IMPARTIAL SELECTION, ADDITIVE APPROXIMATION GUARANTEES, AND PRIORS

Ioannis Caragiannis

iannis.dk

OVERVIEW OF THE TALK

Impartial selection: definition, examples, previous work

Additive approximation guarantees

- C., Christodoulou, & Protopapas (2019)

Using prior information

- C., Christodoulou, & Protopapas (2021)

IMPARTIAL SELECTION

IMPARTIAL SELECTION

Story:

- the members of a society wish to give their annual award to one of the members
- each member can vote (any number of) any other member(s)

Goal: give the award to the **most distinguished member**

PFA MEN'S PLAYERS' PLAYER OF THE YEAR

"the ultimate accolade to be voted for by your fellow professionals", John Terry, 2005 Awardee (BBC sport)

IMPARTIAL SELECTION

Story:

- the members of a society wish to give their annual award to one of the members
- each member can vote (any number of) any other member(s)

Goal: give the award to the **most distinguished member**

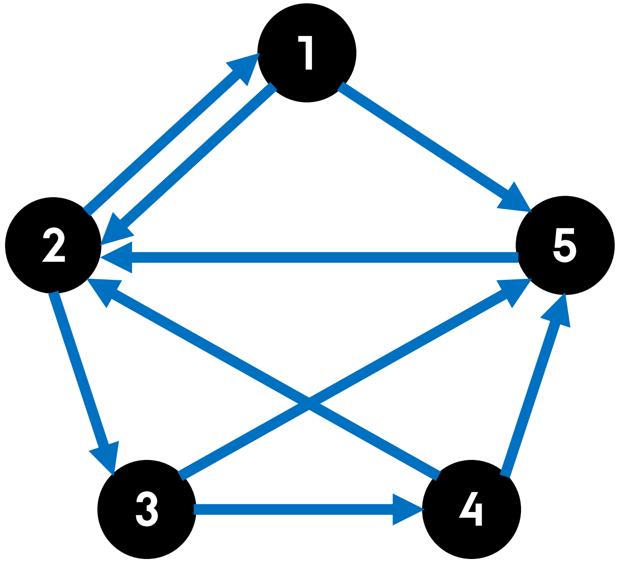
Other examples: selecting the chair of a committee, scientific grants/awards, Papal conclave, many more

Major requirement: **impartiality**

- Agents should not be able to increase **their chance of being selected** by acting strategically

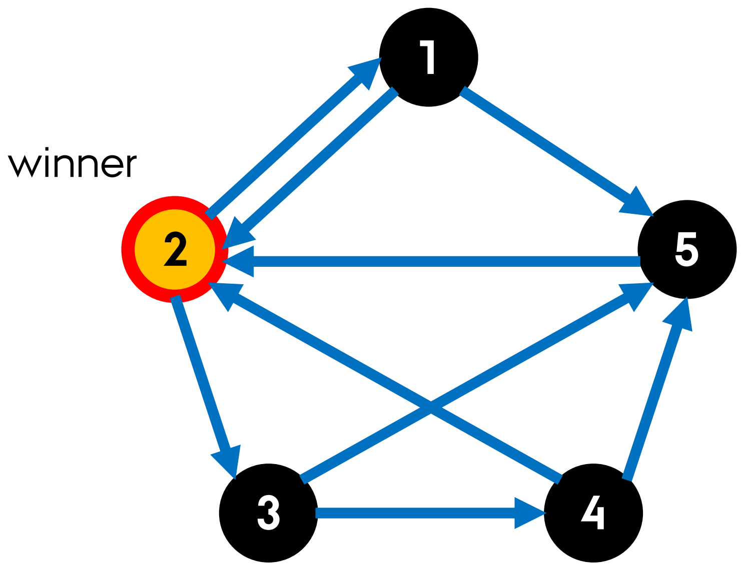
AN EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



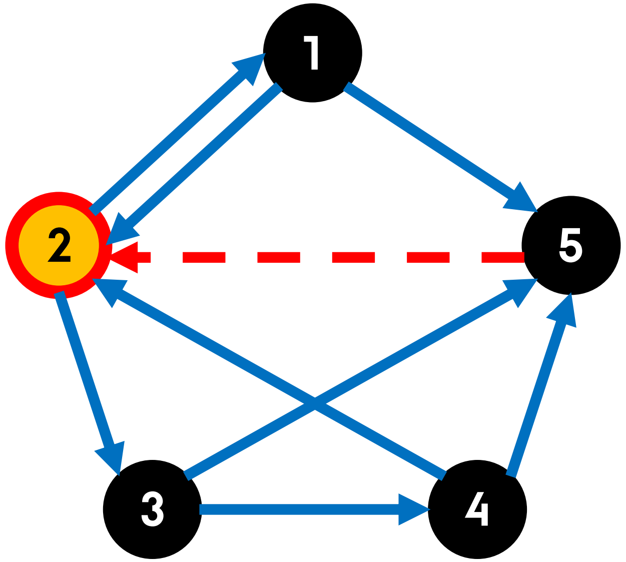
AN EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



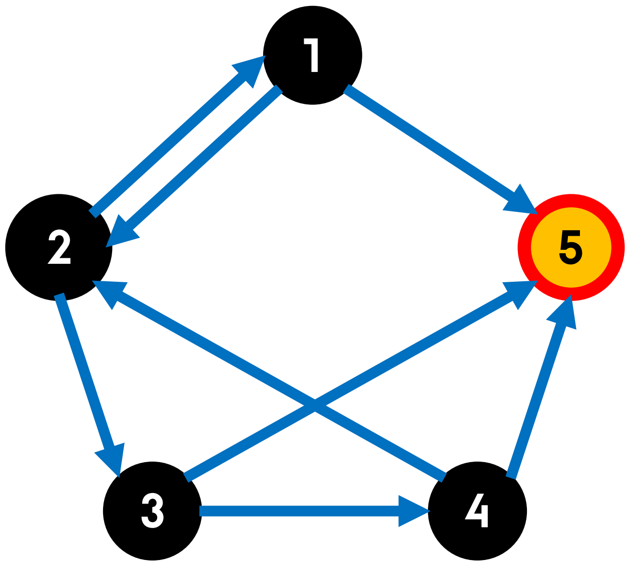
AN EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



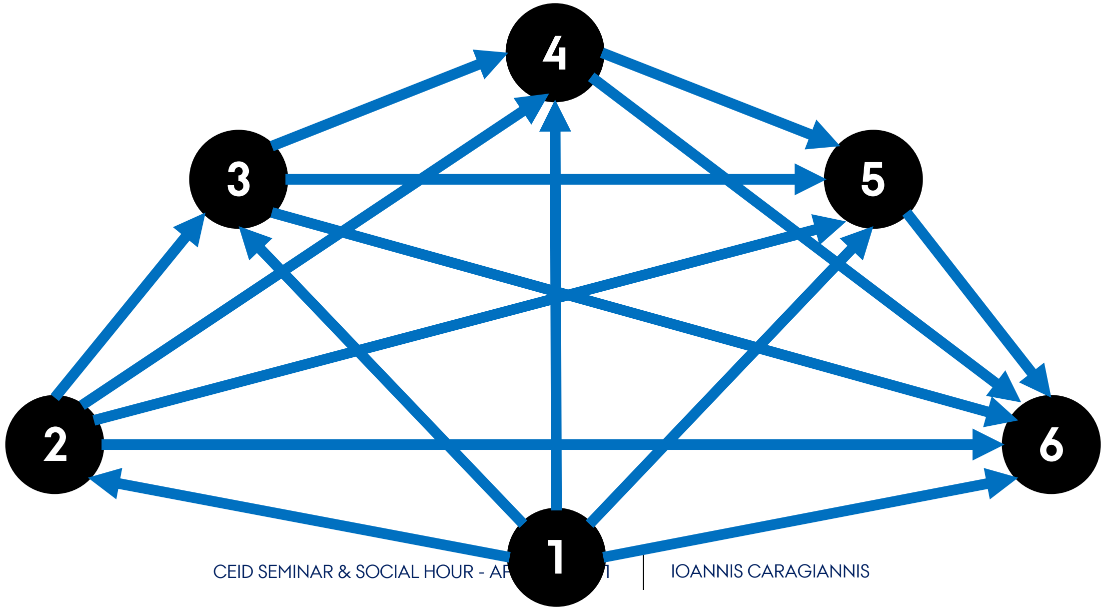
AN EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



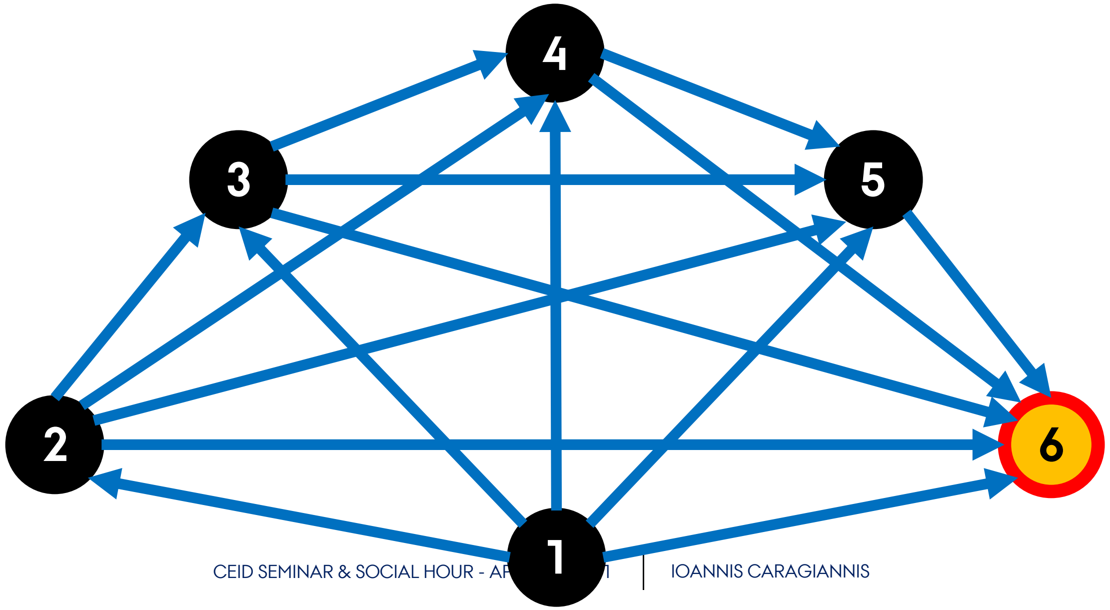
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



