(12)

What about mixing time?

The reverse chain has $t_{mix}(\epsilon) = n$ for every $\epsilon \leq \frac{1}{4}$. 
Winning Streak - Revisited

What about the original one?
What about the original one?
For any initial state \( a, b \in \{0, ..., n\} \), define \( x \) and \( y \) to be the current bitstrings with exactly \( a \) and \( b \) ending 1’s respectively.

Example:

\[
x = 0 \ 0 \ 1 \ 1 \ 1, \ a=3
y = 1 \ 0 \ 0 \ 1 \ 1, \ b=2
\]
Winning Streak - Revisited

What about the original one?
For any initial state \( a, b \in \{0, \ldots, n\} \), define \( x \) and \( y \) to be the current bitstrings with exactly \( a \) and \( b \) ending 1’s respectively.

Example:

\[
\begin{align*}
x &= 0 \ 0 \ 1 \ 1 \ 1, \ a=3 \\
y &= 1 \ 0 \ 0 \ 1 \ 1, \ b=2
\end{align*}
\]

Note:
for \( y=0 \ 0 \ 0 \ 1 \ 1 \), \( b \) is also 2
What about the original one?
For any initial state \( a, b \in \{0, ..., n\} \), define \( x \) and \( y \) to be the current bitstrings with exactly \( a \) and \( b \) ending 1’s respectively.

Example:

\[
x = 0 \ 0 \ 1 \ 1 \ 1, \ a=3 \\
y = 1 \ 0 \ 0 \ 1 \ 1, \ b=2
\]

Note:
for \( y=0 \ 0 \ 0 \ 1 \ 1 \), \( b \) is also 2
Any suggestion for coupling?
Winning Streak - Revisited

We will use the following rule to create coupling:
Generate a random bit and put it at the end of \( x \) and \( y \), then we count the number of ending 1’s.
Winning Streak - Revisited

$x$

$a = 3 \quad 00111$

$y$

$b = 2 \quad 10011$

Flip Coin Generator
Winning Streak - Revisited

Easy to check that it is indeed coupling and has the stickyness property.
Winning Streak - Revisited

Easy to check that it is indeed coupling and has the stickyness property. We saw that as soon as we generate random bit that is 0, than both chain fall into state 0 (therefore coupled).
Winning Streak - Revisited

Easy to check that it is indeed coupling and has the stickyness property.
We saw that as soon as we generate random bit that is 0, than both chain fall into state 0 (therefore coupled).
So for every a and b:

\[ P_{a,b}(\tau_{couple} > t) \leq 2^{-t} \]  \hspace{1cm} (13)

Therefore:

\[ t_{mix}(\epsilon) \leq \left\lceil \log_2 \left( \frac{1}{\epsilon} \right) \right\rceil \]  \hspace{1cm} (14)
Winning Streak - Revisited

Notice that the bound is not depend on n, only on $\epsilon$ !
And:

$$ t_{mix} \leq 2 $$  (15)
Wait, we really get the stationary distribution after 2 steps?
Winning Streak - Revisited

Wait, we really get the stationary distribution after 2 steps? No, consider that the chain started at state $i < (n - 2)$. What happen after 2 steps in the chain?
Winning Streak - Revisited

Wait, we really get the stationary distribution after 2 steps? No, consider that the chain started at state \( i < (n - 2) \). What happen after 2 steps in the chain?

\[
P^2(i, \cdot) = \left( \frac{1}{2}, \frac{1}{4}, 0, \ldots, \frac{1}{4}, 0, \ldots, 0 \right)
\]
Random Colorings - Reminder

\( G = (V, E) \)
Random Colorings - Reminder

- $G = (V, E)$
- $S = \{1, \ldots, q\}$
Random Colorings - Reminder

- $G = (V, E)$
- $S = \{1, \ldots, q\}$
- $\Omega \subseteq S^V$ - proper coloring of $G$
Random Colorings - Reminder

- $G = (V, E)$
- $S = \{1, \ldots, q\}$
- $\Omega \subseteq S^V$ - proper coloring of $G$
- $\pi(i) = \frac{1}{|\Omega|}$
Random Colorings - Metropolis Chain

\[ G = (V, E) \]
Theorem

Let $G$ be a graph with $n$ vertices and maximal degree $\Delta$. For the Metropolis chain on proper colorings of $G$, if $q > 3\Delta$ and $c_{\text{met}}(\Delta, q) = 1 - \frac{3\Delta}{q}$, then:

$$t_{\text{mix}}(\epsilon) \leq c_{\text{met}}(\Delta, q)^{-1} n \left( \log(n) + \log\left(\frac{1}{\epsilon}\right) \right)$$
Random Colorings Mixing Time

Lets have a closer look at the condition $q > 3\Delta$. 
If \( q \geq \Delta + 1 \), what we can say about existence of proper coloring?
Random Colorings Mixing Time

$G = (V, E)$
So it seems to be easy to find a proper coloring when $q = \Delta + 1$. But what about irreducibility when $q = \Delta + 1$ (respect to the Metropolis chain)?
Random Colorings Mixing Time

So it seems to be easy to find a proper coloring when $q = \Delta + 1$. But what about irreducibility when $q = \Delta + 1$ (respect to the Metropolis chain)?

