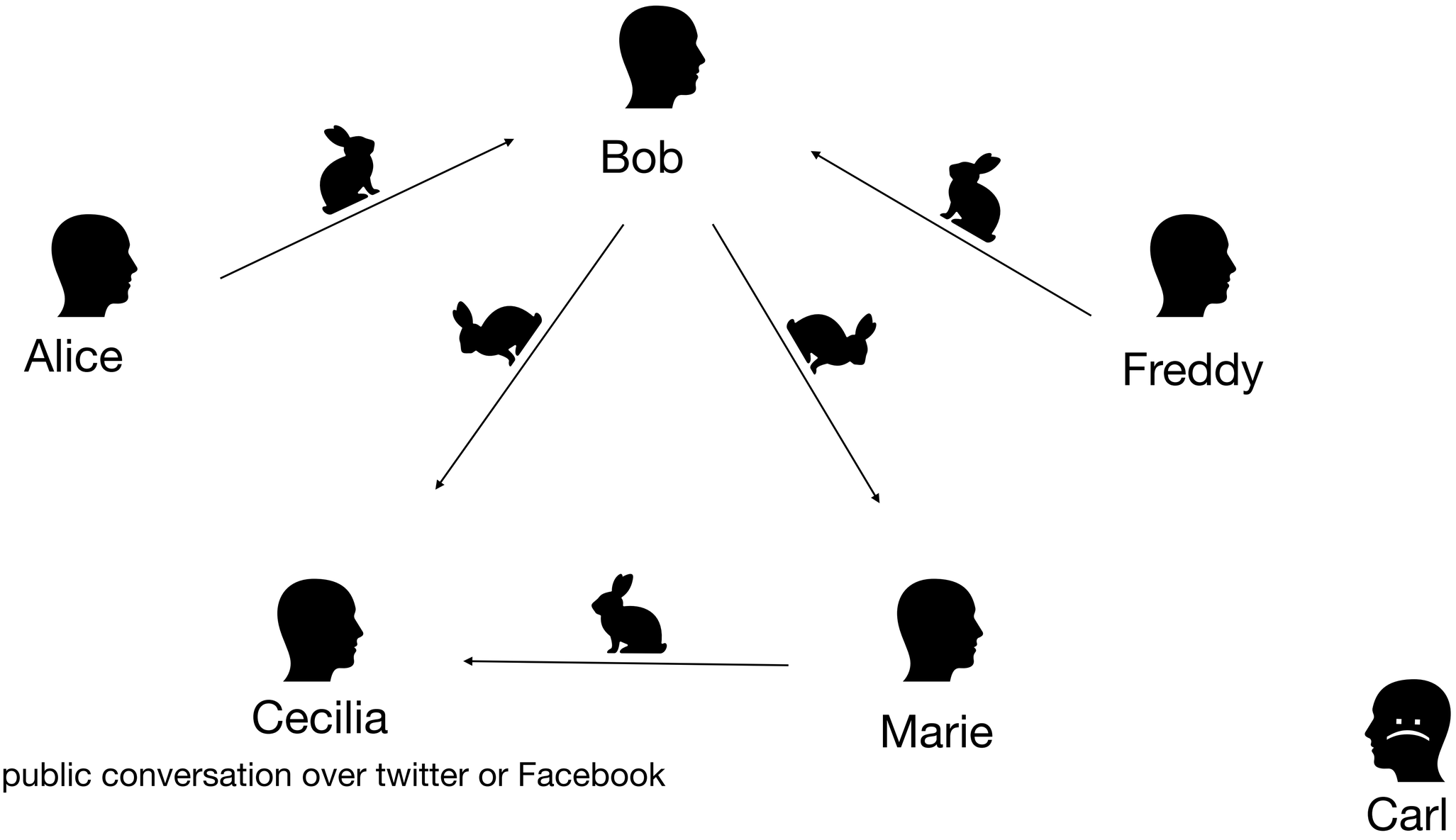
Lower Bounds for Anonymous Whistleblowing