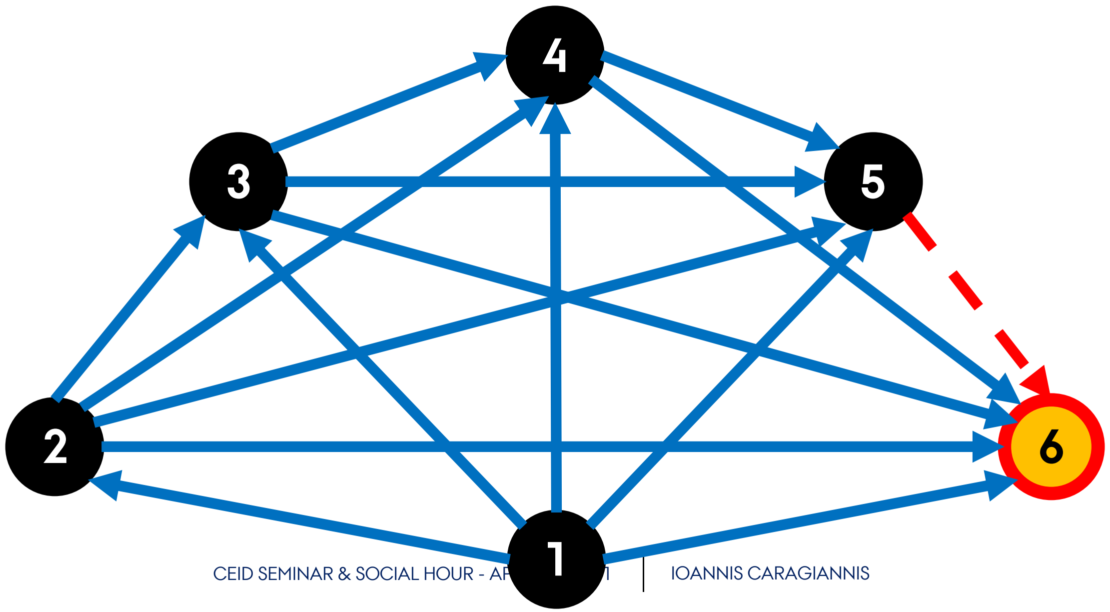
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



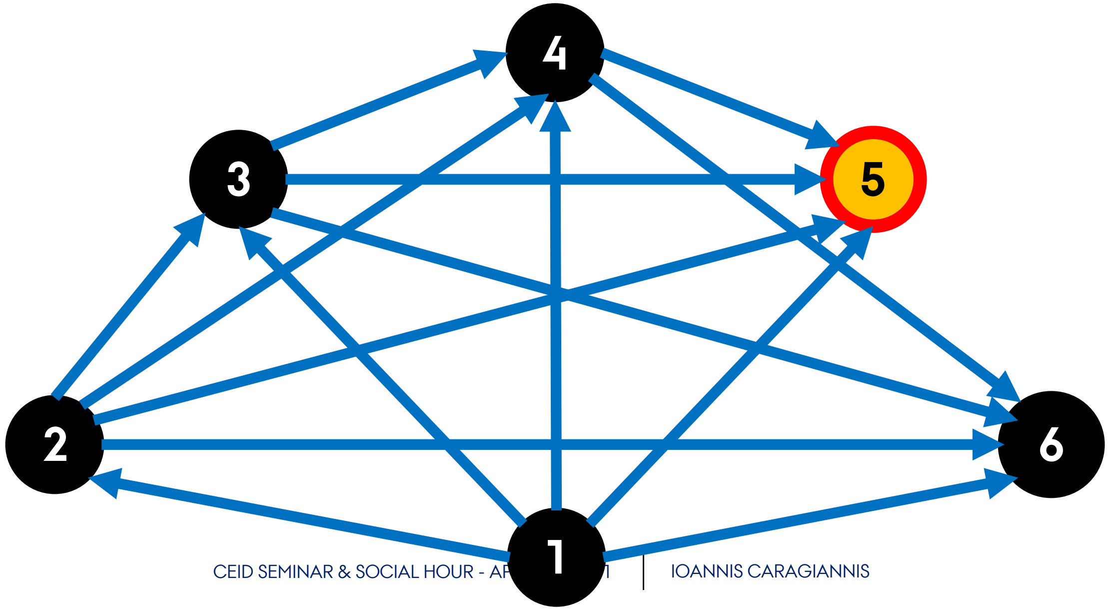
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



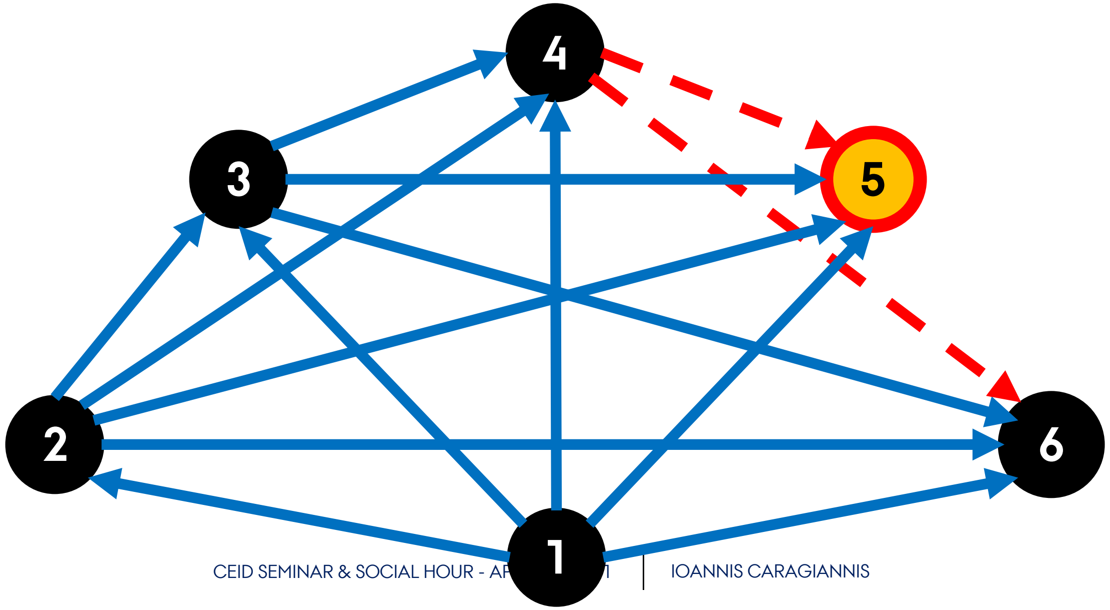
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



ANOTHER EXAMPLE

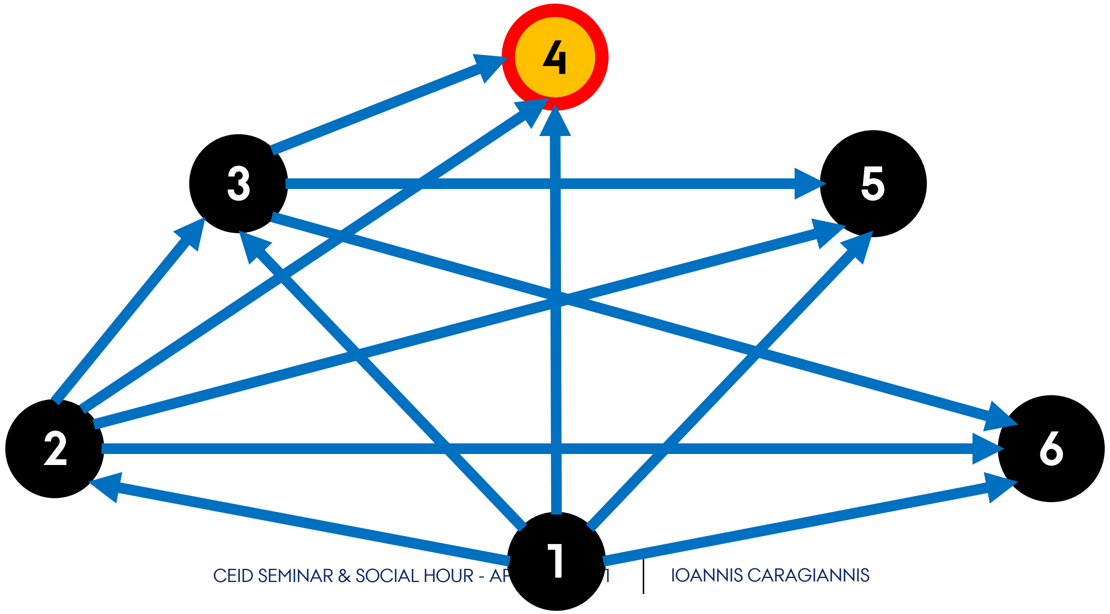
- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



