Lecture 19: The Evolving Set Process

Lecturer: Yuval Peres Week: 10

The ***evolving set process (ESP)*** is a Markov chain whose states are subsets of the vertex set V. The goal of this lecture is to show how to use martingales with the ESP in order to obtain bounds for the mixing time of a Markov chain. To begin with, we study the case of a random walk on a graph.

Let $G = (V, E)$ be a graph and consider the lazy simple random walk on G, which is given by the transition kernel

$$p(x, y) = \begin{cases} 1/(2d(x)) & \text{if } \{x, y\} \in E, \\ 1/2 & \text{if } x = y, \end{cases}$$

where $d(x)$ denotes the degree of x. Let $p^t(x, y)$ be the t-step transition kernel. The ***attraction*** of a vertex x to a set S as

$$p(x, S) = \sum_{z \in S} p(x, z) = \frac{1}{2}\frac{e(x, S)}{d(x)} + \frac{1}{2}1(x \in S).$$

We define the ESP as follows: given the current state S_0, pick U uniformly from $[0, 1]$ and let the next state be

$$S_1 = \{y \colon p(y, S_0) \geq U\}.$$

The following result states that the ***volume*** of the sets S_t, that is, the sum of the degree of their vertices, form a martingale.

Proposition 19.126. *Let (S_t) be the previously defined ESP. Then, the sequence $\{vol(S_t)\}$ is a martingale:*

$$\mathbb{E}(vol(S_{t+1})|S_t) = vol(S_t).$$

The usefulness of the ESP to study random walks comes from the following property.

Proposition 19.127. *For all $t \geq 0$ and x, $y \in V$,*

$$\mathbb{P}_x(y \in S_t) = p^t(x, y)\frac{d(x)}{d(y)}.$$

The ESP that we have considered has two absorbing states, the empty set and the set V of all vertices. We will condition on the fact that V will be the absorbing state and study the biased ESP.

The ***volume-biased ESP (vb-ESP)*** is a Markov chain on subsets of V defined as follows. Let (X_t) be a random walk with initial state $X_0 = x_0$. Start with $S_0 = \{x_0\}$. Next, update the random walk taking $X_1 = x_1$ with probability $p(x_0, x_1)$. Then, pick a threshold U uniformly

from $[0, p(x_1, S_0)]$ and take $S_1 = \{y\colon p(y, S) \geq U\}$. This ensures that $x_1 \in S_1$. Then, repeat the procedure to obtain the sequence (S_t).

A nice property that this biased ESP satisfies is the following. For any vertex y and time $t \geq 0$,

$$\mathbb{P}^*_x(X_t = y | S_1, \ldots, S_n) = 1(y \in S_t)\frac{d(t)}{vol(S_t)},$$

where $\mathbb{P}^*_x$ is the probability measure for the coupling $\{(S_t, X_t)\}$. In other words, given $S_1, \ldots, S_t$, the distribution of the random walk X_t is just the stationary distribution restricted to S_t. The transition kernel for the volume-biased ESP is

$$\widehat{K}(S, S`) = \frac{vol(S')}{vol(S)}K(S, S'),$$

where $K(S, S')$ is the transition kernel for the ESP. Thus, the vb-ESP is equivalent to the ESP conditioned to absorb in V, that is, it is the Doob h-transform of the ESP.

Proposition 19.128. *For any function f and any starting set $S_0 \neq \emptyset$,*

$$\widehat{\mathbb{E}}_{S_0}(f(S_0, \ldots, S_n)) = \mathbb{E}_{S_0}\left(\frac{vol(S_n)}{vol(S_0)}f(S_0, \ldots, S_n)\right).$$

Next, we move to more general situations. We will need a stronger version of the Optional Stopping Theorem.