G = (V, E)

$S$
One can check that the Metropolis chain is irreducible for $q \geq \Delta + 2$. 
Random Coloring Coupling

Proof:
We will use the following coupling step \((X^x_t, X^y_t) \rightarrow (X^x_{t+1}, X^y_{t+1})\):

1. Generate pair \((v, k)\) u.a.r from \(V \times S\)
Random Coloring Coupling

Proof:
We will use the following coupling step \((X^x_t, X^y_t) \rightarrow (X^x_{t+1}, X^y_{t+1})\):

1. Generate pair \((v, k)\) u.a.r from \(V \times S\)
2. Attempt to color \(v\) with \(k\)

It's easy to check that it's indeed a coupling and has the stickyness property.
Proof:
We will use the following coupling step \((X^x_t, X^y_t) \rightarrow (X^x_{t+1}, X^y_{t+1})\):

1. Generate pair \((v, k)\) u.a.r from \(V \times S\)
2. Attempt to color \(v\) with \(k\)
3. If no neighbor of \(v\) has color \(k\), update
4. otherwise, no changes are made
Random Coloring Coupling

Proof:
We will use the following coupling step \((X_t^x, X_t^y) \rightarrow (X_{t+1}^x, X_{t+1}^y)\):

1. Generate pair \((v, k)\) u.a.r from \(V \times S\)
2. Attempt to color \(v\) with \(k\)
3. If no neighbor of \(v\) has color \(k\), update
4. otherwise, no changes are made

It's easy to check that it's indeed coupling and have the stickyness property.
Random Colorings Coupling

Definition (Hamming Distance)

\[ \rho(x, y) = \sum_{v \in V} 1_{\{x(v) \neq y(v)\}} \]

Proposition (\( \rho \) is metric)

for \( \forall x, y, w \in \Omega \):

\[ \rho(x, y) \leq \rho(x, w) + \rho(y, w) \]
Proof.

If \( x(v) = y(v) \) then:

\[
0 = 1\{x(v)\neq y(v)\} \leq 1\{x(v)\neq w(v)\} + 1\{w(v)\neq y(v)\}
\]
Random Colorings Coupling

Proof.

If $x(v) = y(v)$ then:

$$0 = 1_{\{x(v) \neq y(v)\}} \leq 1_{\{x(v) \neq w(v)\}} + 1_{\{w(v) \neq y(v)\}}$$

If $x(v) \neq y(v)$ then:

$$1 = 1_{\{x(v) \neq y(v)\}} \leq 1_{\{x(v) \neq w(v)\}} + 1_{\{w(v) \neq y(v)\}} =$$

$$\begin{cases} 
2 & w(v) \neq y(v) \text{ and } w(v) \neq x(v) \\
1 & w(v) = x(v) \\
1 & w(v) = y(v) 
\end{cases} \quad (16)$$

Sum over all $v \in V$ completes the proof.
Random Colorings Coupling

Proof Cont.
Assume $\rho(x, y) = 1$, so $x$ and $y$ agree everywhere except at $v_0$.
Let $N$ be the set of colors on the neighbors of $v_0$ in the coloring $x$ ($y$).
First let's see what can happen after one update, meaning that we have interest with \( \rho(X^x_1, X^y_1) \):
First let's see what can happen after one update, meaning that we have interest with \( \rho(X_1^x, X_1^y) \):

Clearly \( \rho(X_1^x, X_1^y) \in \{0, 1, 2\} \) (Because we are using the same update for both chains).
When $\rho(X^x_1, X^y_1) = 0$?
When $\rho(X_i^x, X_i^y) = 0$?
It happens iff $\nu_0$ is selected and $k \notin N$:

$$\Pr\{\rho(X_1^x, X_1^y) = 0\} = \frac{1}{n} \cdot \frac{q - N}{q} \geq \frac{q - \Delta}{nq}$$
Random Colorings Coupling

When $\rho(X_1^x, X_1^y) = 2$?
Random Colorings Coupling

When $\rho(X_t^x, X_t^y) = 2$?
It happens iff a vertex $w$ which is neighbor of $v_0$ is selected and $k = x(v_0)$ or $k = y(v_0)$:

$$\Pr\{\rho(X_1^x, X_1^y) = 2\} \leq \frac{\Delta}{n} \cdot \frac{2}{q} \leq \frac{2\Delta}{nq}$$
Now we have a look at $E(\rho(X_1^x, X_1^y))$:

$$E(\rho(X_1^x, X_1^y) - 1) \leq \frac{2\Delta}{nq} - \frac{q - \Delta}{nq} = \frac{3\Delta - q}{nq}$$

$$\Rightarrow E(\rho(X_1^x, X_1^y) \leq 1 - \frac{q - 3\Delta}{nq} \underbrace{\text{c}_{\text{met}(\Delta, q)}}_{\text{c}_{\text{met}(\Delta, q)}}$$
Random Colorings Coupling

Now we have a look at $E(\rho(X_1^x, X_1^y))$:

$$E(\rho(X_1^x, X_1^y) - 1) \leq \frac{2\Delta}{nq} - \frac{q - \Delta}{nq} = \frac{3\Delta - q}{nq}$$

$$\Rightarrow E(\rho(X_1^x, X_1^y)) \leq 1 - \frac{q - 3\Delta}{nq}$$

If $c_{met}(\Delta, q) > 0$ we have:

$$E(\rho(X_1^x, X_1^y)) < 1$$
Random Colorings Coupling

What about $E(\rho(X_t^x, X_t^y))$?
What about $E(\rho(X^x_t, X^y_t))$? From the Markov behavior of the chain we can see that the event 
\[
\{ \rho(X^x_t, X^y_t) \mid X^x_{t-1} = x_{t-1}, X^y_{t-1} = y_{t-1} \}
\]
has the same distribution as the event $\rho(X^x_{t-1}, X^y_{t-1})$. We know how to bound this expectation when $\rho(x_{t-1}, y_{t-1}) = 1$. But what about the general case?
Random Colorings Coupling

What about $E(\rho(X_t^x, X_t^y))$?

