## L07 Price of Anarchy

### CS 280 Algorithmic Game Theory Ioannis Panageas

Suppose 100 drivers commute from A to B. Drivers want to minimize the time.





Suppose 100 drivers commute from A to B. Drivers want to minimize the time.

Question: What if we add a new link?





Adding a fast link is not always a good idea!



Adding a fast link is not always a good idea!

 $PoA = \frac{\text{performance of worst case NE}}{\text{optimal performance if agents do not decide on their own}}$ Price of Anarchy (Koutsoupias, Papadimitriou 99').

Intro to AGT

#### Example: Simpler example. Pigou network.







#### Example: Simpler example. **Pigou network.**



A non-atomic selfish routing game is defined by:

- Graph G(V, E).
- Source destination pairs  $(s_1, t_1), ..., (s_k, t_k)$ .
- $r_i$  traffic from  $s_i \to t_i$ .
- $c_e(.) \ge 0$  cost function of edge e, continuous and non-decreasing.
- Flow is an equilibrium if all traffic is routed on cheapest paths.



- $r_i$  traffic from  $s_i \to t_i$ .
- $c_e(.) \ge 0$  cost function of edge e, continuous and non-decreasing.
- Flow is an equilibrium if all traffic is routed on cheapest paths.



- $c_e(.) \ge 0$  cost function of edge e, continuous and non-decreasing.
- Flow is an equilibrium if all traffic is routed on cheapest paths.

A bad Example. **Pigou network with large degree** *d*.



#### A bad Example. **Pigou network with large degree** *d*.



#### Questions:

- 1. When is PoA small (bounded)?
- 2. Can we find bounds on PoA for specific classes of cost functions?

**Definition** (Linear costs). *Linear costs are of the form*  $c_e(x) = a_e \cdot x + b_e$ .

**Definition** (Linear costs). *Linear costs are of the form*  $c_e(x) = a_e \cdot x + b_e$ .

**Theorem** (Roughgarden-Tardos 00', PoA for linear costs). *For every network with linear costs*:

$$cost of Nash flow \leq \frac{4}{3} \cdot cost of optimal flow.$$

**Definition** (Linear costs). *Linear costs are of the form*  $c_e(x) = a_e \cdot x + b_e$ .

**Theorem** (Roughgarden-Tardos 00', PoA for linear costs). *For every network with linear costs*:

*cost of Nash flow* 
$$\leq \frac{4}{3} \cdot cost$$
 *of optimal flow.*

*Proof.* Let  $f^*$  be a Nash flow and f another flow. We first show (Variational Inequality)

$$\sum_{e} f_e^* c_e(f_e^*) \le \sum_{e} f_e c_e(f_e^*).$$

**Definition** (Linear costs). *Linear costs are of the form*  $c_e(x) = a_e \cdot x + b_e$ .

**Theorem** (Roughgarden-Tardos 00', PoA for linear costs). *For every network with linear costs*:

*cost of Nash flow* 
$$\leq \frac{4}{3} \cdot cost$$
 *of optimal flow.*

*Proof.* Let  $f^*$  be a Nash flow and f another flow. We first show (Variational Inequality)

$$\sum_{e} f_e^* c_e(f_e^*) \le \sum_{e} f_e c_e(f_e^*).$$

Observe that

 $f^*$  equilibrium flow  $\Rightarrow$  if  $f_p^* > 0$  then  $c_p(f^*) \le c_{p'}(f^*)$  for all paths p'.

Proof cont. Therefore all paths p so that  $f_p^* > 0$  have same cost say L. Hence  $\sum_p f_p^* c_p(f^*) = L \cdot F$  where  $F = \sum_p f_p^*$  is the total flow.

Proof cont. Therefore all paths p so that  $f_p^* > 0$  have same cost say L. Hence  $\sum_p f_p^* c_p(f^*) = L \cdot F$  where  $F = \sum_p f_p^*$  is the total flow.

Since  $c_p(f^*) \ge L$  we conclude

$$\sum_{p} f_p c_p(f^*) \ge L \sum_{p} f_p = L \cdot F$$

Proof cont. Therefore all paths p so that  $f_p^* > 0$  have same cost say L. Hence  $\sum_p f_p^* c_p(f^*) = L \cdot F$  where  $F = \sum_p f_p^*$  is the total flow.