Theorem 19.129. *(Optional Stopping Theorem) Let (M_t) be a martingale with respect to a filtration $\{mathcalF_t\}$ having bounded increments; that is, $|M_{t+1} - M_t| \leq B$ for all y, where B is a non-random constant. Suppose that τ is a stopping time for $\{F_t\}$ with $\mathbb{E}(\tau) < \infty$. Then, $\mathbb{E}(M_\tau) = \mathbb{E}(M_0)$.*

Proof. Note that

$$|M_{\tau\wedge t}| = \left|\sum_{s=1}^{\tau\wedge t}(M_s - M_{s-1}) + M_0\right| \leq \sum_{s=1}^{\tau\wedge t}|M_s - M_{s-1}| + |M_0| \leq B\tau + |M_0|.$$

Since $\mathbb{E}(B\tau + |M_0|) < \infty$, by the Dominated Convergence Theorem and the Optional Stopping Theorem that we studied in the previous lecture,

$$\mathbb{E}(M_0) = \lim_{t\to\infty}\mathbb{E}(M_{\tau\wedge t}) = \mathbb{E}(M_\tau).$$

□

Consider now a Markov chain on subsets of $\mathcal{X}$. Suppose the current state is $S \subseteq \mathcal{X}$. Let U be a random variable which is uniform on $[0, 1]$. The next state of the chain is the set

$$\tilde{S} = \{y \in \mathcal{X}\colon \frac{Q(S, y)}{\pi(y)} \geq U\},$$

where $Q(x, y) = \pi(x)P(x, y)$ and $Q(S, y) = \sum_{x\in S} Q(x, y)$. The very definition gives

$$\mathbb{P}(y \in S_{t+1} | S_t) = \frac{Q(S_t, y)}{\pi(y)}.$$

Lemma 19.130. *If (S_t) is the evolving-set process associated to the transition matrix P, then*

$$P^t(x,y) = \frac{\pi(y)}{\pi(x)}\mathbb{P}_{\{x\}}(y \in S_t).$$

Proof. We will prove this by induction on t. When $t = 0$ the equality is trivial. Assume now that it holds for $t = s$. By conditioning on the position of th chain after s steps and then using the induction hypothesis, we have

$$P^{s+1}(x,y) = \sum_{z\in\mathcal{X}} P^s(x,z)P(z,y) = \sum_{z\in\mathcal{X}} \frac{\pi(z)}{\pi(x)}\mathbb{P}_{\{x\}}(z \in S_s)P(z,y).$$

By switching summation and expectation,

$$\begin{aligned}\sum_{z\in\mathcal{X}} \pi(z)\mathbb{P}_{\{x\}}(z \in S_s)P(z,y) &= \sum_{z\in\mathcal{X}} \mathbb{E}_{\{x\}}(1_{\{z\in S_s\}}\pi(z)P(z,y)) \\ &= \mathbb{E}_{\{x\}}\left(\sum_{z\in S_s} Q(z,y)\right) = \mathbb{E}_{\{x\}}(Q(S_s,y)).\end{aligned}$$

Therefore,

$$P^{s+1}(x,y) = \frac{1}{\pi(x)}\mathbb{E}_{\{x\}}(\pi(y)\mathbb{P}(y \in S_{s+1}|S_s)) = \frac{\pi(y)}{\pi(x)}\mathbb{P}_{\{x\}}(y \in S_{s+1}).$$

□

Next, we prove the martingale property for the general ESP.

Lemma 19.131. *The sequence $\{\pi(S_t)\}$ is a martingale.*

Proof. We have

$$\mathbb{E}(\pi(S_{t+1})|S_t) = \mathbb{E}\left(\sum_{z\in\mathcal{X}} 1_{\{z\in S_{t+1}\}}\pi(z)\middle| S_t\right).$$

To conclude the proof, we just need to note that the right-hand side above equals

$$\sum_{z\in\mathcal{X}} \mathbb{P}(z \in S_{t+1}|S_t)\pi(z) = \sum_{z\in\mathcal{X}} Q(S_t,z) = Q(S_t,\mathcal{X}) = \pi(S_t).$$

□

We need one more result about the ESP. Recall that $\Phi(S) = Q(S,S^c)/\pi(S)$ is the bottleneck ratio of the set S defined in previous lectures.