From the Markov behavior of the chain we can see that the event
\[
\{\rho(X_t^x, X_t^y) \mid X_{t-1}^x = x_{t-1}, X_{t-1}^y = y_{t-1}\}
\]
has the same distribution as the event $\rho(X_{t-1}^{x_{t-1}}, X_{t-1}^{y_{t-1}})$.

So the events have the same expectation:

\[
E(\rho(X_t^x, X_t^y) \mid X_{t-1}^x = x_{t-1}, X_{t-1}^y = y_{t-1}) = E(\rho(X_{t-1}^{x_{t-1}}, X_{t-1}^{y_{t-1}}))
\]

We know how to bound this expectation when $\rho(x_{t-1}, y_{t-1}) = 1$.

But what about the general case?
Random Colorings Coupling

The case when $\rho(x, y) = r$:
Random Colorings Coupling

The case when $\rho(x, y) = r$:

Intuitively, we want to find a sequence of colorings $x = x_0, x_1, \ldots, x_r = y$ such that $\rho(x_i, x_{i+1}) = 1$ for $0 \leq i \leq r - 1$.
Random Colorings Coupling

The case when $\rho(x, y) = r$:

Intuitively we want to find sequence of colorings $x = x_0, x_1, \ldots, x_r = y$ such that $\rho(x_i, x_{i+1}) = 1$ for $0 \leq i \leq r - 1$.

And since we saw that $\rho$ is metric:

$$
\rho(X^x_1, X^y_1) \leq \sum_{k=1}^{r} \rho(X^x_k, X^{x_k-1}_1)
$$
Random Colorings Coupling

The case when $\rho(x, y) = r$:
Intuitively We want to find sequence of colorings $x = x_0, x_1, \ldots, x_r = y$ such that $\rho(x_i, x_{i+1}) = 1$ for $0 \leq i \leq r - 1$
And since we saw that $\rho$ is metric:

$$\rho(X_1^x, X_1^y) \leq \sum_{k=1}^{r} \rho(X_1^{x_k}, X_1^{x_{k-1}})$$

Then we will want to calculate expectation on both side of the inequality.
Random Colorings Coupling

The case when $\rho(x, y) = r$:
Intuitively, we want to find a sequence of colorings $x = x_0, x_1, \ldots, x_r = y$ such that $\rho(x_i, x_{i+1}) = 1$ for $0 \leq i \leq r - 1$.
And since we saw that $\rho$ is metric:

$$
\rho(X_1^x, X_1^y) \leq \sum_{k=1}^{r} \rho(X_1^{x_k}, X_1^{x_{k-1}})
$$

Then we will want to calculate expectation on both sides of the inequality.
But wait, our coupling is defined for two copies only...
So in order to calculate the expectation on the right side, we need to "expand" our coupling to contain all possible copies:

$$
\{ X_t^x \mid x \in \Omega \}
$$

This is often called as "Grand coupling".
But wait again...
Our "Grand Coupling" is defined only over proper colorings ($x \in \Omega$). Why it can be a problem?
Random Colorings Coupling

Even when we have irreducibility \((q \geq 3\Delta)\).
Random Colorings Coupling

So what if we extend our "Grand Coupling" for all colorings (not just proper ones)?
Random Colorings Coupling

So what if we extend our "Grand Coupling" for all colorings (not just proper ones)?
We have to handle some technical aspects...
Random Colorings Coupling

First, we have to extend our Metropolis chain to handle improper colorings to. The original transition rule can be easily extend to all colorings. But does it break our coupling?
Random Colorings Coupling

First, we have to extend our Metropolis chain to handle improper colorings to. The original transition rule can be easily extend to all colorings. But does it break our coupling? Remember that in order to use the coupling bound we need a faithful copy of the original chain, but only for starting states that are proper coloring!
Random Colorings Coupling

How $X_t^x$ look when $x$ is proper coloring?
Random Colorings Coupling

How $X_t^x$ look when $x$ is proper coloring?
Just like the original Metropolis chain!
Random Colorings Coupling

How $X_t^x$ look when $x$ is proper coloring?
Just like the original Metropolis chain!
Thats because we can’t get to improper coloring from proper one.
Random Colorings Coupling

How $X_t^x$ look when $x$ is proper coloring?
Just like the original Metropolis chain!
Thats because we can’t get to improper coloring from proper one.
So the extended ”Grand Coupling” will be:

$$\{X_t^x \mid x \in \tilde{\Omega}\}$$

When $\tilde{\Omega} = S^V$. 