AARHUS
UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

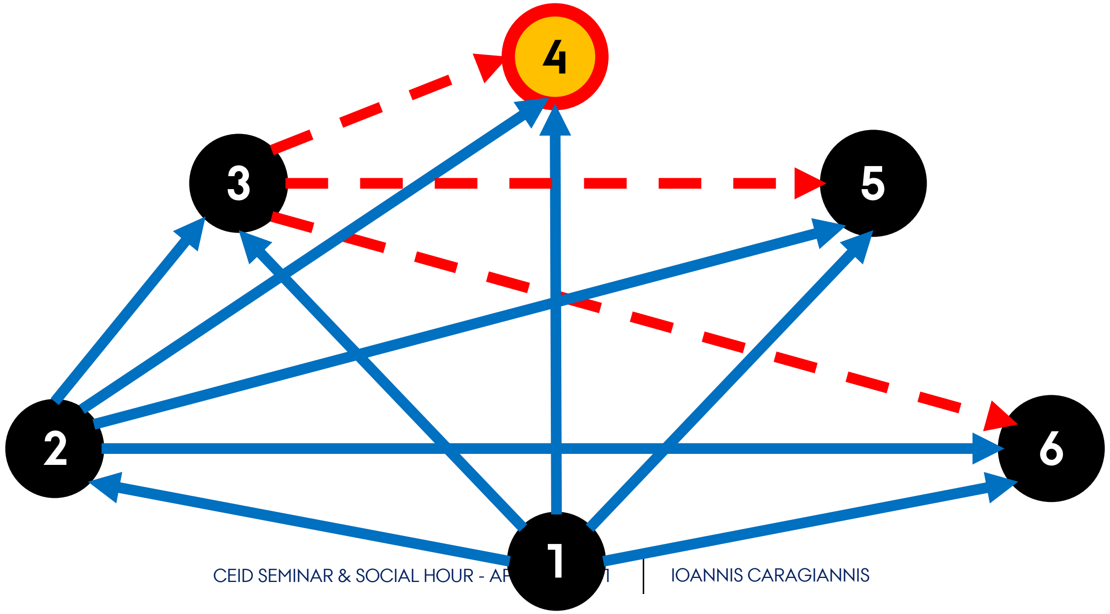
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



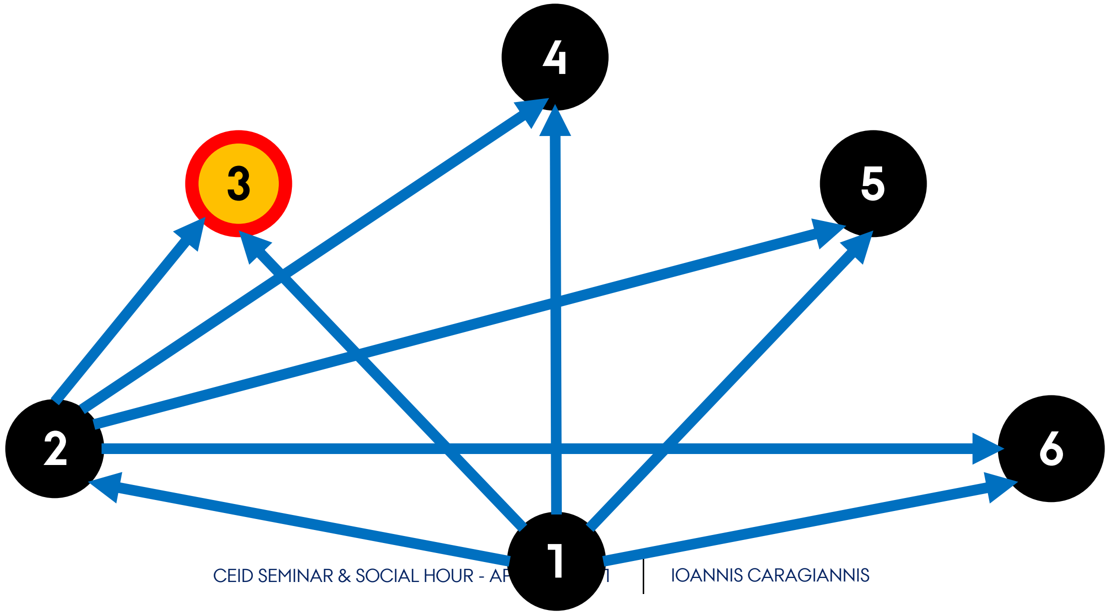
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



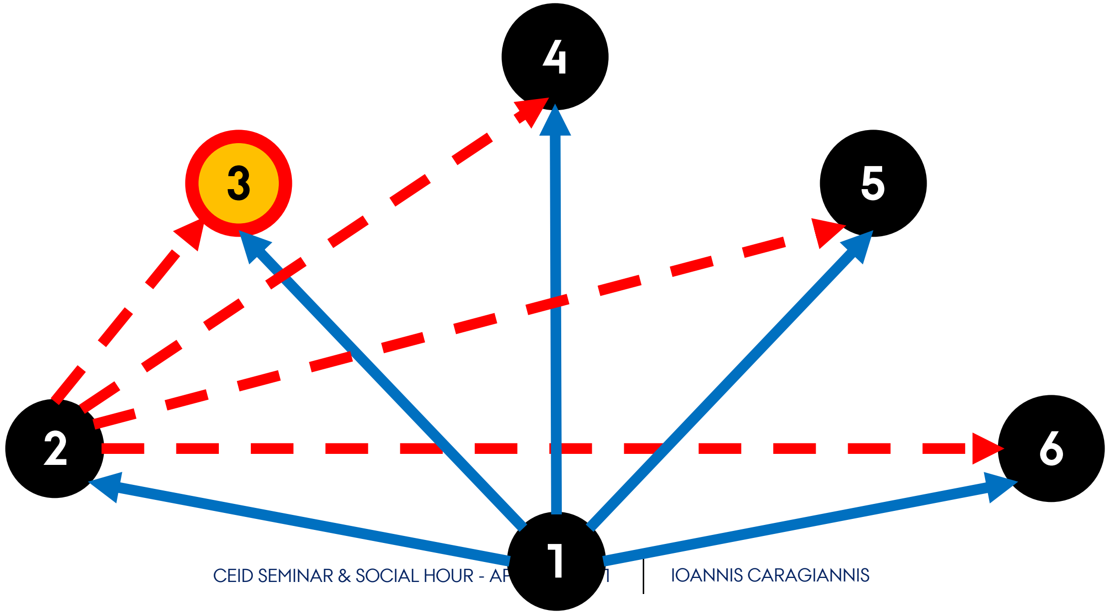
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



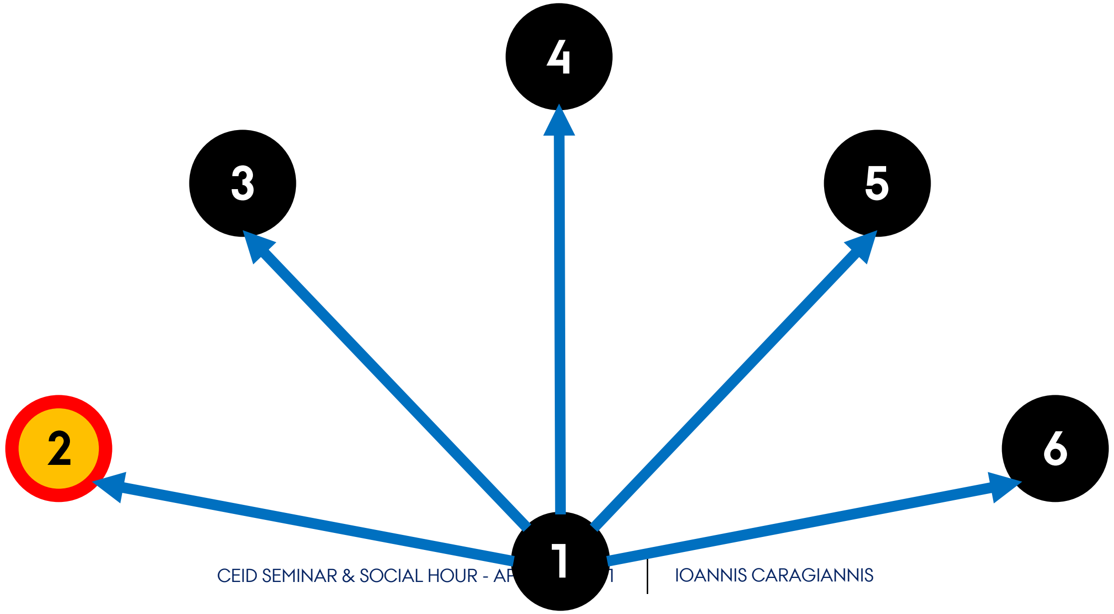
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



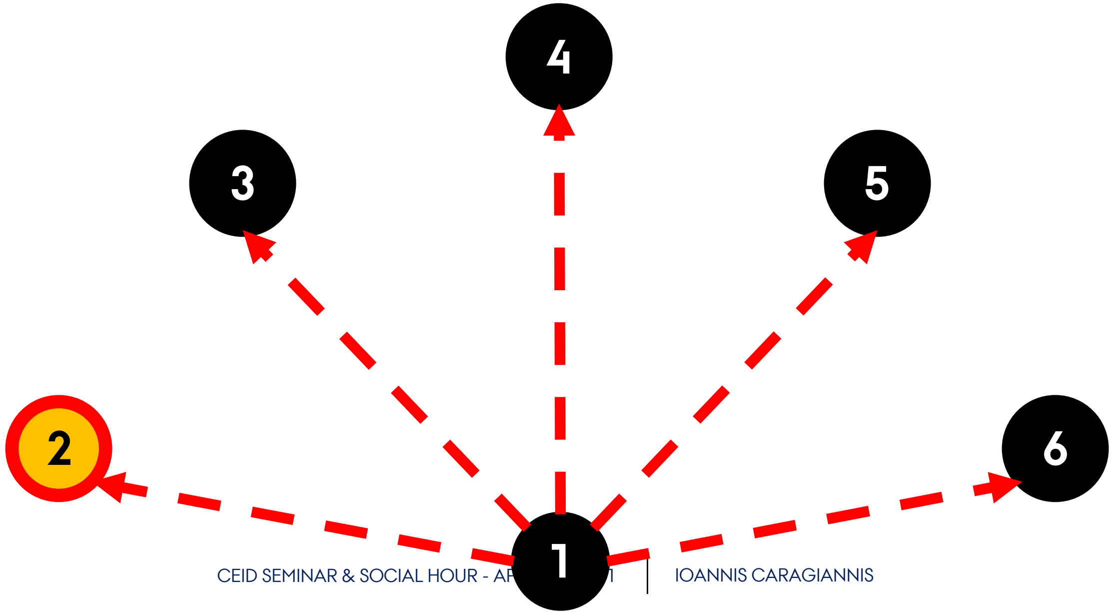
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