Since  $c_p(f^*) \ge L$  we conclude

$$\sum_{p} f_p c_p(f^*) \ge L \sum_{p} f_p = L \cdot F$$

Combining the above

$$\sum_{e} f_e c_e(f^*) = \sum_{p} f_p c_p(f^*) \ge L \cdot F = \sum_{p} f_p^* c_p(f^*) = \sum_{e} f_e^* c_e(f^*)$$

$$\sum_e f_e c_e(f^*) \ge \sum_e f_e^* c_e(f^*).$$

*Proof cont.* We get that

$$\sum_{e} f_{e}^{*} c_{e}(f^{*}) \leq \sum_{e} f_{e} c_{e}(f) + \sum_{e} f_{e}(c_{e}(f^{*}) - c_{e}(f))$$

$$\sum_e f_e c_e(f^*) \ge \sum_e f_e^* c_e(f^*).$$

*Proof cont.* We get that

$$\sum_{e} f_{e}^{*} c_{e}(f^{*}) \leq \sum_{e} f_{e} c_{e}(f) + \sum_{e} f_{e}(c_{e}(f^{*}) - c_{e}(f))$$

We also have that

$$\sum_{e} f_e(c_e(f^*) - c_e(f)) \le \frac{1}{4} \sum_{e} f_e^* c_e(f^*)$$

$$\sum_e f_e c_e(f^*) \ge \sum_e f_e^* c_e(f^*).$$

*Proof cont.* We get that

$$\sum_{e} f_{e}^{*} c_{e}(f^{*}) \leq \sum_{e} f_{e} c_{e}(f) + \sum_{e} f_{e}(c_{e}(f^{*}) - c_{e}(f))$$

We also have that

$$\sum_{e} f_e(c_e(f^*) - c_e(f)) \le \frac{1}{4} \sum_{e} f_e^* c_e(f^*)$$

• Case  $c_e(f^*) < c_e(f)$  trivially  $f_e(c_e(f^*) - c_e(f)) \le \frac{1}{4} f_e^* c_e(f^*)$ .

$$\sum_e f_e c_e(f^*) \ge \sum_e f_e^* c_e(f^*).$$

*Proof cont.* We get that

$$\sum_{e} f_{e}^{*} c_{e}(f^{*}) \leq \sum_{e} f_{e} c_{e}(f) + \sum_{e} f_{e}(c_{e}(f^{*}) - c_{e}(f))$$

We also have that

$$\sum_{e} f_e(c_e(f^*) - c_e(f)) \le \frac{1}{4} \sum_{e} f_e^* c_e(f^*)$$

- Case  $c_e(f^*) < c_e(f)$  trivially  $f_e(c_e(f^*) c_e(f)) \le \frac{1}{4} f_e^* c_e(f^*)$ .
- Case  $c_e(f^*) \ge c_e(f) \Rightarrow f_e^* \ge f_e$ . Linear costs  $\Rightarrow$  LHS =  $a_e f_e(f_e^* f_e)$  and RHS  $\ge \frac{1}{4}a_e f_e^{*2}$ .

$$\sum_e f_e c_e(f^*) \ge \sum_e f_e^* c_e(f^*).$$

*Proof cont.* We get that

$$\sum_{e} f_{e}^{*} c_{e}(f^{*}) \leq \sum_{e} f_{e} c_{e}(f) + \sum_{e} f_{e}(c_{e}(f^{*}) - c_{e}(f))$$

We also have that

$$\sum_{e} f_e(c_e(f^*) - c_e(f)) \le \frac{1}{4} \sum_{e} f_e^* c_e(f^*)$$

• Case 
$$c_e(f^*) < c_e$$
 Since  $xy - y^2 \le \frac{x^2}{4} \Rightarrow \text{LHS} \le \text{RHS}$ .  $c_e(f^*)$ .

