Rapid Mixing in Symmetric MC's

January 31, 2018

MC's with Constrictions

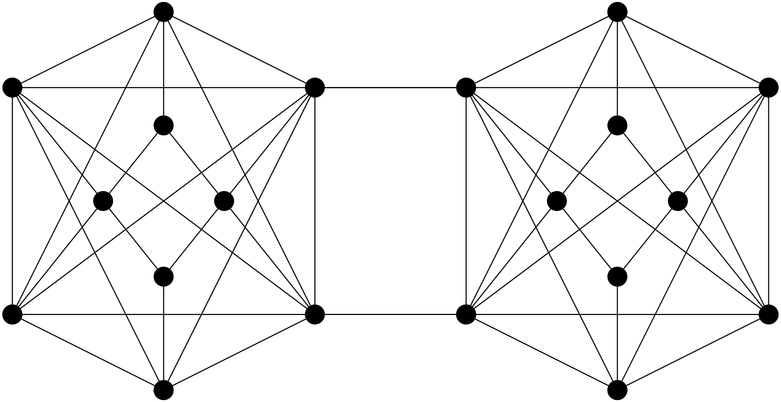
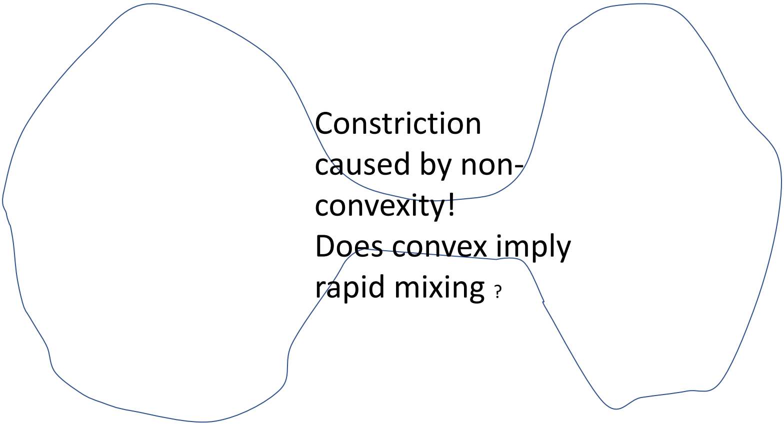


Figure: A network with a constriction. All edges have weight 1.


Constriction
caused by non-
convexity!
Does convex imply
rapid mixing ?

Conductance, Rapid Mixing of Symmetric MC's

- Deal with Symmetric Chains first for better intuition. Same theorem for all time-reversible MC's. Will prove later.

Conductance, Rapid Mixing of Symmetric MC's

- Deal with Symmetric Chains first for better intuition. Same theorem for all time-reversible MC's. Will prove later.
- For a connected MC, *conductance* $\Phi(S)$ of a subset S of states is defined as $\frac{\sum_{i\in S, j\in \bar{S}} \pi_i P_{ij}}{\min(\pi(S), \pi(\bar{S}))}$. (If symmetric, can drop π_i and replace $\pi(S)$ by $|S|$. We stick to the first form.)

Conductance, Rapid Mixing of Symmetric MC's

- Deal with Symmetric Chains first for better intuition. Same theorem for all time-reversible MC's. Will prove later.
- For a connected MC, *conductance* $\Phi(S)$ of a subset S of states is defined as $\frac{\sum_{i\in S, j\in \bar{S}} \pi_i P_{ij}}{\min(\pi(S), \pi(\bar{S}))}$. (If symmetric, can drop π_i and replace $\pi(S)$ by $|S|$. We stick to the first form.)
- *Conductance* Φ of the MC is min over all S of $\Phi(S)$.

Conductance, Rapid Mixing of Symmetric MC's

- Deal with Symmetric Chains first for better intuition. Same theorem for all time-reversible MC's. Will prove later.
- For a connected MC, *conductance* $\Phi(S)$ of a subset S of states is defined as $\frac{\sum_{i\in S, j\in \bar{S}} \pi_i P_{ij}}{\min(\pi(S),\pi(\bar{S}))}$. (If symmetric, can drop π_i and replace $\pi(S)$ by $|S|$. We stick to the first form.)
- *Conductance* Φ of the MC is min over all S of $\Phi(S)$.
- **Informal Theorem** MC converges to stationarity in time $O(1/\Phi^2)$.

Conductance, Rapid Mixing of Symmetric MC's

- Deal with Symmetric Chains first for better intuition. Same theorem for all time-reversible MC's. Will prove later.
- For a connected MC, *conductance* $\Phi(S)$ of a subset S of states is defined as $\frac{\sum_{i\in S, j\in \bar{S}} \pi_i P_{ij}}{\min(\pi(S), \pi(\bar{S}))}$. (If symmetric, can drop π_i and replace $\pi(S)$ by $|S|$. We stick to the first form.)
- *Conductance* Φ of the MC is min over all S of $\Phi(S)$.
- **Informal Theorem** MC converges to stationarity in time $O(1/\Phi^2)$.
- Lower bound: One needs time at least $\Omega(1/\Phi)$ (Simpler than Theorem).