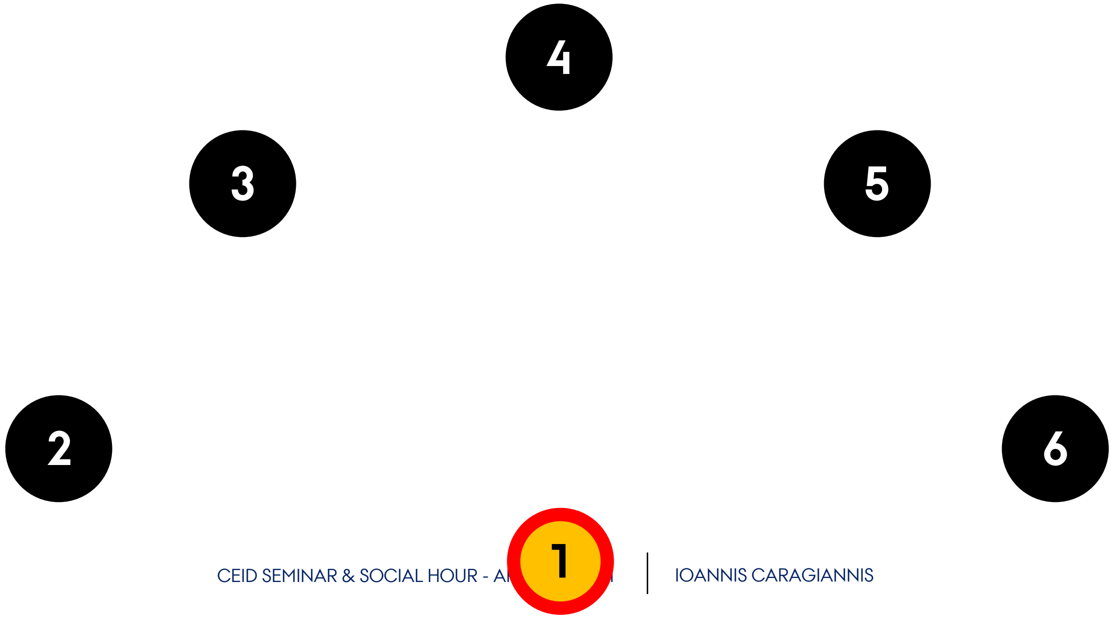
ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



ANOTHER EXAMPLE

- The **highest-degree** node wins
- In case of **ties**, lowest id wins
- Each node wants to win



IMPARTIAL SELECTION

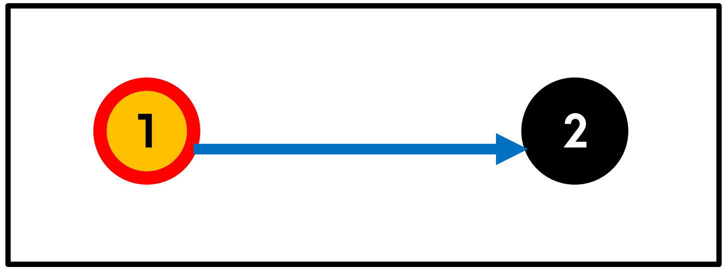
- Goal: design an **impartial selection algorithm** with **good approximation ratio**

IMPARTIAL SELECTION

- Goal: design an **impartial selection algorithm** with **good approximation ratio**
- **Randomization** is important here!

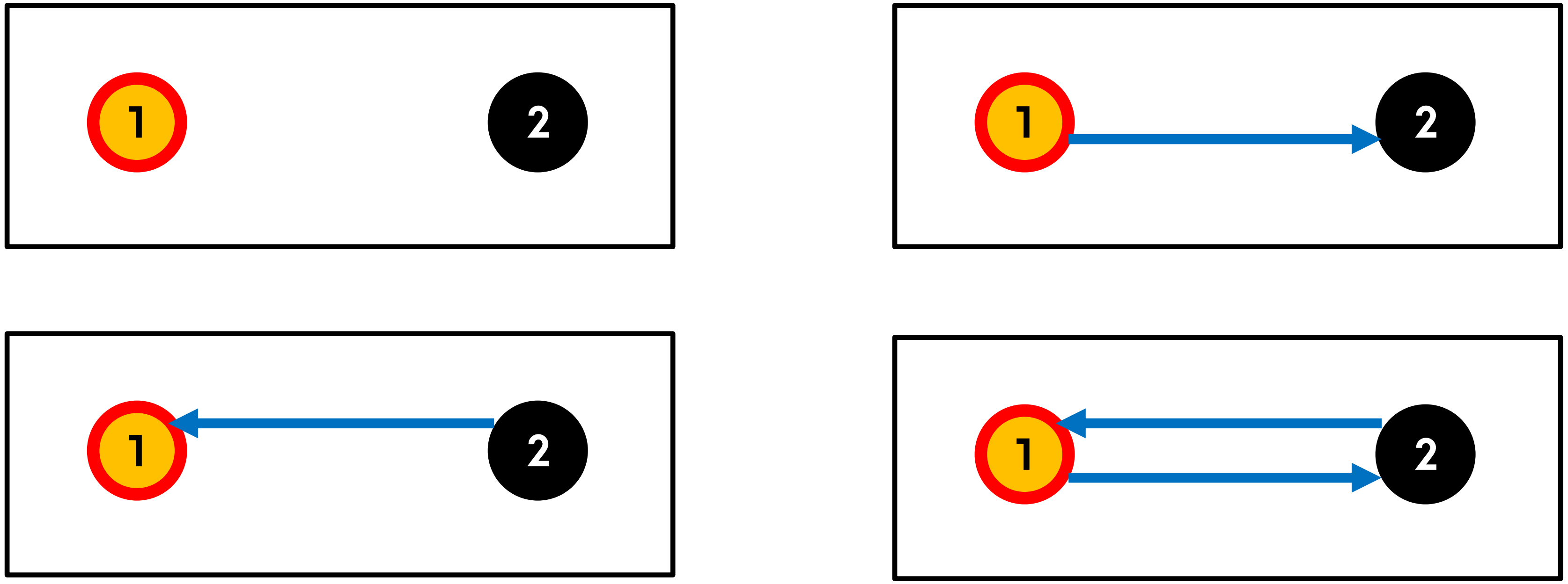
IMPARTIAL SELECTION

- Goal: design an **impartial selection algorithm** with **good approximation ratio**
- **Randomization** is important here!



IMPARTIAL SELECTION

- Goal: design an **impartial selection algorithm** with **good approximation ratio**
- **Randomization** is important here!



A 4-APPROXIMATION IMPARTIAL MECHANISM

Alon, Fischer, Procaccia, & Tennenholz (2011)

Input: a directed graph

1. Randomly partition the nodes into two sets S and W
2. The node of set W with the **highest number** of **incoming edges from set S** wins

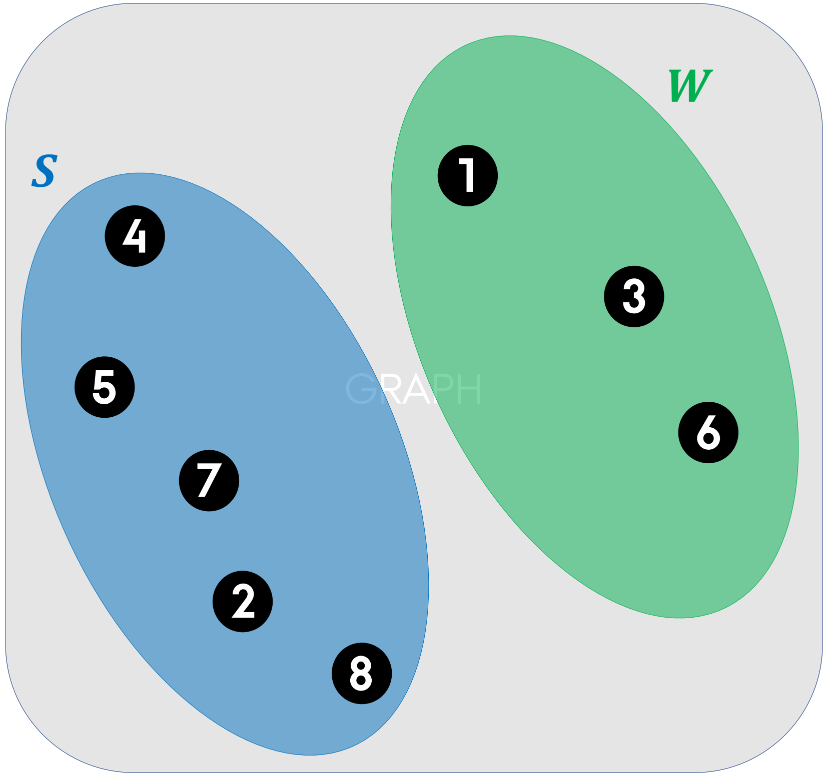
GRAPH

A 4-APPROXIMATION IMPARTIAL MECHANISM

Alon, Fischer, Procaccia, & Tennenholz (2011)

Input: a directed graph

1. Randomly partition the nodes into two sets S and W
2. The node of set W with the **highest number** of **incoming edges from set S** wins