Lemma 19.132. *Suppose that the Markov chain is lazy, let $R_t = \pi(S_{t+1})/\pi(S_t)$, and let (U_t) be a sequence of independent random variables, each uniform on $[0,1]$, such that S_{t+1} is generated from S_t using U_{t+1}. Then*

$$\begin{aligned}\mathbb{E}(R_t|U_{t+1} \le 1/2, S_t = S) &= 1 + 2\Phi(S), \\ \mathbb{E}(R_t|U_{t+1} > 1/2, S_t = S) &= 1 - 2\Phi(S).\end{aligned}$$

Proof. Since the chain is lazy, $Q(y,y) \geq \pi(y)/2$, so if $y \notin S$, then

$$\frac{Q(S,y)}{\pi(y)} = \sum_{x \in S} \frac{Q(x,y)}{\pi(y)} \leq \sum_{x \neq y} \frac{Q(x,y)}{\pi(y)} = 1 - \frac{Q(y,y)}{\pi(y)} \leq \frac{1}{2}.$$

Given $U_{t+1} \leq \frac{1}{2}$, the distribution of U_{t+1} is uniform on $[0, 1/2]$. For $y \notin S$,

$$\mathbb{P}(\frac{Q(S,y)}{\pi(y)} \geq U_{t+1} \Big| U_{t+1} \leq 1/2, S_t = S) = 2\frac{Q(S,y)}{\pi(y)}.$$

Since $y \in S_{t+1}$ if and only if $U_{t+1} \leq Q(S_t, y)/\pi(y)$,

$$\mathbb{P}(y \in S_{t+1} | U_{t+1} \leq 1/2, S_t = S) = 2\frac{Q(S,y)}{\pi(y)} \quad \text{for } y \notin S.$$

Also, since $Q(S,y)/\pi(y) \geq Q(y,y)/\pi(y) \geq 1/2$ for $y \in S$, it follows that

$$\mathbb{P}(y \in S_{t+1} | U_{t+1} \leq 1/2, S_t = S) = 1 \quad \text{for } y \in S.$$

We have

$$\begin{aligned}
\mathbb{E}(\pi(S_{t+1}) | U_{t+1} \leq 1/2, S_t = S) &= \mathbb{E}\left(\sum_{y \in \mathcal{X}} 1_{\{y \in S_{t+1}\}} \pi(y) \Big| U_{t+1} \leq 1/2, S_t = S \right) \\
&= \sum_{y \in S} \pi(y) \mathbb{P}(y \in S_{t+1} | U_{t+1} \leq 1/2, S_t = S) \\
&\quad + \sum_{y \notin S} \pi(y) \mathbb{P}(y \in S_{t+1} | U_{t+1} \leq 1/2, S_t = S).
\end{aligned}$$

By the above,

$$\mathbb{E}(\pi(S_{t+1}) | U_{t+1} \leq 1/2, S_t = S) = \pi(S) + 2Q(S, S^c).$$

Since $(\pi(S_t))$ is a martingale (by the previous lemma), we have

$$\begin{aligned}
\pi(S) &= \mathbb{E}(\pi(S_{t+1}) | S_t = S) \\
&= \frac{1}{2}\mathbb{E}(\pi(S_{t+1}) | U_{t+1} \leq 1/2, S_t = S) + \frac{1}{2}\mathbb{E}(\pi(S_{t+1}) | U_{t+1} > 1/2, S_t = S) \\
&= \frac{\pi(S)}{2} + Q(S, S^c) + \frac{1}{2}\mathbb{E}(\pi(S_{t+1}) | U_{t+1} > 1/2, S_t = S).
\end{aligned}$$

Rearranging shows that

$$\mathbb{E}(\pi(S_{t+1}) | U_{t+1} > 1/2, S_t = S) = \pi(S) - 2Q(S, S^c).$$

We just need to divide both sides by $\pi(S)$ to obtain the result. □

We want to show that the ESP is a martingale that converges rapidly to the absorbing state. For this, we will use a straightforward quantitative version of Jensen's inequality:

Lemma 19.133. *For $\alpha \in [0, 1/2]$,*

$$\frac{\sqrt{1+2\alpha} + \sqrt{1-2\alpha}}{2} \leq \sqrt{1-\alpha^2} \leq 1 - \frac{\alpha^2}{2}.$$

Lemma 19.134. *Let (S_t) be the evolving-set process. If*

$$S_t^{\#} = \begin{cases} S_t & \text{if } \pi(S_t) \leq 1/2, \\ S_t^c & \text{otherwise,} \end{cases}$$

then if Φ_* denotes the bottleneck ratio of the chain we have

$$\mathbb{E}\left(\sqrt{\pi(S_{t+1}^{\#})/\pi(S_t^{\#})}\middle| S_t\right) \leq 1 - \frac{\Phi_*^2}{2}.$$

Proof. First, letting $R_t = \pi(S_{t+1})/\pi(S_t)$, applying Jensen's inequality shows

$$\begin{aligned}\mathbb{E}(\sqrt{R_t}|S_t = S) &= \frac{\mathbb{E}(\sqrt{R_t}|U_{t+1} \leq 1/2, S_t = S) + \mathbb{E}(\sqrt{R_t}|U_{t+1} > 1/2, S_t = S)}{2} \\ &\leq \frac{\sqrt{\mathbb{E}(\sqrt{R_t}|U_{t+1} \leq 1/2, S_t = S)} + \sqrt{\mathbb{E}(\sqrt{R_t}|U_{t+1} > 1/2, S_t = S)}}{2}.\end{aligned}$$

Applying Lemma 19.132 and Lemma 19.133 shows that, for $\pi(S) \leq 1/2$,

$$\mathbb{E}(\sqrt{R_t}|S_t = S) \leq \frac{\sqrt{1+2\Phi(S)} + \sqrt{1-2\Phi(S)}}{2} \leq 1 - \frac{\Phi(S)^2}{2} \leq 1 - \frac{\Phi_*^2}{2}.$$

Now assume that $\pi(S_t) \leq 1/2$. Then

$$\sqrt{\pi(S_{t+1}^{\#})/\pi(S_t^{\#})} = \sqrt{\pi(S_{t+1}^{\#})/\pi(S_t)} \leq \sqrt{\pi(S_{t+1})/\pi(S_t)},$$

from where the result follows. If $\pi(S_t) > 1/2$, then replace S_t by S_t^c in the previous argument. □

Now we are in position to present a useful inequality regarding mixing time. We studied this inequality previously using eigenvalues for reversible chains.

Theorem 19.135. *Let P be a lazy irreducible transition matrix, so that $P(x,x) \geq \frac{1}{2}$ for all $x \in \mathcal{X}$, with stationary distribution π. The mixing time satisfies*

$$t_{\text{mix}}(\varepsilon) \leq t_{\text{mix}}^{(\infty)}(\varepsilon) \leq \frac{2}{\Phi_*^2} \log\left(\frac{1}{\varepsilon \pi_{\min}}\right).$$

Proof. From Lemma 19.134 we know

$$\mathbb{E}\left(\sqrt{\pi(S_{t+1}^{\#})}\right) \leq \mathbb{E}\left(\sqrt{\pi(S_t^{\#})}\right)\left(1 - \frac{\Phi_*^2}{2}\right).$$

Iterating,

$$\mathbb{E}_S\left(\sqrt{\pi(S_t^{\#})}\right) \le \left(1-\frac{\Phi_*^2}{2}\right)^t \sqrt{\pi(S)}.$$

Since $\sqrt{\pi_{\min}}\mathbb{P}_S(S_t^{\#} \neq \emptyset) \le \mathbb{E}_S\left(\sqrt{\pi(S_t^{\#})}\right)$, we have

$$\mathbb{P}_S(S_t^{\#} \neq \emptyset) \le \sqrt{\frac{\pi(S)}{\pi_{\min}}}\left(1-\frac{\Phi_*^2}{2}\right)^t.$$