Conductance, Rapid Mixing of Symmetric MC's

- Deal with Symmetric Chains first for better intuition. Same theorem for all time-reversible MC's. Will prove later.
- For a connected MC, *conductance* $\Phi(S)$ of a subset S of states is defined as $\frac{\sum_{i\in S, j\in \bar{S}} \pi_i P_{ij}}{\min(\pi(S),\pi(\bar{S}))}$. (If symmetric, can drop π_i and replace $\pi(S)$ by $|S|$. We stick to the first form.)
- *Conductance* Φ of the MC is min over all S of $\Phi(S)$.
- **Informal Theorem** MC converges to stationarity in time $O(1/\Phi^2)$.
- Lower bound: One needs time at least $\Omega(1/\Phi)$ (Simpler than Theorem).
- Formally, for any $\varepsilon > 0$, the ε mixing time of a symmetric connected MC is min t such that $\left|\mathbf{a(t)} - \frac{1}{n}\mathbf{1}\right|_1 \leq \varepsilon$, where, $\mathbf{a(t)} = \frac{1}{t}(\mathbf{p(0)} + \mathbf{p(1)} + \cdots + \mathbf{p(t-1)})$.
- **Rapid Mixing Theorem**: The ε mixing time of a connected symmetric MC is $O(\ln n/\Phi^2\varepsilon^3)$

Idea of Proof

- Recall proof of Fundamental Theorem: Showed first that if we run one step starting from **a**(**t**), the change in the probability vector is $O(1/t)$ in l_1 norm. Let $v_i = a(t)_i/\pi_i = na_i(t)$.

Idea of Proof

- Recall proof of Fundamental Theorem: Showed first that if we run one step starting from $\mathbf{a}(\mathbf{t})$, the change in the probability vector is $O(1/t)$ in l_1 norm. Let $v_i = a(t)_i/\pi_i = na_i(t)$.
- Start with $\mathbf{a}(\mathbf{t})$, run 1 step. *Net probability flow* $h(i,j)$ from i to j: $h(i,j) = \frac{1}{n}(v_i P_{ij} - v_j P_{ji}) = \frac{1}{n}(v_i - v_j)P_{ij}$ by symmetry.

Idea of Proof

- Recall proof of Fundamental Theorem: Showed first that if we run one step starting from $\mathbf{a}(\mathbf{t})$, the change in the probability vector is $O(1/t)$ in l_1 norm. Let $v_i = a(t)_i/\pi_i = na_i(t)$.
- Start with $\mathbf{a}(\mathbf{t})$, run 1 step. *Net probability flow* $h(i, j)$ from i to j: $h(i, j) = \frac{1}{n}(v_i P_{ij} - v_j P_{ji}) = \frac{1}{n}(v_i - v_j)P_{ij}$ by symmetry.
- If $v_i > v_j$, $h(i, j) > 0$; net loss. Probability flows from "heavy" to "light" vertices.
- Probability Flow Lemma:

Idea of Proof

- Recall proof of Fundamental Theorem: Showed first that if we run one step starting from $\mathbf{a(t)}$, the change in the probability vector is $O(1/t)$ in l_1 norm. Let $v_i = a(t)_i/\pi_i = na_i(t)$.
- Start with $\mathbf{a(t)}$, run 1 step. *Net probability flow* $h(i,j)$ from i to j: $h(i,j) = \frac{1}{n}(v_i P_{ij} - v_j P_{ji}) = \frac{1}{n}(v_i - v_j)P_{ij}$ by symmetry.
- If $v_i > v_j$, $h(i,j) > 0$; net loss. Probability flows from "heavy" to "light" vertices.
- Probability Flow Lemma:
  - Take two disjoint sets of vertices - with each vertex in 1st set heavier by at least $\gamma > 0$ than each vertex in second set. This gives lower bound in terms of γ on net flow from 1st set to 2nd.

Idea of Proof

- Recall proof of Fundamental Theorem: Showed first that if we run one step starting from $\mathbf{a}(\mathbf{t})$, the change in the probability vector is $O(1/t)$ in l_1 norm. Let $v_i = a(t)_i/\pi_i = na_i(t)$.
- Start with $\mathbf{a}(\mathbf{t})$, run 1 step. *Net probability flow* $h(i,j)$ from i to j: $h(i,j) = \frac{1}{n}(v_i P_{ij} - v_j P_{ji}) = \frac{1}{n}(v_i - v_j)P_{ij}$ by symmetry.
- If $v_i > v_j$, $h(i,j) > 0$; net loss. Probability flows from "heavy" to "light" vertices.
- Probability Flow Lemma:
    - Take two disjoint sets of vertices - with each vertex in 1st set heavier by at least $\gamma > 0$ than each vertex in second set. This gives lower bound in terms of γ on net flow from 1st set to 2nd.
    - But we showed (and will recap), net-flow from all heavy vertices to all light vertices which is $|\mathbf{a} - \mathbf{a}P|_1$ is $O(1/t) \to 0$.

Idea of Proof

- Recall proof of Fundamental Theorem: Showed first that if we run one step starting from $\mathbf{a}(\mathbf{t})$, the change in the probability vector is $O(1/t)$ in l_1 norm. Let $v_i = a(t)_i/\pi_i = na_i(t)$.
- Start with $\mathbf{a}(\mathbf{t})$, run 1 step. *Net probability flow* $h(i,j)$ from i to j: $h(i,j) = \frac{1}{n}(v_i P_{ij} - v_j P_{ji}) = \frac{1}{n}(v_i - v_j)P_{ij}$ by symmetry.
- If $v_i > v_j$, $h(i,j) > 0$; net loss. Probability flows from "heavy" to "light" vertices.
- Probability Flow Lemma:
    - Take two disjoint sets of vertices - with each vertex in 1st set heavier by at least $\gamma > 0$ than each vertex in second set. This gives lower bound in terms of γ on net flow from 1st set to 2nd.
    - But we showed (and will recap), net-flow from all heavy vertices to all light vertices which is $|\mathbf{a} - \mathbf{a}P|_1$ is $O(1/t) \to 0$.
    - So heavy vertices cannot be much heavier than the light ones (if they were, flow would be greater than $1/t$.)