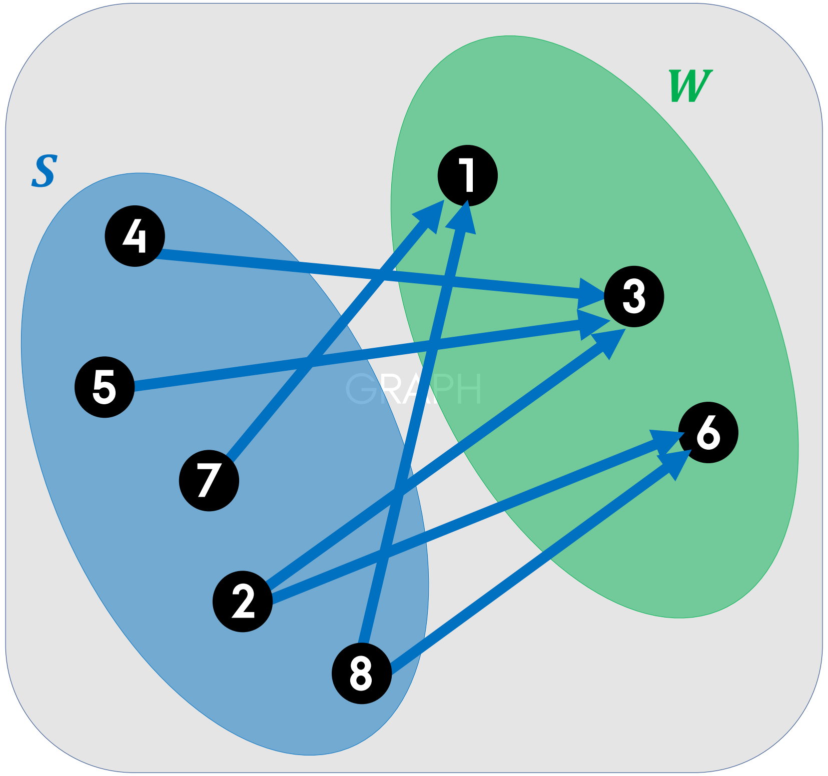


A 4-APPROXIMATION IMPARTIAL MECHANISM

Alon, Fischer, Procaccia, & Tennenholz (2011)

Input: a directed graph

1. Randomly partition the nodes into two sets ***S*** and ***W***
2. The node of set W with the **highest number** of **incoming edges from set S** wins

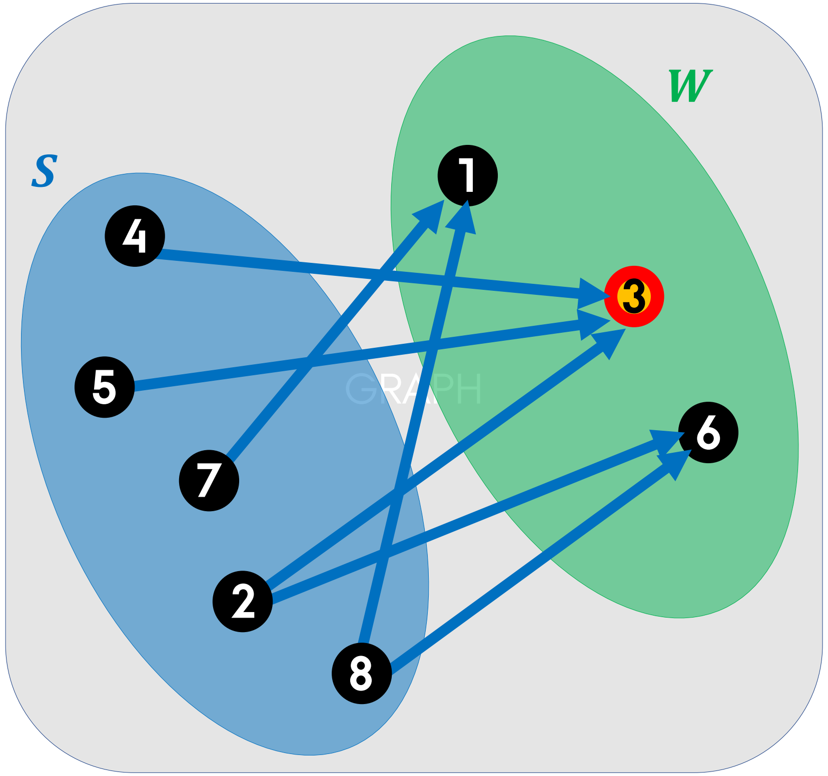


A 4-APPROXIMATION IMPARTIAL MECHANISM

Alon, Fischer, Procaccia, & Tennenholz (2011)

Input: a directed graph

1. Randomly partition the nodes into two sets S and W
2. The node of set W with the **highest number** of **incoming edges from set S** wins



A 4-APPROXIMATION IMPARTIAL MECHANISM

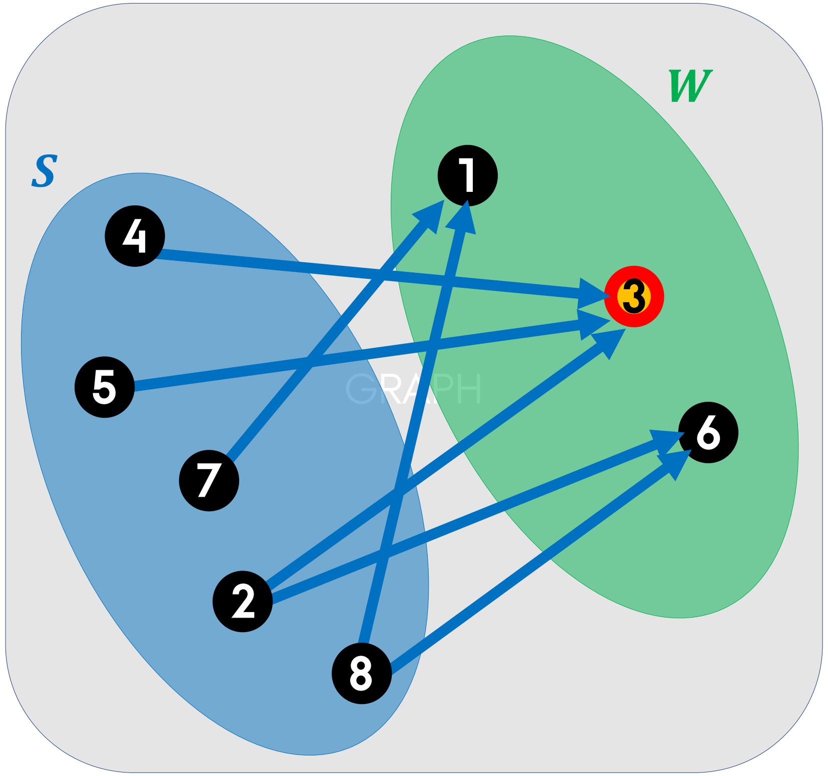
Alon, Fischer, Procaccia, & Tennenholz (2011)

Input: a directed graph

1. Randomly partition the nodes into two sets S and W
2. The node of set W with the **highest number** of **incoming edges from set S** wins

Impartiality:

- A node of S cannot win
- Edges from nodes in W are not taken into account



A 4-APPROXIMATION IMPARTIAL MECHANISM

Alon, Fischer, Procaccia, & Tennenholz (2011)

Input: a directed graph

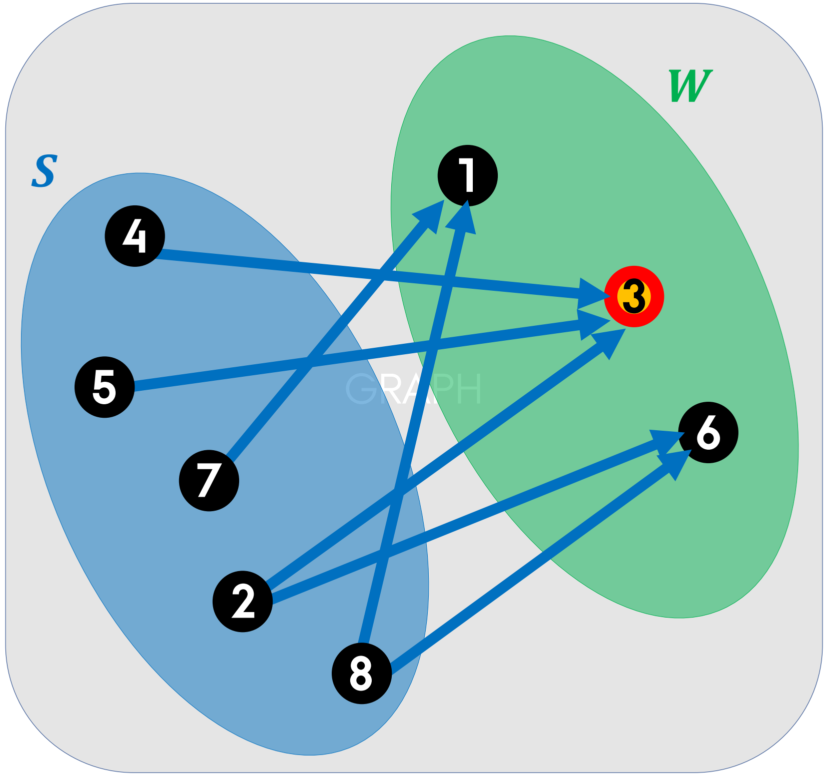
1. Randomly partition the nodes into two sets ***S*** and ***W***
2. The node of set W with the **highest number** of **incoming edges from set S** wins

Impartiality:

- A node of S cannot win
- Edges from nodes in W are not taken into account

Approximation ratio:

- The highest degree node u^* belongs to set W with **probability 1/2**
- Then, its expected in-degree from edges originating from set $\boldsymbol{S}$ is **half** the total in-degree



OPTIMAL RESULTS

Lower bound of 2

- Alon, Fischer, Procaccia, & Tennenholz (2011)

2-approximate impartial selection mechanism

- Fischer and Klimm (2015)
- Extends the random partition method

Other results

- Holzman & Moulin (2013)
- Busquet, Norin, & Vetta (2014)
- Bjalde, Fischer, & Klimm (2017)

ADDITIVE APPROXIMATION GUARANTEES

WHY ADDITIVE APPROXIMATION?