Random Colorings Coupling

So now:

\[ E(\rho(X_1^x, X_1^y)) \leq \sum_{k=1}^{r} E(\rho(X_1^{x_k}, X_1^{x_{k-1}})) \]
Random Colorings Coupling

So now:

\[ E(\rho(X_1^x, X_1^y)) \leq \sum_{k=1}^{r} E(\rho(X_1^{x_k}, X_1^{x_{k-1}})) \]

\[ \leq r \left( 1 - \frac{c_{met}(\Delta, q)}{n} \right) \]

\[ = \rho(x, y) \left( 1 - \frac{c_{met}(\Delta, q)}{n} \right) \]
Random Colorings Coupling

Back to:

$$E(\rho(X_t^x, X_t^y) \mid X_{t-1}^x = x_{t-1}, X_{t-1}^y = y_{t-1}) = E(\rho(X_{1_{t-1}}^x, X_{t-1}^{y_{t-1}}))$$
Random Colorings Coupling

Back to:

\[ E(\rho(X_t^x, X_t^y) \mid X_{t-1}^x = x_{t-1}, X_{t-1}^y = y_{t-1}) = E(\rho(X_{t-1}^{x_{t-1}}, X_{t-1}^{y_{t-1}})) \]

\[ \leq \rho(x_{t-1}, y_{t-1})\left(1 - \frac{c_{\text{met}}(\Delta, q)}{n}\right) \]
Random Colorings Coupling

Back to:

\[
E(\rho(X^x_t, X^y_t) \mid X^x_{t-1} = x_{t-1}, X^y_{t-1} = y_{t-1}) = E(\rho(X^{x_{t-1}}_1, X^{y_{t-1}}_1))
\]

\[
\leq \rho(x_{t-1}, y_{t-1})(1 - \frac{c_{met}(\Delta, q)}{n})
\]

Using the law of total expectation:

\[
E(\rho(X^x_t, X^y_t)) \leq E(\rho(X^x_{t-1}, X^y_{t-1}))(1 - \frac{c_{met}(\Delta, q)}{n})
\]
Random Colorings Coupling

Using iteration we get:

\[ E(\rho(X_t^x, X_t^y)) \leq \rho(x, y)(1 - \frac{c_{met}(\Delta, q)}{n})^t \]
Random Colorings Coupling

Using iteration we get:

\[ E(\rho(X_t^x, X_t^y)) \leq \rho(x, y) \left(1 - \frac{c_{met}(\Delta, q)}{n}\right)^t \]

Notice that since \( c_{met}(\Delta, q) < 1 \) when \( q > 3\Delta \) we found that the expectation of the distance between the coloring is decreasing exponentially in \( c_{met}(\Delta, q)/n \).
Random Colorings Coupling

Using iteration we get:

$$E(\rho(X_t^x, X_t^y)) \leq \rho(x, y) \left(1 - \frac{c_{met}(\Delta, q)}{n}\right)^t$$

Notice that since $c_{met}(\Delta, q) < 1$ when $q > 3\Delta$ we found that the expectation of the distance between the coloring is decreasing exponentially in $c_{met}(\Delta, q)/n$.

Now bound the couple time:

$$\Pr(X_t^x \neq X_t^y) = \Pr(\rho(X_t^x, X_t^y) \geq 1)$$
Random Colorings Coupling

Using iteration we get:

\[ E(\rho(X_t^x, X_t^y)) \leq \rho(x, y) \left(1 - \frac{c_{\text{met}}(\Delta, q)}{n}\right)^t \]

Notice that since \( c_{\text{met}}(\Delta, q) < 1 \) when \( q > 3\Delta \) we found that the expectation of the distance between the coloring is decreasing exponentially in \( c_{\text{met}}(\Delta, q)/n \).

Now bound the couple time:

\[ \Pr(X_t^x \neq X_t^y) = \Pr(\rho(X_t^x, X_t^y) \geq 1) \]

\[ \leq E(\rho(X_t^x, X_t^y)) \]

\[ \text{Markov} \]
Random Colorings Coupling

Using iteration we get:

\[ E(\rho(X^x_t, X^y_t)) \leq \rho(x, y) (1 - \frac{c_{met}(\Delta, q)}{n})^t \]

Notice that since \( c_{met}(\Delta, q) < 1 \) when \( q > 3\Delta \) we found that the expectation of the distance between the coloring is decreasing exponentially in \( c_{met}(\Delta, q)/n \).

Now bound the couple time:

\[ \Pr(X^x_t \neq X^y_t) = \Pr(\rho(X^x_t, X^y_t) \geq 1) \]

\[ \leq E(\rho(X^x_t, X^y_t)) \]

Markov

\[ \leq \rho(x, y) (1 - \frac{c_{met}(\Delta, q)}{n})^t \]
Random Colorings Coupling

Using iteration we get:

\[ E(\rho(X^x_t, X^y_t)) \leq \rho(x, y) \left(1 - \frac{c_{met}(\Delta, q)}{n}\right)^t \]

Notice that since \( c_{met}(\Delta, q) < 1 \) when \( q > 3\Delta \) we found that the expectation of the distance between the coloring is decreasing exponentially in \( c_{met}(\Delta, q)/n \).