Proof of Rapid Mixing - I

- Vertex i is **Heavy** if $v_i > 1$. Renumber so that $v_1 \geq v_2 \ldots v_{i_0} > 1 \geq v_{i_0+1} \geq \ldots v_n$. i_0 is last heavy vertex.

Proof of Rapid Mixing - I

- Vertex i is **Heavy** if $v_i > 1$. Renumber so that $v_1 \geq v_2 \ldots v_{i_0} > 1 \geq v_{i_0+1} \geq \ldots v_n$. i_0 is last heavy vertex.
- **Claim** $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = 2\sum_{i=1}^{i_0}(a_i - (1/n))$. [So need only to measure how heavy the heavy vertices are.]

Proof of Rapid Mixing - I

- Vertex i is **Heavy** if $v_i > 1$. Renumber so that $v_1 \geq v_2 \ldots v_{i_0} > 1 \geq v_{i_0+1} \geq \ldots v_n$. i_0 is last heavy vertex.
- **Claim** $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = 2\sum_{i=1}^{i_0}(a_i - (1/n))$. [So need only to measure how heavy the heavy vertices are.]
  - **Proof**: Since $\sum_{i=1}^{n} a_i = 1$, $\sum_{i=1}^{i_0}(a_i - (1/n)) = \sum_{i=i_0+1}^{n}((1/n) - a_i)$. [Total heaviness of heavy vertices equals total lightness of light vertices.]

Proof of Rapid Mixing - I

- Vertex i is **Heavy** if $v_i > 1$. Renumber so that $v_1 \geq v_2 \dots v_{i_0} > 1 \geq v_{i_0+1} \geq \dots v_n$. i_0 is last heavy vertex.
- **Claim** $||\mathbf{a} - (\mathbf{1/n})||_1 = 2\sum_{i=1}^{i_0}(a_i - (1/n))$. [So need only to measure how heavy the heavy vertices are.]
  - **Proof**: Since $\sum_{i=1}^{n} a_i = 1$, $\sum_{i=1}^{i_0}(a_i - (1/n)) = \sum_{i=i_0+1}^{n}((1/n) - a_i)$. [Total heaviness of heavy vertices equals total lightness of light vertices.]
  - Since $||\mathbf{a} - (\mathbf{1/n})||_1 = \sum_{i=1}^{n} |a_i - (1/n)|$, claim follows.

Proof of Rapid Mixing - I

- Vertex i is **Heavy** if $v_i > 1$. Renumber so that $v_1 \geq v_2 \ldots v_{i_0} > 1 \geq v_{i_0+1} \geq \ldots v_n$. i_0 is last heavy vertex.
- **Claim** $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = 2\sum_{i=1}^{i_0}(a_i - (1/n))$. [So need only to measure how heavy the heavy vertices are.]
  - **Proof**: Since $\sum_{i=1}^{n} a_i = 1$, $\sum_{i=1}^{i_0}(a_i - (1/n)) = \sum_{i=i_0+1}^{n}((1/n) - a_i)$. [Total heaviness of heavy vertices equals total lightness of light vertices.]
  - Since $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = \sum_{i=1}^{n} |a_i - (1/n)|$, claim follows.
- Define $f : [0, i_0/n]$ by $f(x) = v_i - 1$ for $x \in [(i-1)/n, i/n)$ for $i = 1, 2 \ldots, i_0$. [Picture coming.]

Proof of Rapid Mixing - I

- Vertex i is **Heavy** if $v_i > 1$. Renumber so that $v_1 \geq v_2 \ldots v_{i_0} > 1 \geq v_{i_0+1} \geq \ldots v_n$. i_0 is last heavy vertex.
- **Claim** $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = 2\sum_{i=1}^{i_0}(a_i - (1/n))$. [So need only to measure how heavy the heavy vertices are.]
  - **Proof**: Since $\sum_{i=1}^{n} a_i = 1$, $\sum_{i=1}^{i_0}(a_i - (1/n)) = \sum_{i=i_0+1}^{n}((1/n) - a_i)$. [Total heaviness of heavy vertices equals total lightness of light vertices.]
  - Since $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = \sum_{i=1}^{n} |a_i - (1/n)|$, claim follows.
- Define $f : [0, i_0/n]$ by $f(x) = v_i - 1$ for $x \in [(i-1)/n, i/n)$ for $i = 1, 2 \ldots, i_0$. [Picture coming.]
- $\sum_{i=1}^{i_0}(a_i - (1/n)) = \frac{1}{n}\sum_{i=1}^{i_0}(v_i - 1) \ = \ \int_0^{i_0/n} f(x)\, dx.$

Proof of Rapid Mixing - I

- Vertex i is **Heavy** if $v_i > 1$. Renumber so that $v_1 \geq v_2 \ldots v_{i_0} > 1 \geq v_{i_0+1} \geq \ldots v_n$. i_0 is last heavy vertex.
- **Claim** $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = 2\sum_{i=1}^{i_0}(a_i - (1/n))$. [So need only to measure how heavy the heavy vertices are.]
  - **Proof**: Since $\sum_{i=1}^{n} a_i = 1$, $\sum_{i=1}^{i_0}(a_i - (1/n)) = \sum_{i=i_0+1}^{n}((1/n) - a_i)$. [Total heaviness of heavy vertices equals total lightness of light vertices.]
  - Since $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = \sum_{i=1}^{n} |a_i - (1/n)|$, claim follows.
- Define $f : [0, i_0/n]$ by $f(x) = v_i - 1$ for $x \in [(i-1)/n, i/n)$ for $i = 1, 2 \ldots, i_0$. [Picture coming.]
- $\sum_{i=1}^{i_0}(a_i - (1/n)) = \frac{1}{n}\sum_{i=1}^{i_0}(v_i - 1) \ = \ \int_0^{i_0/n} f(x)\, dx$.
- Henceforth: Want to prove: $\int_0^{i_0/n} f(x)\, dx \leq \varepsilon$.