Worst-case scenario for approximation ratio is for **small graphs**

- Fischer & Klimm (2015)

If the **maximum degree is large**, approximation ratio is nearly optimal

- Bousquet, Norin, & Vetta (2014)

Definition: a mechanism yields an **$\delta(n)$-additive approximation** if for every n-node graph,

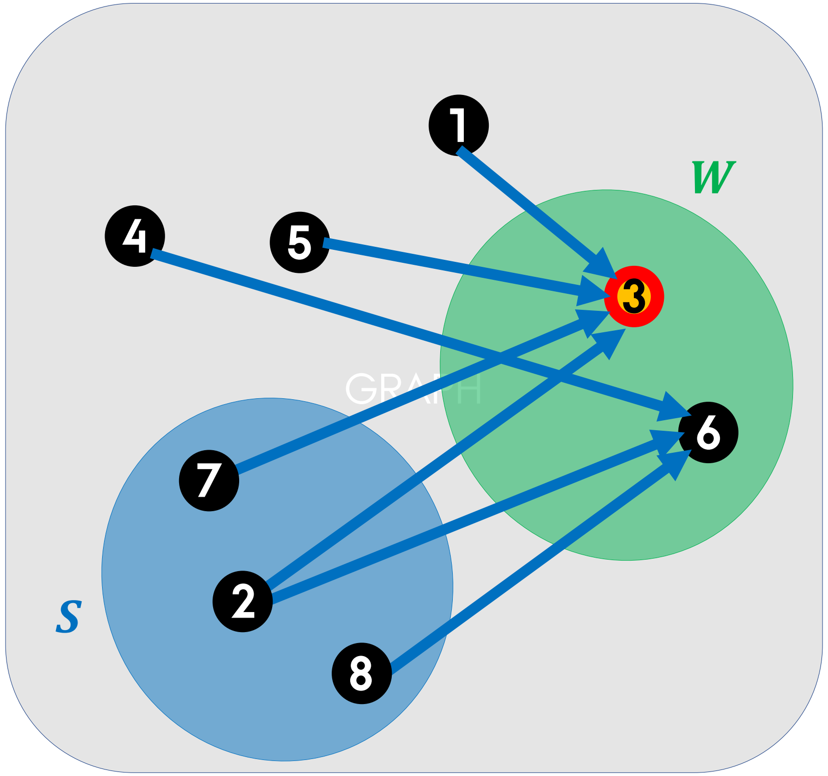
$$\text{maximum degree} - \text{expected degree of the winner} \leq \delta(n)$$

SAMPLE MECHANISMS

1. Given an input graph, select a **sample set** of nodes S
2. Let W be the **nodes nominated** by the nodes in S
3. Select the **winner from set W**

Strong sample mechanisms

- select the sample set **impartially**



OUR RESULTS

Upper bounds: two randomized strong sample mechanisms

- $O(\sqrt{n})$-additive approximation when each node has out-degree 1 (single nomination)
- $\boldsymbol{O(n^{2/3}\ln^{1/3}n)}$**-additive approximation** in general

Lower bounds on the additive approximation of strong sample mechanisms in the single-nomination model:

- $n-2$ for deterministic sample mechanisms
- $\Omega(\sqrt{n})$ for randomized sample mechanisms

General lower bound of 3

A SIMPLE K-SAMPLE MECHANISM

1. Form a sample set S by repeating k node selections uniformly at random with replacement
2. The node of set W with **highest in-degree** **from edges originating from S** wins

A SIMPLE K-SAMPLE MECHANISM (ANALYSIS)

1. Form a sample set S by repeating k node selections uniformly at random with replacement
2. The node of set W with **highest in-degree** **from edges originating from S** wins

Analysis idea:

- For every node v, $\deg_S(v)$ is a sum of Bernoulli random variables with expectation $\frac{k}{n}\deg(v)$
- Let u^* be a node of highest degree Δ
- A node of degree at least $\Delta - k$ wins (at least) when
  - node u^* is not selected in the sample and
  - gets more incoming edges than any node of degree less than $\Delta - k$

A SIMPLE K-SAMPLE MECHANISM (ANALYSIS)

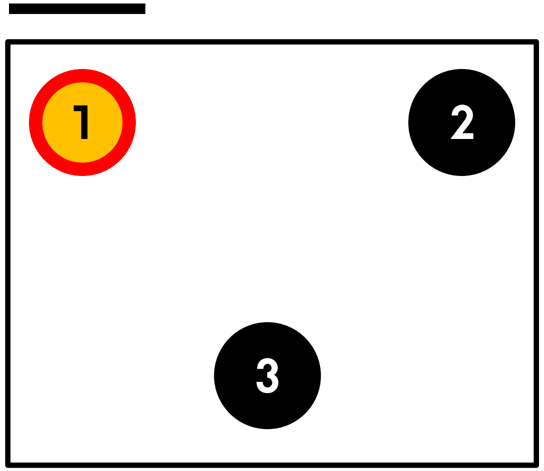
1. Form a sample set S by repeating k node selections uniformly at random with replacement
2. The node of set W with **highest in-degree** **from edges originating from S** wins

Analysis idea:

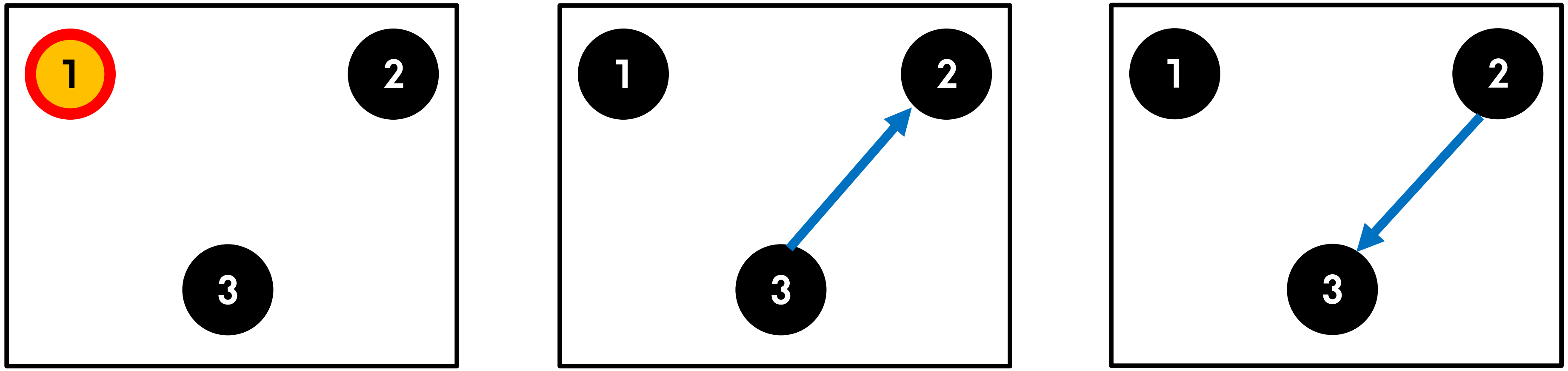
- For every node v, $\deg_S(v)$ is a sum of Bernoulli random variables with expectation $\frac{k}{n}\deg(v)$
- Let u^* be a node of highest degree Δ
- A node of degree at least $\Delta - k$ wins (at least) when
  - node u^* is not selected in the sample and [**so, k should be small**]
  - gets more incoming edges than any node of degree less than $\Delta - k$ [**so, k should be large**, **analysis using a Hoeffding bound**]

An $O(n^{2/3}\ln^{1/3}n)$–additive approximation follows by setting $k = \Theta(n^{2/3}\ln^{1/3}n)$