Therefore, this implies

$$\mathbb{P}_S(S_t^{\#} \neq \emptyset \text{ for all } t \ge 0) = \mathbb{P}_S\left(\bigcap_{t=1}^{\infty}\{S_t^{\#} \neq \emptyset\}\right) = \lim_{t\to\infty}\mathbb{P}_S(S_t^{\#} \neq \emptyset) = 0.$$

That is, $(S_t\#)$ is eventually absorbed in the state $\emptyset$. Let

$$\tau = \min\{t \ge 0 \colon S_t^{\#} = \emptyset\}.$$

We have $S_\tau \in \{\emptyset, \mathcal{X}\}$ and $\pi(S_\tau) = 1_{\{S_\tau = \mathcal{X}\}}$. Note that the Optional Stopping Theorem gives

$$\pi(x) = \mathbb{E}_{\{x\}}(\pi(S_0)) = \mathbb{E}_{\{x\}}(\pi(S_\tau)) = \mathbb{P}_{\{x\}}(S_\tau = \mathcal{X}).$$

Next, observe that

$$\begin{aligned}|P^t(x,y) - \pi(y)| &= \frac{\pi(y)}{\pi(x)}|\mathbb{P}_{\{x\}}(y \in S_t) - \pi(x)| \\ &= \frac{\pi(y)}{\pi(x)}|\mathbb{P}_{\{x\}}(y \in S_t) - \mathbb{P}_{\{x\}}(S_t = \mathcal{X})|.\end{aligned}$$

Using the identity

$$\begin{aligned}\mathbb{P}_{\{x\}}(y \in S_t) &= \mathbb{P}_{\{x\}}(y \in S_t, \tau > t) + \mathbb{P}_{\{x\}}(y \in S_t, \tau \le t) \\ &= \mathbb{P}_{\{x\}}(y \in S_t) + \mathbb{P}_{\{x\}}(S_t = \mathcal{X}, \tau \le t)\end{aligned}$$

we get

$$\begin{aligned}|P^t(x,y) - \pi(y)| &= \frac{\pi(y)}{\pi(x)}|\mathbb{P}_{\{x\}}(y \in S_t) + \mathbb{P}_{\{x\}}(S_t = \mathcal{X}, \tau \le t)| \\ &\le \frac{\pi(y)}{\pi(x)}\mathbb{P}_{\{x\}}(\tau > t).\end{aligned}$$

Recalling the definition of $d^{(\infty)}(t)$,

$$d(t) \le d^{(\infty)}(t) \le \max_{x,y}\frac{|P^t(x,y) - \pi(y)|}{\pi(y)} \le \frac{1}{\pi_{\min}}\left(1-\frac{\Phi_*^2}{2}\right)^t.$$

It follows that if $t \ge \frac{2}{\Phi_*^2}\log\left(\frac{1}{\varepsilon\pi_{\min}}\right)$, then $d(t) \le d^{(\infty)}(t) \le \varepsilon$. □

Markov Chains Mixing Times **December 21, 2022**

Lecture 20: Cutoff for a stratified random walk on the hypercube

Lecturer: Yuval Peres Week: 10

Mixing times of Markov chains

Let P be the transition matrix of an irreducible finite Markov chain with stationary distribution π. The total-variation distance between $P^t(x,)$ and $\pi(.)$ is

$$d_x(t) = \max_{A\subset\Omega}(P^t(x,A) - \pi(A)) = \sum_{y\in\Omega}(P^t(x,y) - \pi(y))_+,$$

and the mixing time is defined as

$$t_{mix}(\epsilon) = inf\{t \geq 0, d(t) \leq \epsilon\}$$

where d(t) = $\max_{x\subset\Omega}(d_x(t))$.
A sequence of Markov chains is said to have cutoff when for all $\epsilon \in$ (0, 1),

$$\frac{t^{(n)}_{mix}(1-\epsilon)}{t_{mix}(\epsilon)} \underset{n\to\infty}{\longrightarrow} 1$$

A matrix random walk

Let $SL_n(Z_2)$ be the set of invertible matrices with coefficients in Z_2.

$$|SL_n(Z_2) = 2^{n^2}$$

Consider the following Markov chain on $SL_n(Z_2)$:

1. Start at the identity matrix I.
2. At each step, pick two distinct rows at random and add one to the other(in Z_2).

Stationary distribution : uniform.
Kassabov (2005) : $t_{mix} = 0(n^3)$
Lower bound : $t_{mix} = \Omega(\frac{n^2}{logn})$(counting argument).
Remark : Diam $\frac{n^2}{logn}$, Andrén et al.(2007).