Proof of Rapid Mixing - I

- Vertex i is **Heavy** if $v_i > 1$. Renumber so that $v_1 \geq v_2 \ldots v_{i_0} > 1 \geq v_{i_0+1} \geq \ldots v_n$. i_0 is last heavy vertex.
- **Claim** $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = 2\sum_{i=1}^{i_0}(a_i - (1/n))$. [So need only to measure how heavy the heavy vertices are.]
  - **Proof**: Since $\sum_{i=1}^{n} a_i = 1$, $\sum_{i=1}^{i_0}(a_i - (1/n)) = \sum_{i=i_0+1}^{n}((1/n) - a_i)$. [Total heaviness of heavy vertices equals total lightness of light vertices.]
  - Since $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1 = \sum_{i=1}^{n} |a_i - (1/n)|$, claim follows.
- Define $f : [0, i_0/n]$ by $f(x) = v_i - 1$ for $x \in [(i-1)/n, i/n)$ for $i = 1, 2 \ldots, i_0$. [Picture coming.]
- $\sum_{i=1}^{i_0}(a_i - (1/n)) = \frac{1}{n}\sum_{i=1}^{i_0}(v_i - 1) \;=\; \int_0^{i_0/n} f(x)\,dx$.
- Henceforth: Want to prove: $\int_0^{i_0/n} f(x)\,dx \leq \varepsilon$.
- $||\mathbf{a} - (\mathbf{1}/\mathbf{n})||_1$ called **Total Variation**.

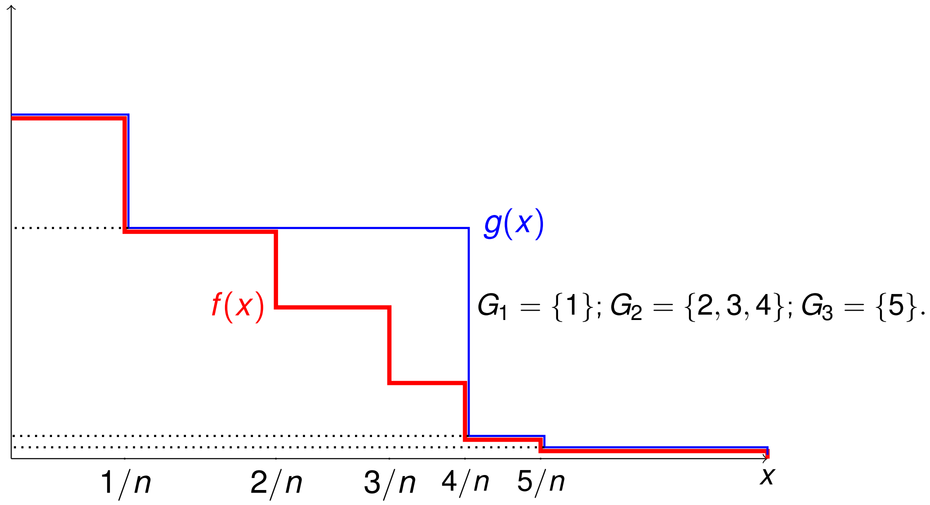


Figure: Bounding l_1 distance.

Proof-II: Groups and Probability Flows

1. Recall central to proof: **Probability Flow Lemma** Verbally: “Take two disjoint sets of vertices - with each vertex in 1st set heavier by at least $\gamma > 0$ than each vertex in second set. Get a lower bound in terms of γ, the “heaviness gap” on net flow from 1st set to 2nd.”

Proof-II: Groups and Probability Flows

1. Recall central to proof: **Probability Flow Lemma** Verbally: "Take two disjoint sets of vertices - with each vertex in 1st set heavier by at least $\gamma > 0$ than each vertex in second set. Get a lower bound in terms of γ, the "heaviness gap" on net flow from 1st set to 2nd."
2. Will implement this by dividing $\{1, 2, \ldots, i_0\}$ into groups $G_1, G_2, \ldots G_r$, where, if G_s ends in some k, G_{s+1} is $\{k+1, k+2, \ldots, l\}$ for some $l \geq k+1$. Will specify the technical detail of groups later.

Proof-II: Groups and Probability Flows

1. Recall central to proof: **Probability Flow Lemma** Verbally: "Take two disjoint sets of vertices - with each vertex in 1st set heavier by at least $\gamma > 0$ than each vertex in second set. Get a lower bound in terms of γ, the "heaviness gap" on net flow from 1st set to 2nd."
2. Will implement this by dividing $\{1, 2, \ldots, i_0\}$ into groups $G_1, G_2, \ldots G_r$, where, if G_s ends in some k, G_{s+1} is $\{k+1, k+2, \ldots, l\}$ for some $l \geq k+1$. Will specify the technical detail of groups later.
3. But, point of groups will be: we will take as "1st set" $G_1 \cup G_2 \cup \ldots G_s$ (for each s) and "2nd set" $G_{s+2}, G_{s+3}, \ldots$. In other words, "net flow from 1st set to 2nd set" is really the "flow across" G_{s+1}. To implement the intuitive idea, we need to (i) express the total variation distance in terms of the γ above, which we do first (this is the more technical part) and (ii) prove Prob. Flow Lemma, later. (See (1)).