Now bound the couple time:

\[ \Pr(X^x_t \neq X^y_t) = \Pr(\rho(X^x_t, X^y_t) \geq 1) \]

\[ \leq E(\rho(X^x_t, X^y_t)) \]

\[ \leq \rho(x, y) \left(1 - \frac{c_{met}(\Delta, q)}{n}\right)^t \]

\[ \leq ne^{-t(c_{met}(\Delta, q)/n)} \]

\[ 1 - x \leq e^{-x} \]
Random Colorings Coupling

Request $t$ such that $d(t) \leq \epsilon$ we get:

$$t_{mix}(\epsilon) \leq c_{met}(\Delta, q)^{-1} n(\log(n) + \log\left(\frac{1}{\epsilon}\right))$$
Random Colorings Coupling

Some observations on the proof:

- Grand coupling is helpful when we need "interpolation capability"
Random Colorings Coupling

Some observations on the proof:

- Grand coupling is helpful when we need "interpolation capability"
- In the most of the examples we saw so far, we tried to find a distance function that when it hits zero we have coupled chains. We tried to figure out a coupling that make this distance function to decrease fast in expectation manner (preferably exponentially). Then we used the Markov inequality to bound the coupling time.
Random Colorings Coupling

Some observations on the proof:

- Grand coupling is helpful when we need ”interpolation capability”
- In the most of the examples we saw so far, we tried to find a distance function that when it hits zero we have coupled chains. We tried to figure out a coupling that make this distance function to decrease fast in expectation manner (preferably exponentially). Then we used the Markov inequality to bound the coupling time.
- Easy to say, hard to do. Coupling in this form is limited (sometimes called one step coupling). Better coupling methods are out there (Path Coupling).
Theorem

Let $G$ be a graph with maximum degree $\Delta$ and $n$ vertices. Assume $q \geq 2\Delta + 1$ then:
Random Colorings - Revisited

Theorem

Let $G$ be a graph with maximum degree $\Delta$ and $n$ vertices. Assume $q \geq 2\Delta + 1$ then:

$$t_{mix}(\epsilon) \leq \frac{q}{q - 2\Delta} n \log \left( \frac{n}{\epsilon} \right)$$

Note: We still get mixing time that is $O(n \log n)$, but the assumption on $q$ is weaker!
Theorem

Let $G$ be a graph with maximum degree $\Delta$ and $n$ vertices. Assume $q \geq 2\Delta + 1$ then:

$$t_{\text{mix}}(\epsilon) \leq \frac{q}{q - 2\Delta} n \log\left(\frac{n}{\epsilon}\right)$$

Note:
We still get mixing time that is $O(n \log n)$, but the assumption on $q$ is weaker!
Random Colorings - Revisited

We will use the same Metropolis chain, just on proper coloring this time. We will build coupling ("regular" not the "grand" one) as follow:

1. choose $v \in V$ u.a.r
2. compute $g = g(G, X_t, Y_t)$ permutation of $S$ (explained later how)
3. choose color $k \in S$ u.a.r
4. In the coloring $X_t$ (respectively $Y_t$) recolor $v$ with $k$ (respectively $g(k)$) to get new colorings $X'_t$ (respectively $Y'_t$).
5. If $X'_t$ (respectively $Y'_t$) is proper coloring then $X_{t+1} = X'_t$ (respectively $Y_{t+1} = Y'_t$), otherwise no changes are made.
We will use the same Metropolis chain, just on proper coloring this time. We will build coupling ("regular" not the "grand" one) as follow:

1. choose $v \in V$ u.a.r
Random Colorings - Revisited

We will use the same Metropolis chain, just on proper coloring this time. We will build coupling ("regular" not the "grand" one) as follow:

1. choose \( v \in V \) u.a.r
2. compute \( g = g(G, X_t, Y_t) \) permutation of \( S \) (explained later how)
We will use the same Metropolis chain, just on proper coloring this time. We will build coupling ("regular" not the "grand" one) as follow:

1. choose $v \in V$ u.a.r

2. compute $g = g(G, X_t, Y_t)$ permutation of $S$ (explained later how)

3. choose color $k \in S$ u.a.r

4. In the coloring $X_t$ (respectively $Y_t$) recolor $v$ with $k$ (respectively $g(k)$) to get new colorings $X'$ (respectively $Y'$).
Random Colorings - Revisited

We will use the same Metropolis chain, just on proper coloring this time. We will build coupling ("regular" not the "grand" one) as follow:

1. choose $v \in V$ u.a.r
2. compute $g = g(G, X_t, Y_t)$ permutation of $S$ (explained later how)
3. choose color $k \in S$ u.a.r
4. In the coloring $X_t$ (respectively $Y_t$) recolor $v$ with $k$ (respectively $g(k)$) to get new colorings $X'$ (respectively $Y'$).
5. If $X'$ (respectively $Y'$) is proper coloring then $X_{t+1} = X'$ (respectively $Y_{t+1} = Y'$), otherwise no changes are made.
Proposition

No matter how we choose $g$, each of the sequences $(X_t), (Y_t)$ is a copy of the Metropolis chain.
Random Colorings - Revisited

Proposition

No matter how we choose $g$, each of the sequences $(X_t), (Y_t)$ is a copy of the Metropolis chain.

Proof.

For $(X_t)$ its easy to verify, since we choose $v$ and $k$ u.a.r from $V \times S$ like the Metropolis chain.
Proposition

No matter how we choose $g$, each of the sequences $(X_t), (Y_t)$ is a copy of the Metropolis chain.