• Case  $c_e(f^*) \ge c_e(f) \ne f_e \le f_e$ . Linear costs  $\Rightarrow$  Line  $a_e f_e(f_e^* - f_e)$  and RHS  $\ge \frac{1}{4}a_e f_e^{*2}$ .

$$\sum_e f_e c_e(f^*) \ge \sum_e f_e^* c_e(f^*).$$

*Proof cont.* We conclude that

$$\sum_{e} f_{e}^{*} c_{e}(f^{*}) \leq \sum_{e} f_{e} c_{e}(f) + \frac{1}{4} \sum_{e} f_{e}^{*} c_{e}(f^{*})$$

$$\sum_e f_e c_e(f^*) \ge \sum_e f_e^* c_e(f^*).$$

*Proof cont.* We conclude that

$$\sum_{e} f_{e}^{*} c_{e}(f^{*}) \leq \sum_{e} f_{e} c_{e}(f) + \frac{1}{4} \sum_{e} f_{e}^{*} c_{e}(f^{*}).$$

Or equivalently

$$\sum_{e} f_e^* c_e(f^*) \le \frac{4}{3} \sum_{e} f_e c_e(f).$$



Or equivalently

$$\sum_{e} f_e^* c_e(f^*) \le \frac{4}{3} \sum_{e} f_e c_e(f).$$

**Theorem** (Roughgarden 02', PoA for polynomial costs). *For every network with polynomial costs with degree d:* 

$$cost of Nash flow \le \Theta\left(\frac{d}{\log d}\right) \cdot cost of optimal flow.$$



Or equivalently

$$\sum_{e} f_{e}^{*} c_{e}(f^{*}) \leq \frac{4}{3} \sum_{e} f_{e} c_{e}(f).$$

**Theorem** (Roughgarden 02', PoA for polynomial costs). For every network with polynomial costs with degree d:

$$cost of Nash flow \le \Theta\left(\frac{d}{\log d}\right) \cdot cost of optimal flow$$



**Theorem** (Christodoulou-Koutsoupias, PoA for linear costs). *For every congestion game with linear costs:* 

*cost of worst Nash* 
$$\leq \frac{5}{2} \cdot cost$$
 *of optimal welfare.*

**Theorem** (Christodoulou-Koutsoupias, PoA for linear costs). *For every congestion game with linear costs:* 

*cost of worst Nash* 
$$\leq \frac{5}{2} \cdot cost$$
 *of optimal welfare.*

*Proof.* Let  $l^*$  be a Nash equilibrium in which i uses path  $P_i$  and assume i deviates to path  $\tilde{P}_i$ . It holds (Variational Inequality)

$$\sum_{e \in P_i} c_e(l_e^*) \le \sum_{e \in P_i \cap \tilde{P}_i} c_e(l_e^*) + \sum_{e \in \tilde{P}_i \setminus P_i} c_e(l_e^* + 1)$$

**Theorem** (Christodoulou-Koutsoupias, PoA for linear costs). *For every congestion game with linear costs:* 

*cost of worst Nash* 
$$\leq \frac{5}{2} \cdot cost$$
 *of optimal welfare.*

*Proof.* Let  $l^*$  be a Nash equilibrium in which i uses path  $P_i$  and assume i deviates to path  $\tilde{P}_i$ . It holds (Variational Inequality)

$$\sum_{e \in P_i} c_e(l_e^*) \leq \sum_{e \in P_i \cap \tilde{P}_i} c_e(l_e^*) + \sum_{e \in \tilde{P}_i \setminus P_i} c_e(l_e^* + 1)$$
$$\leq \sum_{e \in P_i \cap \tilde{P}_i} c_e(l_e^* + 1) + \sum_{e \in \tilde{P}_i \setminus P_i} c_e(l_e^* + 1)$$