Groups and Prob Flows -II

1. Now, more rigorously: Assume we have the groups. Define for $s = 1, 2, \ldots, r$: $u_s = \text{Max}_{j \in G_s} v_j$ and $u_{r+1} = 1$. Define $g(x) = u_s - 1 \text{for} x \in \cup_{i \in G_s} [(i-1)/n, i/n)$.

Groups and Prob Flows -II

1. Now, more rigorously: Assume we have the groups. Define for $s = 1, 2, \ldots, r$: $u_s = \text{Max}_{j \in G_s} v_j$ and $u_{r+1} = 1$. Define $g(x) = u_s - 1 \text{for} x \in \cup_{i \in G_s} [(i-1)/n, i/n)$.
2. $g(x) \geq f(x) \forall x$. So enough to bound $\int g(x)$.

Groups and Prob Flows -II

1. Now, more rigorously: Assume we have the groups. Define for $s = 1, 2, \ldots, r$: $u_s = \text{Max}_{j \in G_s} v_j$ and $u_{r+1} = 1$. Define $g(x) = u_s - 1 \text{for} x \in \cup_{i \in G_s} [(i-1)/n, i/n)$.
2. $g(x) \geq f(x) \forall x$. So enough to bound $\int g(x)$.
3. Intuition for next assertion:
$\int_0^{i_0/n} \underbrace{g(x)}_{Y} \underbrace{dx}_{dX} = \underbrace{[xg(x)]_0^{i_0/n}}_{0} - \int_0^{i_0/n} xg'(x) dx$.

Groups and Prob Flows -II

1. Now, more rigorously: Assume we have the groups. Define for $s = 1, 2, \ldots, r$: $u_s = \text{Max}_{j \in G_s} v_j$ and $u_{r+1} = 1$. Define $g(x) = u_s - 1 \text{for} x \in \cup_{i \in G_s} [(i-1)/n, i/n)$.
2. $g(x) \geq f(x) \forall x$. So enough to bound $\int g(x)$.
3. Intuition for next assertion: $\int_0^{i_0/n} \underbrace{g(x)}_{Y} \underbrace{dx}_{dX} = \underbrace{[xg(x)]_0^{i_0/n}}_{0} - \int_0^{i_0/n} xg'(x)dx$.
4. Total Variation $= \sum_{s=1}^{r} \pi(G_1 \cup G_2 \cup \ldots G_s)(u_s - u_{s+1}) \ldots\ldots(*)$

Groups and Prob Flows -II

1. Now, more rigorously: Assume we have the groups. Define for $s = 1, 2, \ldots, r$: $u_s = \text{Max}_{j \in G_s} v_j$ and $u_{r+1} = 1$. Define $g(x) = u_s - 1 \text{for} x \in \cup_{i \in G_s} [(i-1)/n, i/n)$.
2. $g(x) \geq f(x) \forall x$. So enough to bound $\int g(x)$.
3. Intuition for next assertion: $\int_0^{i_0/n} \underbrace{g(x)}_{Y} \underbrace{dx}_{dX} = \underbrace{[xg(x)]_0^{i_0/n}}_{0} - \int_0^{i_0/n} xg'(x)dx$.
4. Total Variation $= \sum_{s=1}^r \pi(G_1 \cup G_2 \cup \ldots G_s)(u_s - u_{s+1}).....(*)$
5. Direct Proof of (*), by expansion: $\sum_{s=1}^r \pi(G_1 \cup G_2 \cup \ldots G_s)(u_s - u_{s+1}) = \pi(G_1)(u_1 - u_2) + (\pi(G_1) + \pi(G_2))(u_2 - u_3) + \cdots + (\pi(G_1) + \pi(G_2) + \cdots + \pi(G_r))(u_r - 1) = \pi(G_1)u_1 + \pi(G_2)u_2 + \pi(G_3)u_3 + \cdots + \pi(G_r)u_r - (\pi(G_1) + \pi(G_2) + \cdots + \pi(G_r)). = \sum_{s=1}^r \pi(G_s)(u_s - 1)$.

Groups and Prob Flows -II

1. Now, more rigorously: Assume we have the groups. Define for $s = 1, 2, \ldots, r$: $u_s = \text{Max}_{j \in G_s} v_j$ and $u_{r+1} = 1$. Define $g(x) = u_s - 1 \text{for} x \in \cup_{i \in G_s} [(i-1)/n, i/n)$.
2. $g(x) \geq f(x) \forall x$. So enough to bound $\int g(x)$.
3. Intuition for next assertion: $\int_0^{i_0/n} \underbrace{g(x)}_{Y} \underbrace{dx}_{dX} = \underbrace{[xg(x)]_0^{i_0/n}}_{0} - \int_0^{i_0/n} xg'(x)dx$.
4. Total Variation $= \sum_{s=1}^{r} \pi(G_1 \cup G_2 \cup \ldots G_s)(u_s - u_{s+1}).....(*)$
5. Direct Proof of (*), by expansion: $\sum_{s=1}^{r} \pi(G_1 \cup G_2 \cup \ldots G_s)(u_s - u_{s+1}) = \pi(G_1)(u_1 - u_2) + (\pi(G_1) + \pi(G_2))(u_2 - u_3) + \cdots + (\pi(G_1) + \pi(G_2) + \cdots + \pi(G_r))(u_r - 1) = \pi(G_1)u_1 + \pi(G_2)u_2 + \pi(G_3)u_3 + \cdots + \pi(G_r)u_r - (\pi(G_1) + \pi(G_2) + \cdots + \pi(G_r)). = \sum_{s=1}^{r} \pi(G_s)(u_s - 1)$.
6. Integrals only for intuition on “integration by parts”.