Proof.
For $(X_t)$ its easy to verify, since we choose $v$ and $k$ u.a.r from $V \times S$ like the Metropolis chain.
For $(Y_t)$, because $k$ is chose u.a.r from $S$ so is $g(k)$. Therefore from $(Y_t)$ perspective its the same chain as the Metropolis chain.
Proposition

No matter how we choose $g$, each of the sequences $(X_t), (Y_t)$ is a copy of the Metropolis chain.

Proof.
For $(X_t)$ its easy to verify, since we choose $v$ and $k$ u.a.r from $V \times S$ like the Metropolis chain.
For $(Y_t)$, because $k$ is chose u.a.r from $S$ so is $g(k)$. Therefore from $(Y_t)$ perspective its the same chain as the Metropolis chain.

Corollary

The coupling proposed above is indeed a Coupling.
Random Colorings - Revisited

Let $A = A_t \subseteq V$ the set of vertices which $X_t$ and $Y_t$ agree

$A = A_t = \{v_2, v_3\}$
Random Colorings - Revisited

- Let $D = D_t \subseteq V$ the set of vertices which $X_t$ and $Y_t$ disagree.

\[ D = D_t = \{v_0, v_1\} \]
Random Colorings - Revisited

- Let $d'(v)$ be the number of edges incident in $v$ which has one vertex in $A$ and one vertex in $D$.

\[ d'(v_0) = 2 \]
\[ d'(v_2) = 2 \]
Proposition

\[
\sum_{v \in A} d'(v) = \sum_{v \in D} d'(v) = m'
\]

When \(m'\) is the number of edges that span \(A\) and \(D\).
Random Colorings - Revisited

Proof.

\[ m' = 4 \]

\( X_t \) \hspace{2cm} \( Y_t \)
"Good" g will try to "block" bad updates (which increase $D$ size).
So, how will g be defined?
\( \nu_0 \in D \)
$v_0 \in A$

$X_t$

$Y_t$
Random Colorings - Revisited

Let $N$ be the set of $v_0$ neighbors.
Let $C_X \subseteq S$ to be the set of colors of $N$ in the coloring $X_t$, which no $w \in N$ has in the coloring $Y_t$.

$v_0 \in A$

$x_t$

$y_t$

$C_X = \{white\}$
Random Colorings - Revisited

Let $C_Y \subseteq S$ be defined similar to above ($X_t$ and $Y_t$ interchange).

$v_0 \in A$

$X_t$

$Y_t$

$C_X = \{white\}$

$C_Y = \{black, brown\}$
Random Colorings - Revisited

$v_0 \in A$

$X_t$

$Y_t$

$C_X = \{\text{white}\}$

$C_Y = \{\text{black, brown}\}$

$g(\text{white}) = \text{black}$
Random Colorings - Revisited

In the previous coupling we bound the probability of choosing ”bad” color (one that make $D$ increase) by $\frac{|C_X| + |C_Y|}{q}$. 
Random Colorings - Revisited

In the previous coupling we bound the probability of choosing "bad" color (one that make $D$ increase) by $\frac{|C_X| + |C_Y|}{q}$.

Without loss of generality if $|C_X| \leq |C_Y|$, we can bound this probability by $\frac{|C_Y|}{q}$, by doing the following "trick":

Choose arbitrary subset of $C'_Y \subseteq C_Y$ such that $|C_X| = |C'_Y|$.

Assume $C_X = \{c_1, \ldots, c_r\}$ and $C'_Y = \{c'_1, \ldots, c'_r\}$. Define $g$ to be:

$$g(k) = \begin{cases} 
k, & k \in C_X \land k = c_i \\
\frac{k}{q}, & k = c'_i
\end{cases}$$
In the previous coupling we bound the probability of choosing "bad" color (one that make $D$ increase) by $\frac{|C_X| + |C_Y|}{q}$.

Without loss of generality if $|C_X| \leq |C_Y|$, we can bound this probability by $\frac{|C_Y|}{q}$, by doing the following "trick":

Choose arbitrary subset of $C'_Y \subseteq C_Y$ such that $|C_X| = |C'_Y|$. 
In the previous coupling we bound the probability of choosing "bad" color (one that make D increase) by $\frac{|C_X| + |C_Y|}{q}$.

Without loss of generality if $|C_X| \leq |C_Y|$, we can bound this probability by $\frac{|C_Y|}{q}$, by doing the following "trick":

Choose arbitrary subset of $C_Y' \subseteq C_Y$ such that $|C_X| = |C_Y'|$.

Assume $C_X = \{c_1, \ldots, c_r\}$ and $C_Y' = \{c'_1, \ldots, c'_r\}$.

Define $g$ to be:

$$g(k) = \begin{cases} 
  k, & k \notin C_X \\
  k', & k = c_i \text{ and } k' = c'_i
\end{cases}$$
Proposition

\[ C_X \cap C_Y = \emptyset \]
Proposition
\[ C_X \cap C_Y = \emptyset \]

Proof.
Easy, by definition of \( C_X \) and \( C_Y \). \qed
Random Colorings - Revisited

Proposition

For $v \in A$: $|C_X|, |C_Y| \leq d'(v)$
Random Colorings - Revisited

Proposition

For $v \in A$: $|C_X|, |C_Y| \leq d'(v)$

Proof.