**Theorem** (Christodoulou-Koutsoupias, PoA for linear costs). *For every congestion game with linear costs:* 

*cost of worst Nash* 
$$\leq \frac{5}{2} \cdot cost$$
 *of optimal welfare.*

*Proof.* Let  $l^*$  be a Nash equilibrium in which i uses path  $P_i$  and assume i deviates to path  $\tilde{P}_i$ . It holds (Variational Inequality)

$$\sum_{e \in P_i} c_e(l_e^*) \leq \sum_{e \in P_i \cap \tilde{P}_i} c_e(l_e^*) + \sum_{e \in \tilde{P}_i \setminus P_i} c_e(l_e^* + 1)$$
$$\leq \sum_{e \in P_i \cap \tilde{P}_i} c_e(l_e^* + 1) + \sum_{e \in \tilde{P}_i \setminus P_i} c_e(l_e^* + 1)$$
$$= \sum_{e \in \tilde{P}_i} c_e(l_e^* + 1).$$

$$\sum_{e \in P_i} c_e(l_e^*) \le \sum_{e \in \tilde{P}_i} c_e(l_e^* + 1).$$

*Proof cont.* Consider any configuration  $\tilde{l}$ , where each agent j uses path  $\tilde{P}_j$ . Summing for all agents i

$$\sum_{i \in [n]} \sum_{e \in P_i} c_e(l_e^*) \le \sum_{i \in [n]} \sum_{e \in \tilde{P}_i} c_e(l_e^* + 1).$$

$$\sum_{e \in P_i} c_e(l_e^*) \le \sum_{e \in \tilde{P}_i} c_e(l_e^* + 1).$$

*Proof cont.* Consider any configuration  $\tilde{l}$ , where each agent j uses path  $\tilde{P}_j$ . Summing for all agents i

$$\sum_{i \in [n]} \sum_{e \in P_i} c_e(l_e^*) \le \sum_{i \in [n]} \sum_{e \in \tilde{P}_i} c_e(l_e^* + 1).$$
$$= \sum_{e} \tilde{l}_e c_e(l_e^* + 1).$$

$$\sum_{e \in P_i} c_e(l_e^*) \le \sum_{e \in \tilde{P}_i} c_e(l_e^* + 1).$$

*Proof cont.* Consider any configuration  $\tilde{l}$ , where each agent j uses path  $\tilde{P}_j$ . Summing for all agents i

$$\sum_{i \in [n]} \sum_{e \in P_i} c_e(l_e^*) \le \sum_{i \in [n]} \sum_{e \in \tilde{P}_i} c_e(l_e^* + 1).$$
  
=  $\sum_e \tilde{l}_e c_e(l_e^* + 1).$   
=  $\sum_e a_e \tilde{l}_e(l_e^* + 1) + b_e \tilde{l}_e.$ 

$$\sum_{e \in P_i} c_e(l_e^*) \le \sum_{e \in \tilde{P}_i} c_e(l_e^* + 1).$$

*Proof cont.* Consider any configuration  $\tilde{l}$ , where each agent j uses path  $\tilde{P}_j$ . Summing for all agents i

$$\sum_{i \in [n]} \sum_{e \in P_i} c_e(l_e^*) \leq \left| \begin{array}{c} \text{Since } y(z+1) \leq \frac{5}{3}y^2 + \frac{1}{3}z^2 \text{ for naturals } y, z \\ \\ = \underbrace{HW2} \\ \\ = \sum_e a_e \tilde{l}_e(l_e^* + 1) + b_e \tilde{l}_e. \\ \\ \\ \leq \sum_e a_e \left(\frac{5}{3}\tilde{l}_e^2 + \frac{1}{3}{l_e^*}^2\right) + b_e \tilde{l}_e. \end{array} \right|$$

Intro to AGT

$$\sum_{i \in [n]} \sum_{e \in P_i} c_e(l_e^*) \le \sum_e a_e\left(\frac{5}{3}\tilde{l}_e^2 + \frac{1}{3}l_e^{*2}\right) + b_e\tilde{l}_e$$

*Proof cont.* Observe that

$$\frac{5}{3}C(\tilde{l}) = \frac{5}{3}\sum_{i\in[n]}\sum_{e\in\tilde{P}_i}c_e(\tilde{l}_e) = \sum_e \frac{5}{3}a_e\tilde{l}_e^2 + \frac{5}{3}b_e\tilde{l}_e \ge \sum_e \frac{5}{3}a_e\tilde{l}_e^2 + b_e\tilde{l}_e$$

$$\sum_{i \in [n]} \sum_{e \in P_i} c_e(l_e^*) \le \sum_e a_e\left(\frac{5}{3}\tilde{l}_e^2 + \frac{1}{3}{l_e^*}^2\right) + b_e\tilde{l}_e$$