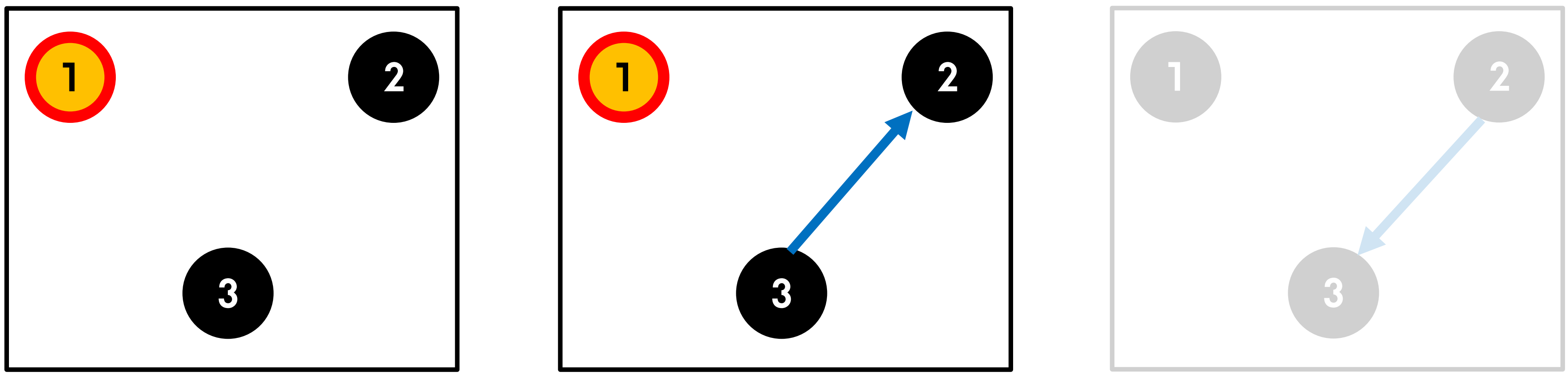
A SIMPLE ADDITIVE LOWER BOUND OF 2



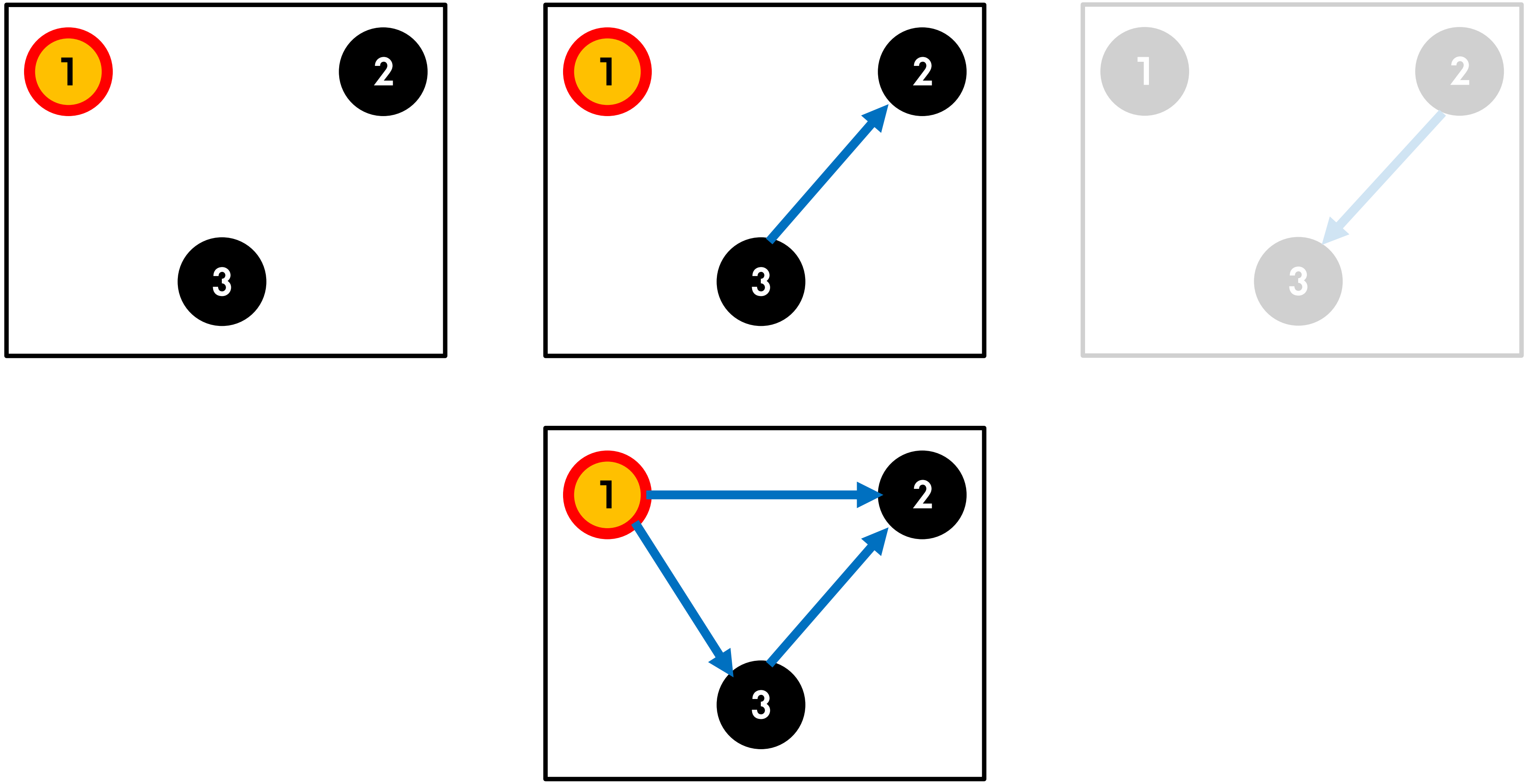
A SIMPLE ADDITIVE LOWER BOUND OF 2



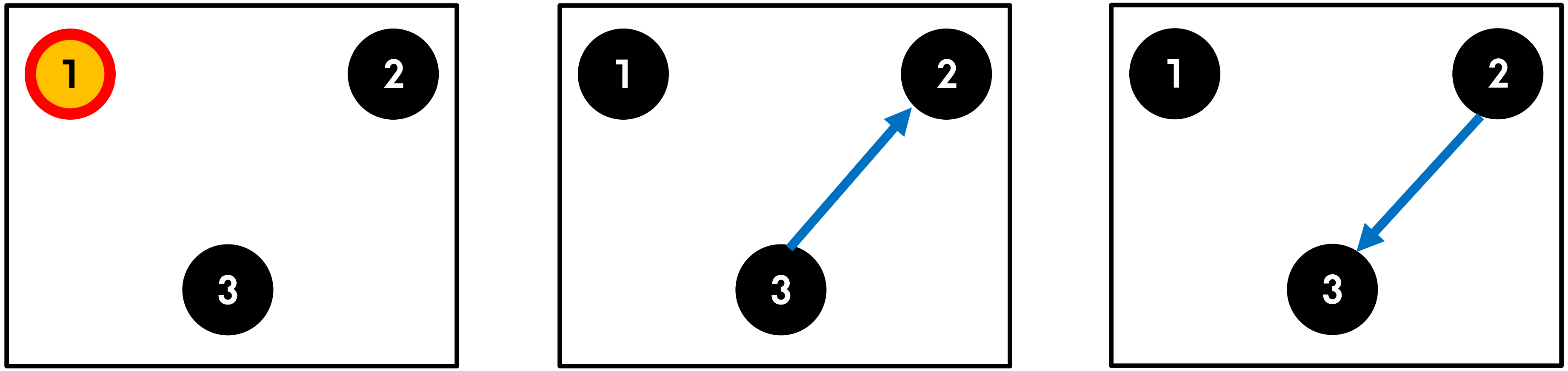
A SIMPLE ADDITIVE LOWER BOUND OF 2



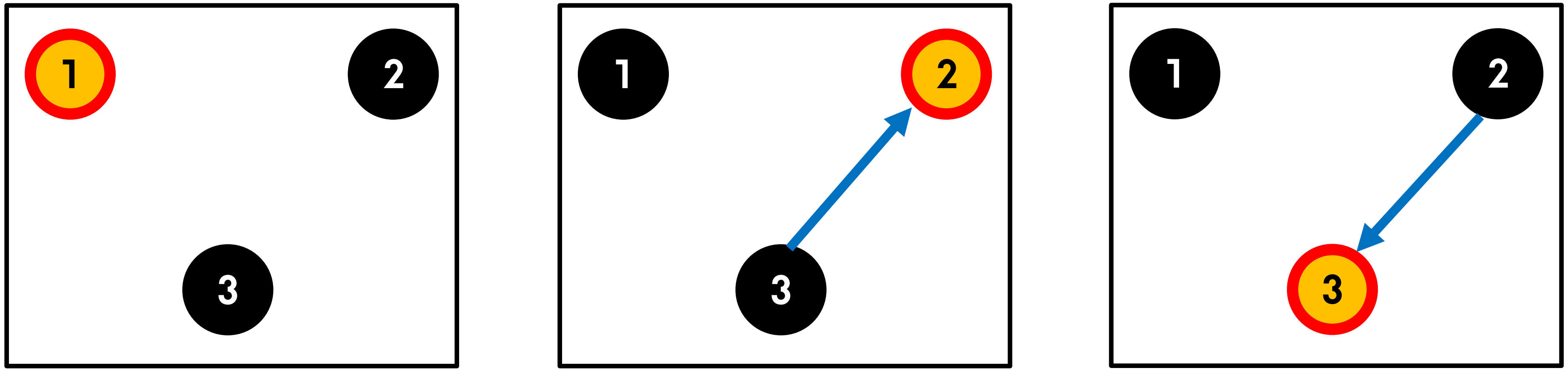
A SIMPLE ADDITIVE LOWER BOUND OF 2



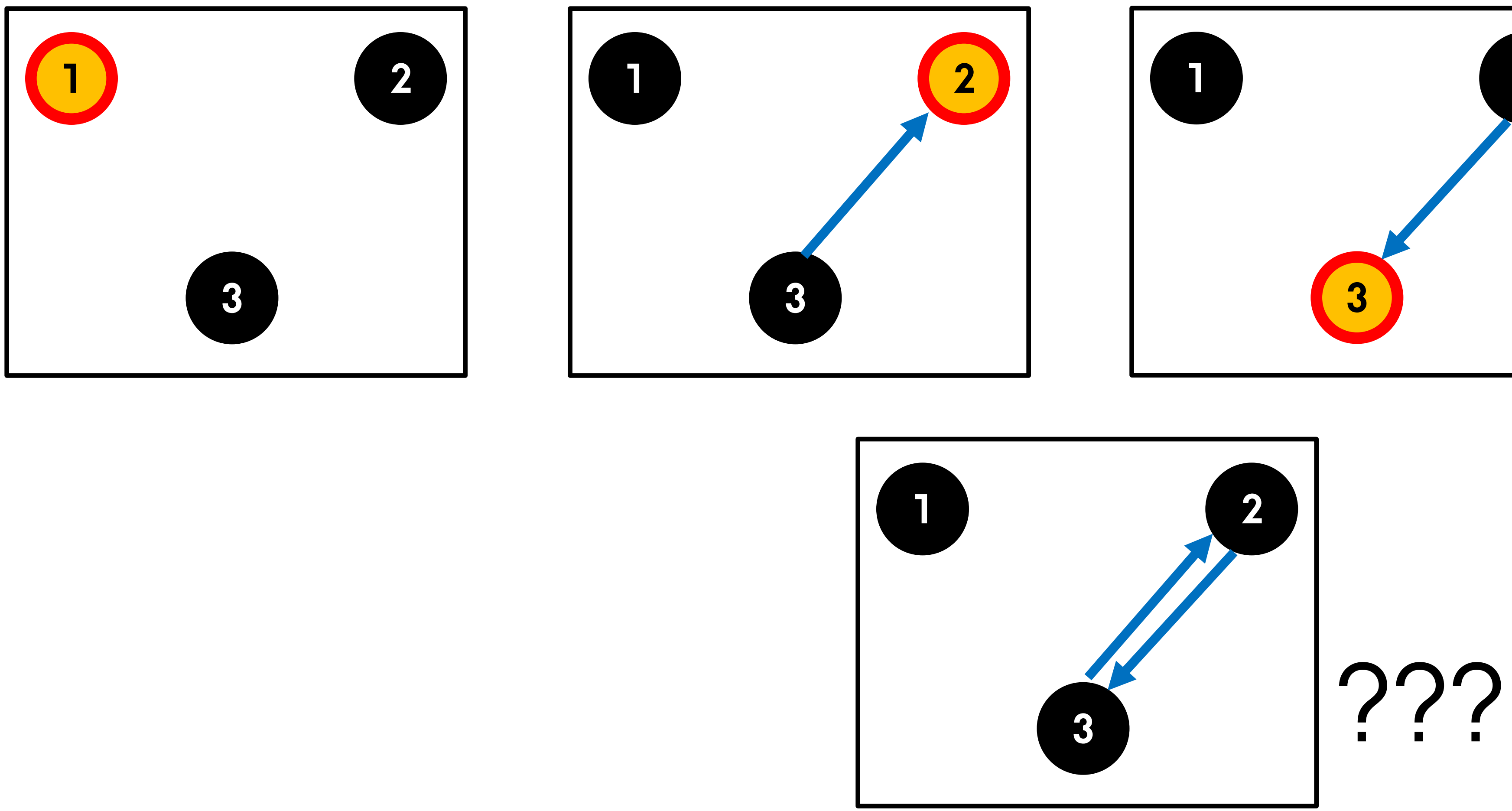
A SIMPLE ADDITIVE LOWER BOUND OF 2



A SIMPLE ADDITIVE LOWER BOUND OF 2



A SIMPLE ADDITIVE LOWER BOUND OF 2



OPEN PROBLEMS

Close the gap between 3 and $n-1$ for deterministic mechanisms

Improve the $O(n^{2/3}\ln^{1/3}n)$ bound for randomized mechanisms

Is **$O(1)$-additive approximation** possible?

USING PRIOR INFORMATION

THE MODEL

Input: random n-node graph, selected according to a **probability distribution $\mathbf{P}$**

Main assumption: **voter independence**

Objective: given (information about) $\mathbf{P}$, design an impartial mechanism with as **low expected additive approximation** as possible

Hierarchy of distributions (models):

- **Opinion poll**: each node v selects its set of outgoing edges according to a probability distribution $\mathbf{P}_v$
- **A priori popularity**: node v has popularity $p_v \in [0,1]$ and the edge (u, v) exists independently with probability p_v
- **Uniform**: a priori popularity with $p_v = 1/2$

THE CONSTANT MECHANISM

Return a **fixed** node

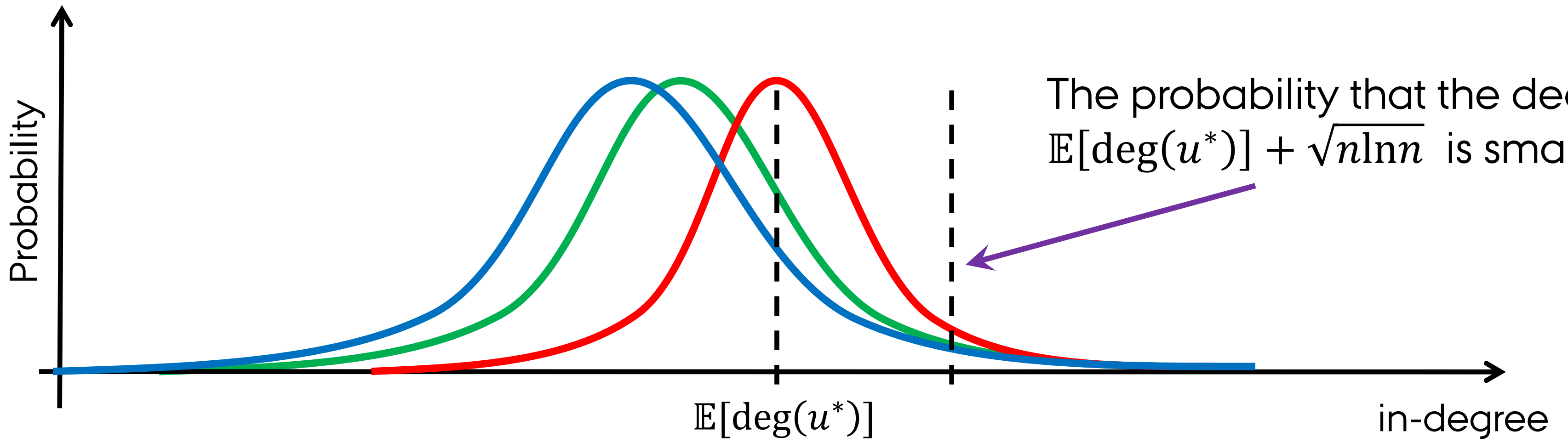
E.g., return the node of **highest expected degree** according to $\mathbf{P}$

THE CONSTANT MECHANISM (ANALYSIS)

Return a **fixed** node

E.g., return the node of **highest expected degree** according to $\mathbf{P}$

Analysis: Due to voter independence, the in-degree of each node is a sum of Bernoulli trials, even in the opinion poll model



The probability that the degree of some node exceeds $\mathbb{E}[\deg(u^*)] + \sqrt{n \ln n}$ is small (Hoeffding bound)