$v \in A$, therefore $d'(v)$ is the number of neighbors that the colorings $X_t$ and $Y_t$ are disagree. On the other hand, $C_X$ and $C_Y$ are colors that not appear in neighbors of $v$ in $Y_t$ and $X_t$ respectively. Therefore, the inequality holds.

$v \in A$

$X_t$

$Y_t$

$v_0$

$C_X = \{white\}$

$C_Y = \{black, brown\}$

$d'(v) = 5$
Random Colorings - Revisited

**Proposition**
\[ |D_{t+1}| - |D_t| \in \{-1, 0, 1\} \]

**Proof.**
Easy, from the definition of the coupling step, we update at most one vertex and the same one for both \(X_t\) and \(Y_t\). \qed
Random Colorings - Revisited

Proposition

We have stickyness

Proof.

When \((X_t)\) and \((Y_t)\) are coupled, the coupling step will choose only in \(A\), but \(C_X = C_Y = \emptyset\). Therefore \(g\) is the identity permutation. Now, the same vertex and same color are chose for both chains, and they will be update (or not) together.

\[ g \text{ is the identity} \]
Random Colorings - Revisited

When $|D_{t+1}| = |D_t| + 1$?

- $v \in A$, $g$ is calculated as in the case 2
Random Colorings - Revisited

When $|D_{t+1}| = |D_t| + 1$?

- $v \in A$, $g$ is calculated as in the case 2
- $k \in C_Y$
Random Colorings - Revisited

When $|D_{t+1}| = |D_t| + 1$?

- $v \in A$, $g$ is calculated as in the case 2
- $k \in C_Y$

$v_0 \in A$  $X_t$  $Y_t$

$C_X = \{\text{white}\}$  $C_Y = \{\text{black, brown}\}$

$g(\text{white}) = \text{black}$
\[ \Pr(|D_{t+1}| = |D_t| + 1) \leq \frac{1}{n} \sum_{v \in A} \frac{d'(v)}{q} = \frac{m'}{nq} \]
Random Colorings - Revisited

When $|D_{t+1}| = |D_t| - 1$?

- $v \in D$, therefore $g$ is the identity
When $|D_{t+1}| = |D_t| - 1$?

- $v \in D$, therefore $g$ is the identity
- $k$ has to be color that not in neighbors of $v$ in either $X_t$ and $Y_t$. How many colors we have to choose from?
When $|D_{t+1}| = |D_t| - 1$?

- $v \in D$, therefore $g$ is the identity
- $k$ has to be color that not in neighbors of $v$ in either $X_t$ and $Y_t$. How many colors we have to choose from?
When \( |D_{t+1}| = |D_t| - 1? \)

- \( v \in D \), therefore \( g \) is the identity
- \( k \) has to be color that not in neighbors of \( v \) in either \( X_t \) and \( Y_t \). How many colors we have to choose from?

Naive bound will be at least \( q - 2\Delta \)
When \( |D_{t+1}| = |D_t| - 1? \)

- \( v \in D \), therefore \( g \) is the identity
- \( k \) has to be color that not in neighbors of \( v \) in either \( X_t \) and \( Y_t \). How many colors we have to choose from?

Naive bound will be at least \( q - 2\Delta \)
But Notice that we count $d'(\nu)$ twice in the subtraction...
Random Colorings - Revisited

At least \( q - 2\Delta + d'(\nu) \).

- \( q \) - is total number of colors
- \( 2\Delta \) - is the maximum number of neighbors
- \( d'(\nu) \) - is the number of neighbors that agree on color in both \( X_t \) and \( Y_t \) (because \( \nu \in D \))
At least \( q - 2\Delta + d'(v) \).

- \( q \) - is total number of colors
- \( 2\Delta \) - is the maximum number of neighbors
- \( d'(v) \) - is the number of neighbors that agree on color in both \( X_t \) and \( Y_t \) (because \( v \in D \))

\[
\Rightarrow \Pr(|D_{t+1}| = |D_t| - 1) \geq \frac{1}{n} \sum_{v \in D} \left( \frac{q - 2\Delta + d'(v)}{q} \right)
\]

\[
= \frac{q - 2\Delta}{nq} |D_t| + \frac{m'}{nq}
\]
Random Colorings - Revisited

- Let \( \alpha = \frac{q - 2\Delta}{nq} > 0 \)
Random Colorings - Revisited

- Let $\alpha = \frac{q - 2\Delta}{nq} > 0$
- Let $\beta = \frac{m'}{nq}$

Note the above calculations hold when $D_t \neq \emptyset$ and $D_t \neq V$. But it's easy to check that the inequality holds for those cases too.
Random Colorings - Revisited

- Let $\alpha = \frac{q-2\Delta}{nq} > 0$
- Let $\beta = \frac{m'}{nq}$
- So,

\[
\Pr(|D_{t+1}| = |D_t| + 1) \leq \beta \\
\Pr(|D_{t+1}| = |D_t| - 1) \geq \alpha |D_t| + \beta
\]

Note the above calculations holds when $D_t \neq \emptyset$ and $D_t \neq V$. But its easy to check that the inequality holds for those cases too.
Random Colorings - Revisited