*Proof cont.* Observe that

$$\frac{5}{3}C(\tilde{l}) = \frac{5}{3}\sum_{i\in[n]}\sum_{e\in\tilde{P}_i}c_e(\tilde{l}_e) = \sum_e \frac{5}{3}a_e\tilde{l}_e^2 + \frac{5}{3}b_e\tilde{l}_e \ge \sum_e \frac{5}{3}a_e\tilde{l}_e^2 + b_e\tilde{l}_e$$

Therefore

$$C(l^*) \le \frac{5}{3}C(\tilde{l}) + \frac{1}{3}\sum_e a_e l_e^{*2}$$

$$\sum_{i \in [n]} \sum_{e \in P_i} c_e(l_e^*) \le \sum_e a_e\left(\frac{5}{3}\tilde{l}_e^2 + \frac{1}{3}{l_e^*}^2\right) + b_e\tilde{l}_e$$

*Proof cont.* Observe that

$$\frac{5}{3}C(\tilde{l}) = \frac{5}{3}\sum_{i\in[n]}\sum_{e\in\tilde{P}_i}c_e(\tilde{l}_e) = \sum_e \frac{5}{3}a_e\tilde{l}_e^2 + \frac{5}{3}b_e\tilde{l}_e \ge \sum_e \frac{5}{3}a_e\tilde{l}_e^2 + b_e\tilde{l}_e$$

Therefore

$$C(l^*) \le \frac{5}{3}C(\tilde{l}) + \frac{1}{3}\sum_e a_e l_e^{*2}$$
$$\le \frac{5}{3}C(\tilde{l}) + \frac{1}{3}C(l^*)$$

$$\sum_{i \in [n]} \sum_{e \in P_i} c_e(l_e^*) \le \sum_e a_e\left(\frac{5}{3}\tilde{l}_e^2 + \frac{1}{3}{l_e^*}^2\right) + b_e\tilde{l}_e$$

$$Proof \ cont. \ Ob$$

$$\frac{5}{3}C(\tilde{l}) = C(l^*) \le \frac{5}{2}C(\tilde{l}).$$

$$a_e\tilde{l}_e^2 + b_e\tilde{l}_e$$

Therefore

$$C(l^*) \le \frac{5}{3}C(\tilde{l}) + \frac{1}{3}\sum_e a_e l_e^{*2}$$
$$\le \frac{5}{3}C(\tilde{l}) + \frac{1}{3}C(l^*)$$

Remark:

- 1. The above bound is tight!
- 2. For polynomial cost functions the PoA is exponential in *d*.

#### **Definition** (Balls and Bins). Consider

- set of *n* balls and *n* bins  $\{e_1, ..., e_n\}$ .
- *Each ball i chooses a bin j and pays the load of the bin j.*
- Define social cost the maximum load.
- What is PoA? Is it  $\frac{5}{2}$ ?

#### **Definition** (Balls and Bins). Consider

- set of *n* balls and *n* bins  $\{e_1, ..., e_n\}$ .
- *Each ball i chooses a bin j and pays the load of the bin j.*
- Define social cost the maximum load.



**Theorem** (Koutsoupias-Papadimitriou, PoA for balls & bins). *The PoA is* 

$$\Omega\left(\frac{\ln n}{\ln\ln n}\right).$$

*Proof.* We will use second moment method.

- Set every ball in a different bin. Hence optimal social cost is 1.
- Uniform  $\left(\frac{1}{n}, ..., \frac{1}{n}\right)$  is a Nash Equilibrium (symmetry).

**Theorem** (Koutsoupias-Papadimitriou, PoA for balls & bins). *The PoA is* 

$$\Omega\left(\frac{\ln n}{\ln\ln n}\right).$$

*Proof.* We will use second moment method.

- Set every ball in a different bin. Hence optimal social cost is 1.
- Uniform  $\left(\frac{1}{n}, ..., \frac{1}{n}\right)$  is a Nash Equilibrium (symmetry).
- With high probability, we show that uniform gives max load  $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ , which implies the expected max load is  $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ .

Theorem (Koutsoupias-Papadimitriou, PoA for balls & bins). The PoA is

$$\Omega\left(\frac{\ln n}{\ln\ln n}\right).$$

*Proof.* We will use second moment method.

- Set every ball in a different bin. Hence optimal social cost is 1.
- Uniform  $\left(\frac{1}{n}, ..., \frac{1}{n}\right)$  is a Nash Equilibrium (symmetry).
- With high probability, we show that uniform gives max load  $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ , which implies the expected max load is  $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ .

Claim 1: Bin *i* has at least  $k \ll n$  balls with probability at least:

$$\binom{n}{k}\frac{1}{n^k}\left(1-\frac{1}{n}\right)^{n-k} \ge \frac{1}{n^k}\left(\frac{n}{k}\right)^k\frac{1}{e} = \frac{1}{ek^k}.$$

**Theorem** (Koutsoupias-Papadimitriou, PoA for balls & bins). *The PoA is* 

$$\Omega\left(\frac{\ln n}{\ln\ln n}\right).$$

*Proof.* We will use second moment method.

Set every ball in a different bin. Hence optimal social cost is 1.
Uniform (
With high which imp
Claim 1: Bin i has at least k ≪ n balls with probability at least:

$$\binom{n}{k}\frac{1}{n^k}\left(1-\frac{1}{n}\right)^{n-k} \ge \frac{1}{n^k}\left(\frac{n}{k}\right)^k\frac{1}{e} = \frac{1}{ek^k}.$$

Proof cont. Choosing  $k = \frac{\ln n}{3 \ln \ln n}$  we have  $k^k \leq (\ln n)^k = (\ln n)^{\frac{\ln n}{3 \ln \ln n}} = n^{1/3}$ . Claim 1: Thus bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls with probability at least  $\frac{1}{en^{1/3}}$ .

Proof cont. Choosing  $k = \frac{\ln n}{3 \ln \ln n}$  we have  $k^k \le (\ln n)^k = (\ln n)^{\frac{\ln n}{3 \ln \ln n}} = n^{1/3}$ .

Claim 1: Thus bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls with probability at least  $\frac{1}{e^{n^{1/3}}}$ .

Let  $X_i$  be the indicator that bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls and X be the expected number of all bins with at least  $\frac{\ln n}{3 \ln \ln n}$  balls.

Proof cont. Choosing  $k = \frac{\ln n}{3 \ln \ln n}$  we have  $k^k \le (\ln n)^k = (\ln n)^{\frac{\ln n}{3 \ln \ln n}} = n^{1/3}$ .

Claim 1: Thus bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls with probability at least  $\frac{1}{e^{n^{1/3}}}$ .

Let  $X_i$  be the indicator that bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls and X be the expected number of all bins with at least  $\frac{\ln n}{3 \ln \ln n}$  balls.

$$X = \sum_{i} X_{i} \Rightarrow E[X] = \sum_{i} E[X_{i}].$$

Proof cont. Choosing  $k = \frac{\ln n}{3 \ln \ln n}$  we have  $k^k \le (\ln n)^k = (\ln n)^{\frac{\ln n}{3 \ln \ln n}} = n^{1/3}$ .

Claim 1: Thus bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls with probability at least  $\frac{1}{e^{n^{1/3}}}$ .

Let  $X_i$  be the indicator that bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls and X be the expected number of all bins with at least  $\frac{\ln n}{3 \ln \ln n}$  balls.

$$X = \sum_{i} X_{i} \Rightarrow E[X] = \sum_{i} E[X_{i}].$$

Observe that  $E[X] \ge \frac{n^{2/3}}{e} \gg 1$  but this does not imply  $X \ge 1$  with high probability. We need to argue about the variance (second moment).

Proof cont. Choosing  $k = \frac{\ln n}{3 \ln \ln n}$  we have  $k^k \le (\ln n)^k = (\ln n)^{\frac{\ln n}{3 \ln \ln n}} = n^{1/3}$ .

Claim 1: Thus bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls with probability at least  $\frac{1}{e^{n^{1/3}}}$ .