APPROVAL VOTING WITH DEFAULT

Mechanism **AVD**

Extends a mechanism by Holzman & Moulin (2013)

Informal definition:

- The highest-degree node wins, if it is **unique**
- In case of **ties**, a preselected **default node t** wins

APPROVAL VOTING WITH DEFAULT

Mechanism **AVD**

Extends a mechanism by Holzman & Moulin (2013)

Informal definition:

- The highest-degree node wins, if it is **unique**
- In case of **ties**, a preselected **default node t** wins

Formal definition:

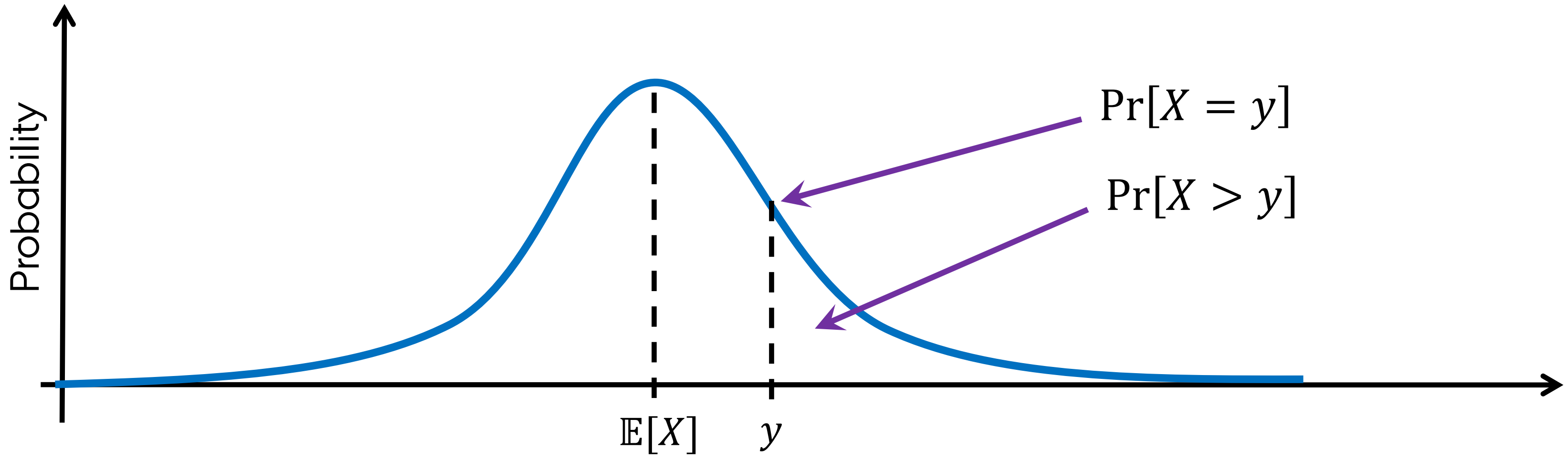
- Compare the degrees of two nodes u and v, ignoring the edges between them and the edges originating from the default node t
- If there is a node that beats all other nodes in their pairwise comparison, it is the winner
- Otherwise, the default node wins

AVD HAS EXPECTED ADDITIVE APPROXIMATION …

- $O(\ln^2 n)$ in the a priori popularity model
- $\Omega(\ln n)$ on uniform instance
- Unfortunately, as bad as $\Theta\left(\sqrt{n \ln n}\right)$ in the opinion poll model

A FEW WORDS ABOUT THE ANALYSIS

- Node degrees follow the **binomial** probability distribution $\mathbf{B}(n, p_k)$
- A node of (almost) highest degree wins unless there is a "**tie at the top**"
- Bounding the expected additive approximation strongly depends on bounding the **hazard rate** $\Pr[X = y]$ / $\Pr[X > y]$ of a random variable $X \sim \mathbf{B}(n, p_k)$
- **Theorem**: The hazard rate of a binomial r.v. X is $\Theta\left(\sqrt{\frac{\ln n}{\min\{\mathbb{E}[X], n-\mathbb{E}[X]\}}}\right)$ for values of y close to $\mathbb{E}[X]$



OPEN PROBLEMS

Polylogarithmic or constant expected additive approximation in the opinion poll model?

Variations of AVD?

What if prior information is **not accurate**?

- Rough estimates of the highest expected degree are enough to get the $O(\sqrt{n \ln n})$ bound.
- Can we recover the polylogarithmic result?

THANK YOU!