Groups and Prob Flows -II

1. Now, more rigorously: Assume we have the groups. Define for $s = 1, 2, \ldots, r$: $u_s = \text{Max}_{j \in G_s} v_j$ and $u_{r+1} = 1$. Define $g(x) = u_s - 1 for x \in \cup_{i \in G_s} [(i-1)/n, i/n)$.
2. $g(x) \geq f(x) \forall x$. So enough to bound $\int g(x)$.
3. Intuition for next assertion: $\int_0^{i_0/n} \underbrace{g(x)}_{Y} \underbrace{dx}_{dX} = \underbrace{[xg(x)]_0^{i_0/n}}_{0} - \int_0^{i_0/n} xg'(x)dx$.
4. Total Variation $= \sum_{s=1}^{r} \pi(G_1 \cup G_2 \cup \ldots G_s)(u_s - u_{s+1})$.....(*)
5. Direct Proof of (*), by expansion: $\sum_{s=1}^{r} \pi(G_1 \cup G_2 \cup \ldots G_s)(u_s - u_{s+1}) = \pi(G_1)(u_1 - u_2) + (\pi(G_1) + \pi(G_2))(u_2 - u_3) + \cdots + (\pi(G_1) + \pi(G_2) + \cdots + \pi(G_r))(u_r - 1) = \pi(G_1)u_1 + \pi(G_2)u_2 + \pi(G_3)u_3 + \cdots + \pi(G_r)u_r - (\pi(G_1) + \pi(G_2) + \cdots + \pi(G_r)). = \sum_{s=1}^{r} \pi(G_s)(u_s - 1)$.
6. Integrals only for intuition on “integration by parts”.
7. Now, we have TV in terms of the “heaviness gap”: $u_s - u_{s+1}$.

Probability Flow Lemma

- Definition of groups: Intuitively: If $G_1, G_2, \ldots, G_{s-1}$ are already defined, G_s has the next $\frac{\varepsilon\Phi}{4}(|G_1 \cup G_2 \cup \ldots G_{s-1}|)$ vertices. (Sizes of $|G_1 \cup G_2 \cup \ldots G_s|$ grow as $(1 + (\varepsilon\Phi/4)^s$.) More precisely,

Probability Flow Lemma

- Definition of groups: Intuitively: If $G_1, G_2, \ldots, G_{s-1}$ are already defined, G_s has the next $\frac{\varepsilon\Phi}{4}(|G_1 \cup G_2 \cup \ldots G_{s-1}|)$ vertices. (Sizes of $|G_1 \cup G_2 \cup \ldots G_s|$ grow as $(1 + (\varepsilon\Phi/4)^s.)$ More precisely,
  - Start with $G_1 = \{1\}$.

Probability Flow Lemma

- Definition of groups: Intuitively: If $G_1, G_2, \ldots, G_{s-1}$ are already defined, G_s has the next $\frac{\varepsilon\Phi}{4}(|G_1 \cup G_2 \cup \ldots G_{s-1}|)$ vertices. (Sizes of $|G_1 \cup G_2 \cup \ldots G_s|$ grow as $(1 + (\varepsilon\Phi/4)^s$.) More precisely,
  - Start with $G_1 = \{1\}$.
  - If $G_1, G_2, \ldots, G_{s-1}$ have been defined, let $i_s = 1+$(end of G_{s-1}).

Probability Flow Lemma

- Definition of groups: Intuitively: If $G_1, G_2, \ldots, G_{s-1}$ are already defined, G_s has the next $\frac{\varepsilon\Phi}{4}(|G_1 \cup G_2 \cup \ldots G_{s-1}|)$ vertices. (Sizes of $|G_1 \cup G_2 \cup \ldots G_s|$ grow as $(1 + (\varepsilon\Phi/4)^s$.) More precisely,
  - Start with $G_1 = \{1\}$.
  - If $G_1, G_2, \ldots, G_{s-1}$ have been defined, let $i_s = 1+$(end of G_{s-1}).
  - Define l, the last element of G_s to be the largest integer in $\{i_s, i_s, \ldots, i_0\}$ such that $l - i_s + 1 \leq \frac{\varepsilon\Phi(i_s-1)}{4}$.