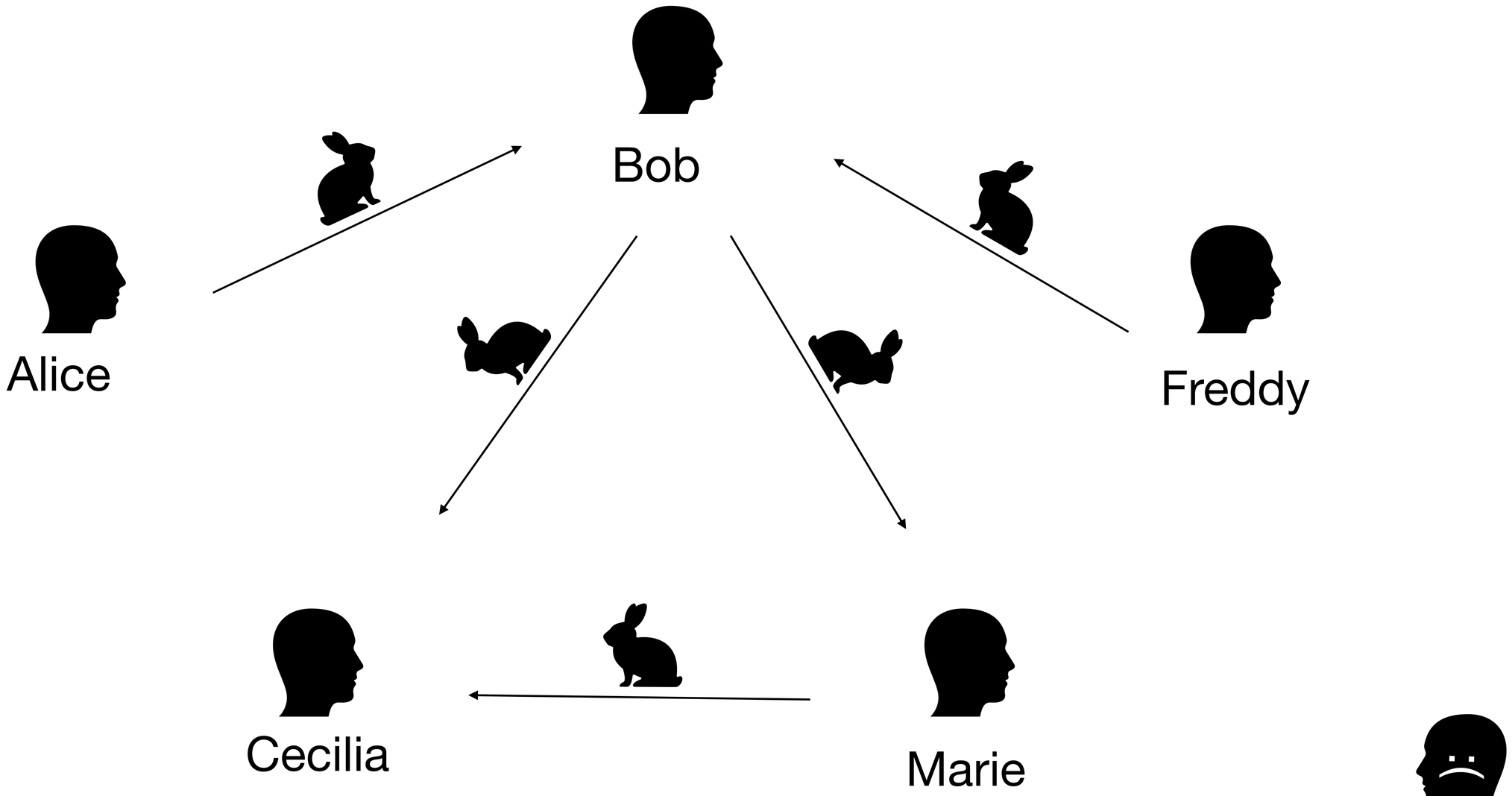
Willy Quach, LaKyah Tyner, Daniel Wichs

Anonymous Transfer [Agricola, Couteau, Maier 22]