Cryptographic motivation

Time-based authentication protocols

A verifier wants to check the identity of a prover, based on the time needed to answer a challenge.

Example (Sotiraki, 2016)

- The prover runs the matrix chain up to time t to create a matrix $A \in SL_n(Z^2)$, and makes it public.
- Whenever he wants to authenticate, the prover asks the verifier for a vector x $\in \{0,1\}^n$.
- The challenge is to compute $y = Ax$.

As the prover knows the trajectory of the chain, he can answer in time of order t. Also, if t is large enough, a dishonest party will not be able to distinguish A from a randomly chosen matrix, and its best solution would be to perform usual matrix-vector multiplication, which typically takes about n^2 operations. Can t be chosen much smaller than n^2?
It is reasonable to assume that if each column is mixed, then no dishonest party would be able to distinguish A from uniform (in polynomial time).

The first column Markov chain

Let X = $\{0,1\}^n \setminus 0$ and consider the following Markov chain $\{Z_t\}_{t\geq 0}$ on X : if the current state is x, choose a pair (i, j) of distinct coordinates, and replace x(j) by x(j) + x(i) (where x(i) denotes the bit at location i).

- Falls into the class of what Chung and Graham (1997) call stratified random walks : the transition probabilities may depend on the current Hamming weight (number of ones in the string).
- They showed that for all $0 < \epsilon < 1$, there exists c_ϵ such that

  $$t_{mix}(\epsilon) \leq c_\epsilon nlogn$$

  and asked whether the dependence in ϵ could be removed.

Theorem 23.1 (Ben-Hamou and Peres): The chain $\{Z_t\}$ has total-variation cutoff at time $\frac{3}{2}nlogn$ with window n, i.e.

$$\lim_{\alpha\to+\infty}\lim_{n\to+\infty} inf(d(\frac{3}{2}nlogn - \alpha n)) = 1$$

and

$$\lim_{\alpha\to+\infty}\lim_{n\to+\infty} sup(d(\frac{3}{2}nlogn + \alpha n)) = 0$$

The Aldous setting

Both the matrix walk and the one-column walk fall into a class of chains introduced by Aldous in 1994.
Consider a finite connected graph G = (V, E) on n vertices, and a finite group Γ.

- Initially, each vertex v ∈ V is labelled by an element σ(v) ∈ Γ, in such a way that {σ(v), v ∈ V } generates Γ.
- At each step, pick a directed edge (u, v) at random and replace σ(v) by σ(u) ± σ(v).
- Irreducible on $\Gamma^* \subset \Gamma^n$, the subset of vectors $(g_1, ..., g_n)$ such that $g_1, ..., g_n$ generate the group.
- May be used to obtain a random generating set of Γ.

Examples

- Take G = K_n and Γ = Z_q .
  Diaconis and Saloff-Coste (1996), Chung and Graham (1996) showed t_{mix} = O($c_q n^2 logn$).
  Kassabov (2005) : t_mix = O(n^2).
- With G = K_n and Γ = Z_2^n: the matrix walk.
- With G = K_n and Γ = Z_2 : the one-column walk.
- With G = directed line and Γ = Z_2 : the East process.

See Diaconis and Saloff-Coste (1996) for other examples and results on general graphs.

Outline of the proof of cutoff

Recall the standard lazy random walk on the hypercube $0, 1^n$: at each step, pick a coordinate uniformly at random and randomize the bit at this location.