Probability Flow Lemma

- Definition of groups: Intuitively: If $G_1, G_2, \ldots, G_{s-1}$ are already defined, G_s has the next $\frac{\varepsilon\Phi}{4}(|G_1 \cup G_2 \cup \ldots G_{s-1}|)$ vertices. (Sizes of $|G_1 \cup G_2 \cup \ldots G_s|$ grow as $(1 + (\varepsilon\Phi/4)^s.)$ More precisely,
  - Start with $G_1 = \{1\}$.
  - If $G_1, G_2, \ldots, G_{s-1}$ have been defined, let $i_s = 1+$(end of G_{s-1}).
  - Define l, the last element of G_s to be the largest integer in $\{i_s, i_s, \ldots, i_0\}$ such that $l - i_s + 1 \leq \frac{\varepsilon\Phi(i_s-1)}{4}$.
- **Probability Flow Lemma**: Suppose groups $G_1, G_2, \ldots, G_r, u_1.u_2, \ldots, u_r, u_{r+1}$ are as above. Then,

$$\pi(G_1 \cup G_2 \cup \ldots G_s)(u_s - u_{s+1}) \leq \frac{8}{t\Phi\varepsilon}.$$

Proof of Prob flow lemma

- This is the main lemma. Use two ways of calculating prob flow from heavy to light states when we execute one step of MC starting at $\mathbf{a}$. Prob's after that step is $\mathbf{a}P$. $\mathbf{a} - \mathbf{a}P$ is net loss of prob for each state.

Proof of Prob flow lemma

- This is the main lemma. Use two ways of calculating prob flow from heavy to light states when we execute one step of MC starting at $\mathbf{a}$. Prob's after that step is $\mathbf{a}P$. $\mathbf{a} - \mathbf{a}P$ is net loss of prob for each state.
- Suppose group $G_s = \{k, k+1, \ldots, l\}$. First consider case: $k < i_0$. Let $A = \{1, 2, \ldots, k\}$. Net loss of prob flow from A is $\sum_{i=1}^{k}(a_i - (\mathbf{a}P)_i)$ which $\leq \frac{2}{t}$ by the proof of Fund Thm.

Proof of Prob flow lemma

- This is the main lemma. Use two ways of calculating prob flow from heavy to light states when we execute one step of MC starting at $\mathbf{a}$. Prob's after that step is $\mathbf{a}P$. $\mathbf{a} - \mathbf{a}P$ is net loss of prob for each state.
- Suppose group $G_s = \{k, k+1, \ldots, l\}$. First consider case: $k < i_0$. Let $A = \{1, 2, \ldots, k\}$. Net loss of prob flow from A is $\sum_{i=1}^{k}(a_i - (\mathbf{a}P)_i)$ which $\leq \frac{2}{t}$ by the proof of Fund Thm.
- But also, net prob flow from A is (Net flow from A to G_s) + (Net flow from A to $\{l+1, l+2, \ldots, n\}$).

Proof of Prob flow lemma

- This is the main lemma. Use two ways of calculating prob flow from heavy to light states when we execute one step of MC starting at $\mathbf{a}$. Prob's after that step is $\mathbf{a}P$. $\mathbf{a} - \mathbf{a}P$ is net loss of prob for each state.
- Suppose group $G_s = \{k, k+1, \ldots, l\}$. First consider case: $k < i_0$. Let $A = \{1, 2, \ldots, k\}$. Net loss of prob flow from A is $\sum_{i=1}^{k}(a_i - (\mathbf{a}P)_i)$ which $\leq \frac{2}{t}$ by the proof of Fund Thm.
- But also, net prob flow from A is (Net flow from A to G_s) + (Net flow from A to $\{l+1, l+2, \ldots, n\}$).
- Saw: Net flow from i to j is non-negative when i is heavier than j, so Net flow from A to G_s is non-negative.

Proof of Prob flow lemma

- This is the main lemma. Use two ways of calculating prob flow from heavy to light states when we execute one step of MC starting at $\mathbf{a}$. Prob's after that step is $\mathbf{a}P$. $\mathbf{a} - \mathbf{a}P$ is net loss of prob for each state.
- Suppose group $G_s = \{k, k+1, \ldots, l\}$. First consider case: $k < i_0$. Let $A = \{1, 2, \ldots, k\}$. Net loss of prob flow from A is $\sum_{i=1}^{k}(a_i - (\mathbf{a}P)_i)$ which $\leq \frac{2}{t}$ by the proof of Fund Thm.
- But also, net prob flow from A is (Net flow from A to G_s) + (Net flow from A to $\{l+1, l+2, \ldots, n\}$).
- Saw: Net flow from i to j is non-negative when i is heavier than j, so Net flow from A to G_s is non-negative.
- So Net flow from A $\geq$ (Net flow from A to $\{l+1, l+2, \ldots, n\}$) $\geq \sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji}(v_i - v_j) \geq (v_k - v_{l+1}) \sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji}$.

Prob Flow Lemma-contd.

- Net loss of flow $\leq 2/t$ (from proof of Fund Thm). So:

$$(v_k - v_{l+1}) \sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} \leq \frac{2}{t}. \tag{1}$$

Prob Flow Lemma-contd.

- Net loss of flow $\leq 2/t$ (from proof of Fund Thm). So:

$$(v_k - v_{l+1}) \sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} \leq \frac{2}{t}. \tag{1}$$

- $\sum_{i=1}^{k} \sum_{j=k+1}^{l} \pi_j p_{ji} \leq \sum_{j=k+1}^{l} \pi_j \leq \varepsilon \Phi \pi(A)/4.$

Prob Flow Lemma-contd.

- Net loss of flow $\leq 2/t$ (from proof of Fund Thm). So:

$$(v_k - v_{l+1}) \sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} \leq \frac{2}{t}. \tag{1}$$

- $\sum_{i=1}^{k} \sum_{j=k+1}^{l} \pi_j p_{ji} \leq \sum_{j=k+1}^{l} \pi_j \leq \varepsilon \Phi \pi(A)/4$.
- By defn of Φ: $\sum_{i \leq k < j} \pi_j p_{ji} \geq \Phi \min(\pi(A), \pi(\bar{A})) \geq \varepsilon \Phi k/(2n)$, (assuming $\pi(\bar{A}) \geq \varepsilon \pi(A)/2$ (to prove later).)