Let  $X_i$  be the indicator that bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls and X be the expected number of all bins with at least  $\frac{\ln n}{3 \ln \ln n}$  balls.

$$X = \sum_{i} X_{i} \Rightarrow E[X] = \sum_{i} E[X_{i}].$$

Observe that  $E[X] \ge \frac{n^{2/3}}{e} \gg 1$  but this does not imply  $X \ge 1$  with high probability. We need to argue about the variance (second moment).

Chebyshev's inequality gives  $Pr[|X - E[X]| \ge tE[X]] \le \frac{Var[X]}{t^2 E^2[X]}$ ,

Proof cont. Choosing  $k = \frac{\ln n}{3 \ln \ln n}$  we have  $k^k \leq (\ln n)^k = (\ln n)^{\frac{\ln n}{3 \ln \ln n}} = n^{1/3}$ .

Claim 1: Thus bin *i* has at least  $\frac{\ln n}{3 \ln \ln}$  balls with probability at least  $\frac{1}{en^{1/3}}$ .

Let  $X_i$  be the indicator that bin *i* has at least  $\frac{\ln n}{3 \ln \ln n}$  balls and X be the expected number of all bins with at least  $\frac{\ln n}{3 \ln \ln n}$  balls.

$$X = \sum_{i} X_{i} \Rightarrow E[X] = \sum_{i} E[X_{i}].$$

Observe that  $E[X] \ge \frac{n^{2/3}}{e} \gg 1$  but this does not imply  $X \ge 1$  with high probability. We need to argue about the variance (second moment).

Chebyshev's inequality gives  $Pr[|X - E[X]| \ge tE[X]] \le \frac{Var[X]}{t^2 E^2[X]}$ , thus  $Pr[X = 0] \le Pr[|X - E[X]| \ge E[X]] \le \frac{Var[X]}{E^2[X]}$ .

Proof cont.  $Pr[X=0] \leq \frac{Var[X]}{E^2[X]}$ .

From negative correlation we have that  $Var[X] \leq \sum_i Var[X_i]$ . . Morever  $Var[X_i] = E[X_i^2] - E^2[X_i] \leq E[X_i^2] = E[X_i] \leq 1$ 

Intro to AGT

Proof cont.  $Pr[X=0] \leq \frac{Var[X]}{E^2[X]}$ .

From negative correlation we have that  $Var[X] \leq \sum_i Var[X_i]$ . . Morever  $Var[X_i] = E[X_i^2] - E^2[X_i] \leq E[X_i^2] = E[X_i] \leq 1$ 

We conclude that

$$Pr[X=0] \le \frac{n}{e^2 n^{4/3}} = \frac{n^{-1/3}}{e^2}$$

Proof cont.  $Pr[X=0] \leq \frac{Var[X]}{E^2[X]}$ .

From negative correlation we have that  $Var[X] \leq \sum_i Var[X_i]$ . . Morever  $Var[X_i] = E[X_i^2] - E^2[X_i] \leq E[X_i^2] = E[X_i] \leq 1$ 

We conclude that

$$Pr[X=0] \le \frac{n}{e^2 n^{4/3}} = \frac{n^{-1/3}}{e^2}$$

Therefore

$$Pr[X \ge 1] = 1 - Pr[X = 0] \ge 1 - \frac{n^{-1/3}}{e^2} \to 1.$$

# **Congestion Games**

- A congestion game is defined by:
- *n* set of players.
- E set of edges/facilities/ bins.
- $S_i \subset 2^E$  the set of strategies of player *i*.
- $c_e: \{1, ..., n\} \to \mathbb{R}^+$  cost function of edge e.

For any  $s = (s_1, ..., s_n)$ 

- $l_e(s)$  number of players (load) that use edge e.
- $c_i(s) = \sum_{e \in s_i} c_e(l_e)$  the cost function of player *i*.

# **Congestion Games**



For this game:

 $n = \{1, 2\}$  (red, green) E are the edges of the network.  $S_i$  is all s - t paths.  $c_e$  on edges.

Remark: Defined by Rosenthal in 1973. Capture atomic routing games!