Theorem 20.136 (***Aldous (1983)***)**.** *The lazy RW on the hypercube has cutoff at* $\frac{nlogn}{2}$ *with window n.*

Proof. Basic intuition : starting from $x_0 = (1, 0, ..., 0)$, the stratified RW (Z_t) on $\chi = \{0, 1\}^n \setminus \{0\}$ needs n log n steps to have "enough ones", and then behaves like a lazy standard RW . Altogether, it mixes in time $\frac{3}{2} n \log n$. □

First step : The Hamming weight

Proposition 20.137. *Let $H_t = H(Z_t)$ be the projected chain given by the Hamming weight. Then, (H_t) has cutoff at $\frac{3}{2}nlogn$ with window n.*

Test Cutoff for birth-and-death chains
$\Longleftrightarrow$
Concentration of hitting time of stationary distribution median. (Ding, Lubetzky, Peres (2010)).

$$T_{n/2} = inf\{t \geq 0, H_t = \frac{n}{2}\}$$

We have

$$||\mathbb{P}_i(H_t \in .) - \pi_H(.)||_{TV} = \mathbb{P}_i(T_{n/2} > t)$$

Expectation : $E_1 T_{n/2} \sim \frac{3}{2}nlogn$, and $T_{n/2}$ is well concentrated.

Second Step: Go from H_t to Z_t.

1. By time nlogn + αn, the Hamming weight is larger than n/3.

2. Let $x \in \chi$ with $\bar{x} :=$ H(x) $\geq$ n/3. Without loss of generality,

$$x = (\underbrace{1, \ldots, 1}_{\bar{x}}, \underbrace{0, \ldots, 0}_{\bar{x}})$$

   Goal: $d_x(\frac{nlogn}{2} + \alpha n) \leq \epsilon(\alpha) \underset{\alpha \to \infty}{\longrightarrow} 0$.

3. Define

$$W_t = (\sum_{i=1}^{\bar{x}} Z_t(i), \sum_{i=\bar{x}+1}^{n} Z_t(i)) := (X_t, Y_t)$$

   It is enough to couple $W_t = (X_t, Y_t)$ and $\tilde{W}_t = (\tilde{X}_t, \tilde{Y}_t)$, for two walks with the same initial Hamming weight $\bar{x}$.

4. Letting $\tau = inf t \geq 0, W_t = \tilde{W}_t$, construct a coupling such that

$$\max_{\substack{i,j \\ i+j=\bar{x}}} \mathbb{P}_{(\bar{x},0),(i,j)}(\tau > \tfrac{nlogn}{2} + \alpha n) \leq \epsilon(\alpha) \text{ with } \epsilon(\alpha) \underset{\alpha \to \infty}{\longrightarrow} 0$$

The Coupling

1. Start at :
$$\left\{ \begin{array}{ll} Z_0 & with \;\; W_0 = (\bar{x}, 0) \\ \tilde{Z}_0 & with \;\; \tilde{W}_0 = (i,j) with i + j = \bar{x} \end{array} \right\} .$$

2. At step t ≥ 0, assuming $W_t \neq \tilde{W}_t$ and $H(Z_t) = H(\tilde{Z}_t)$, choose

   - a uniform matching between the ones in Z_t and the ones in $\tilde{Z}_t$,
   - a uniform matching between the zeros in Z_t and the zeros in $\tilde{Z}_t$.

   Let π_t the corresponding permutation on indices.

3. If $Z_t + 1$ is obtained by the update
$$Z_t(j_t) \longrightarrow Z_t(j_t) + Z_t(i_t)$$
   then move to $\tilde{Z}_{t+1}$ by the update
$$\tilde{Z}_t(\pi_t(j_t)) \longrightarrow \tilde{Z}_t(\pi_t(j_t)) + \tilde{Z}_t(\pi_t(i_t))$$

4. Analysis of the process $D_t = X_t - \tilde{X}_t$.

   - After $\frac{nlogn}{2}$, the variable D_t has fallen below order $\sqrt{n}$.
   - Comparison to SRW on Z : after additional time of order n, the variable D_t has touched 0.