Prob Flow Lemma-contd.

- Net loss of flow $\leq 2/t$ (from proof of Fund Thm). So:

$$(v_k - v_{l+1}) \sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} \leq \frac{2}{t}. \quad (1)$$

- $\sum_{i=1}^{k} \sum_{j=k+1}^{l} \pi_j p_{ji} \leq \sum_{j=k+1}^{l} \pi_j \leq \varepsilon \Phi \pi(A)/4.$
- By defn of Φ: $\sum_{i \leq k < j} \pi_j p_{ji} \geq \Phi \min(\pi(A), \pi(\bar{A})) \geq \varepsilon \Phi k/(2n)$, (assuming $\pi(\bar{A}) \geq \varepsilon \pi(A)/2$ (to prove later).)
- $\sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} = \sum_{i \leq k < j} \pi_j p_{ji} - \sum_{i \leq k; j \leq l} \pi_j p_{ji} \geq \varepsilon \Phi k/4n.$

Prob Flow Lemma-contd.

- Net loss of flow $\leq 2/t$ (from proof of Fund Thm). So:

$$(v_k - v_{l+1}) \sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} \leq \frac{2}{t}. \tag{1}$$

- $\sum_{i=1}^{k} \sum_{j=k+1}^{l} \pi_j p_{ji} \leq \sum_{j=k+1}^{l} \pi_j \leq \varepsilon \Phi \pi(A)/4.$
- By defn of Φ: $\sum_{i \leq k < j} \pi_j p_{ji} \geq \Phi \min(\pi(A), \pi(\bar{A})) \geq \varepsilon \Phi k/(2n)$, (assuming $\pi(\bar{A}) \geq \varepsilon \pi(A)/2$ (to prove later).)
- $\sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} = \sum_{i \leq k < j} \pi_j p_{ji} - \sum_{i \leq k; j \leq l} \pi_j p_{ji} \geq \varepsilon \Phi k/4n.$
- $\frac{k}{n}(v_k - v_{l+1}) \leq \frac{8}{t\varepsilon\Phi} \equiv (\pi(G_1) + \pi(G_2) + \cdots \pi(G_s))(u_s - u_{s+1}) \leq \frac{8}{t\varepsilon\Phi}$ proving lemma provided $k < i_0$.

Prob Flow Lemma-contd.

- Net loss of flow $\leq 2/t$ (from proof of Fund Thm). So:

$$(v_k - v_{l+1}) \sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} \leq \frac{2}{t}. \quad (1)$$

- $\sum_{i=1}^{k} \sum_{j=k+1}^{l} \pi_j p_{ji} \leq \sum_{j=k+1}^{l} \pi_j \leq \varepsilon \Phi \pi(A)/4$.
- By defn of Φ: $\sum_{i \leq k < j} \pi_j p_{ji} \geq \Phi \min(\pi(A), \pi(\bar{A})) \geq \varepsilon \Phi k/(2n)$, (assuming $\pi(\bar{A}) \geq \varepsilon \pi(A)/2$ (to prove later).)
- $\sum_{\substack{i \leq k \\ j > l}} \pi_j p_{ji} = \sum_{i \leq k < j} \pi_j p_{ji} - \sum_{i \leq k; j \leq l} \pi_j p_{ji} \geq \varepsilon \Phi k/4n$.
- $\frac{k}{n}(v_k - v_{l+1}) \leq \frac{8}{t\varepsilon\Phi} \equiv (\pi(G_1) + \pi(G_2) + \cdots \pi(G_s))(u_s - u_{s+1}) \leq \frac{8}{t\varepsilon\Phi}$ proving lemma provided $k < i_0$.
- If $k = i_0$, the proof is similar but simpler.

Bounding number of groups

- If $\sum_{i>i_0}(1-v_i)\pi_i \leq \varepsilon$, we would be done. Wlg assume $\sum_{i>i_0}(1-v_i)\pi_i > \varepsilon$ implies $\pi(\bar{A}) \geq \varepsilon$. [Needed in last slide.]

Bounding number of groups

- If $\sum_{i>i_0}(1-v_i)\pi_i \leq \varepsilon$, we would be done. Wlg assume $\sum_{i>i_0}(1-v_i)\pi_i > \varepsilon$ implies $\pi(\bar{A}) \geq \varepsilon$. [Needed in last slide.]
- Proved: $(\pi(G_1)+\pi(G_2)+\cdots\pi(G_s))(u_s-u_{s+1}) \leq \frac{8}{t\varepsilon\Phi}$, but TV is sum over all s. Need now to bound the number of groups.

Bounding number of groups

- If $\sum_{i>i_0}(1-v_i)\pi_i \le \varepsilon$, we would be done. Wlg assume $\sum_{i>i_0}(1-v_i)\pi_i > \varepsilon$ implies $\pi(\bar{A}) \ge \varepsilon$. [Needed in last slide.]
- Proved: $(\pi(G_1)+\pi(G_2)+\cdots\pi(G_s))(u_s-u_{s+1}) \le \frac{8}{t\varepsilon\Phi}$, but TV is sum over all s. Need now to bound the number of groups.
- Since group sizes grow geometrically, get number r of groups is at most: $r \le \ln_{1+(\varepsilon\Phi/2)}(1/\pi_1)+2 \le \ln(1/\pi_1)/(\varepsilon\Phi/2)+2$. $\pi_1 = 1/n$, so get an extra log factor.