- Let $\alpha = \frac{q-2\Delta}{nq} > 0$
- Let $\beta = \frac{m'}{nq}$
- So,

\[
\Pr(|D_{t+1}| = |D_t| + 1) \leq \beta
\]
\[
\Pr(|D_{t+1}| = |D_t| - 1) \geq \alpha|D_t| + \beta
\]

\[
\Rightarrow E(|D_{t+1}| - |D_t|) \leq \beta - (\alpha|D_t| + \beta) = (1 - \alpha)|D_t|
\]
Let $\alpha = \frac{q-2\Delta}{nq} > 0$

Let $\beta = \frac{m'}{nq}$

So,

$$\Pr(|D_{t+1}| = |D_t| + 1) \leq \beta$$
$$\Pr(|D_{t+1}| = |D_t| - 1) \geq \alpha |D_t| + \beta$$

$$\Rightarrow E(|D_{t+1}| - |D_t|) \leq \beta - (\alpha |D_t| + \beta) = (1 - \alpha)|D_t|$$

Note the above calculations holds when $D_t \neq \emptyset$ and $D_t \neq V$. But its easy to check that the inequality holds for those cases too.
Random Colorings - Revisited

By Iteration we get:

\[ E(|D_t|) \leq (1 - \alpha)^t |D_0| \leq n(1 - \alpha)^t \]
By Iteration we get:

\[ E(|D_t|) \leq (1 - \alpha)^t |D_0| \leq n(1 - \alpha)^t \]

\[ \Rightarrow \Pr(|D_t| \geq 1) \leq n(1 - \alpha)^t \leq ne^{-\alpha t} \]
By Iteration we get:

\[ E(|D_t|) \leq (1 - \alpha)^t |D_0| \leq n(1 - \alpha)^t \]

\[ \Rightarrow \Pr(|D_t| \geq 1) \leq n(1 - \alpha)^t \leq ne^{-\alpha t} \]

\[ \Rightarrow \Pr(\tau_{couple} > t) \leq \Pr(|D_t| \geq 1) \leq ne^{-\alpha t} \]
Random Colorings - Revisited

By Iteration we get:

\[ E(\mid D_t \mid) \leq (1 - \alpha)^t \mid D_0 \mid \leq n(1 - \alpha)^t \]

\[ \Rightarrow \Pr(\mid D_t \mid \geq 1) \leq n(1 - \alpha)^t \leq ne^{-\alpha t} \]

\[ \Rightarrow \Pr(\tau_{couple} > t) \leq \Pr(\mid D_t \mid \geq 1) \leq ne^{-\alpha t} \]

Request for \( d(t) \leq \epsilon \) we get:

\[ t_{mix}(\epsilon) \leq \alpha^{-1} \log\left(\frac{n}{\epsilon}\right) \]
Vigoda reduced the requirement for $q$ to be $q \geq \frac{11\Delta}{6}$ but with mixing time $O(n^2)$. (2000)
Dyer and Frieze reduced the requirement for $q$ to $q > \alpha_0 \Delta$ when $\alpha_0 \approx 1.763$ ($\frac{11}{6} \approx 1.833$), but $\Delta = \Omega(\log n)$ and girth $g = \Omega(\log(\Delta))$. (2001)
girth - is the length of shortest cycle.
The main "trick" was to show that after a "burn in" period ($O(n)$) with the above conditions, the colors on neighbors of a vertex are close to uniform distribution. Thus, we can ignore the "bad" states (in terms of coupling time).
Molloy reduced the previous result to $q > \alpha_1 \Delta$ when $\alpha_1 \approx 1.489$ under same assumptions. (2002)
Hayes reduced the girth requirement to $g \geq 6$. (2003)
Last known to writer:
Vigoda and Hayes show $O(n \log n)$ mixing time when
$q > (1 + \epsilon) \Delta \ \forall \epsilon > 0$, girth $g \geq 11$ and $\Delta \geq \Omega(\log n)$. (2005)
Random Colorings - Why it's even interesting?

Antiferromagnetism
Let $G=(V,E)$ be a finite graph. For each spin configuration $\sigma : V \to \{1, \ldots, q\}$, define a Hamiltonian:

$$H(\sigma) = \sum_{\{x,y\} \in E} 1_{\{\sigma(x)=\sigma(y)\}}$$
Let $G=(V,E)$ be a finite graph. For each spin configuration $\sigma : V \rightarrow \{1, \ldots, q\}$, define a Hamiltonian:

$$H(\sigma) = \sum_{\{x,y\} \in E} 1_{\{\sigma(x) = \sigma(y)\}}$$

and for each inverse temperature $\beta \geq 0$, define Gibbs measure:

$$\mu_\beta(\sigma) = \frac{1}{Z_\beta} e^{-\beta H(\sigma)}$$

when $Z_\beta = \sum_\sigma e^{-\beta H(\sigma)}$. 
Random Colorings - Why its even interesting?

In zero temperature $\beta \rightarrow \infty$, we get uniform distribution on proper coloring (if exist):

$$\mu_\beta(\sigma) = \frac{1}{\sum_{\sigma'} e^{-\beta(H(\sigma')-H(\sigma))}}$$

and for proper coloring:

$$H(\sigma) = 0$$
Now that we have an efficient way to sample random coloring, one can show a fpras (fully polynomial randomised approximation scheme) for estimating the number of total proper colorings.