- Friends having public conversation over twitter or Facebook

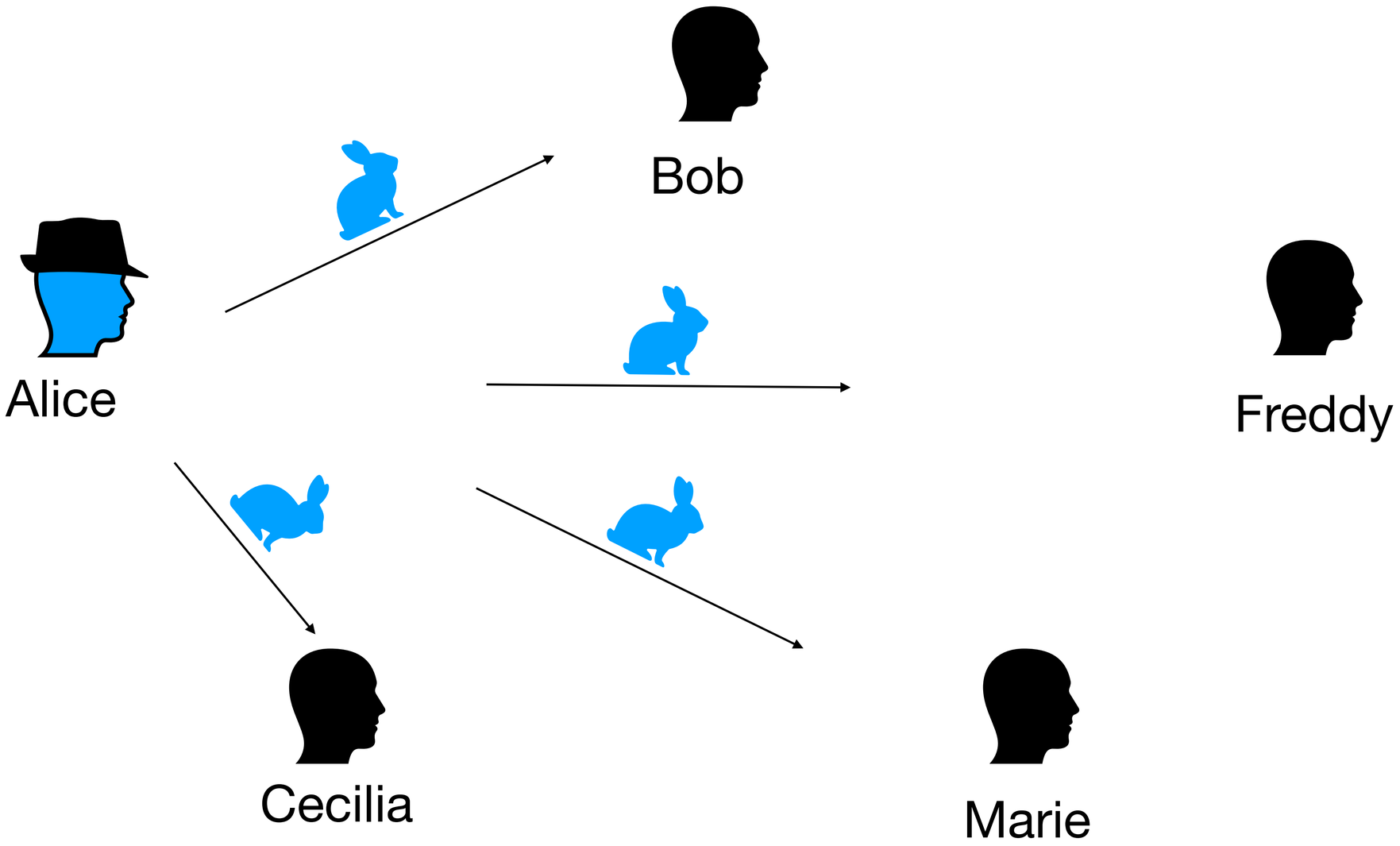
Anonymous Transfer [Agricola, Couteau, Maier 22]



Carl

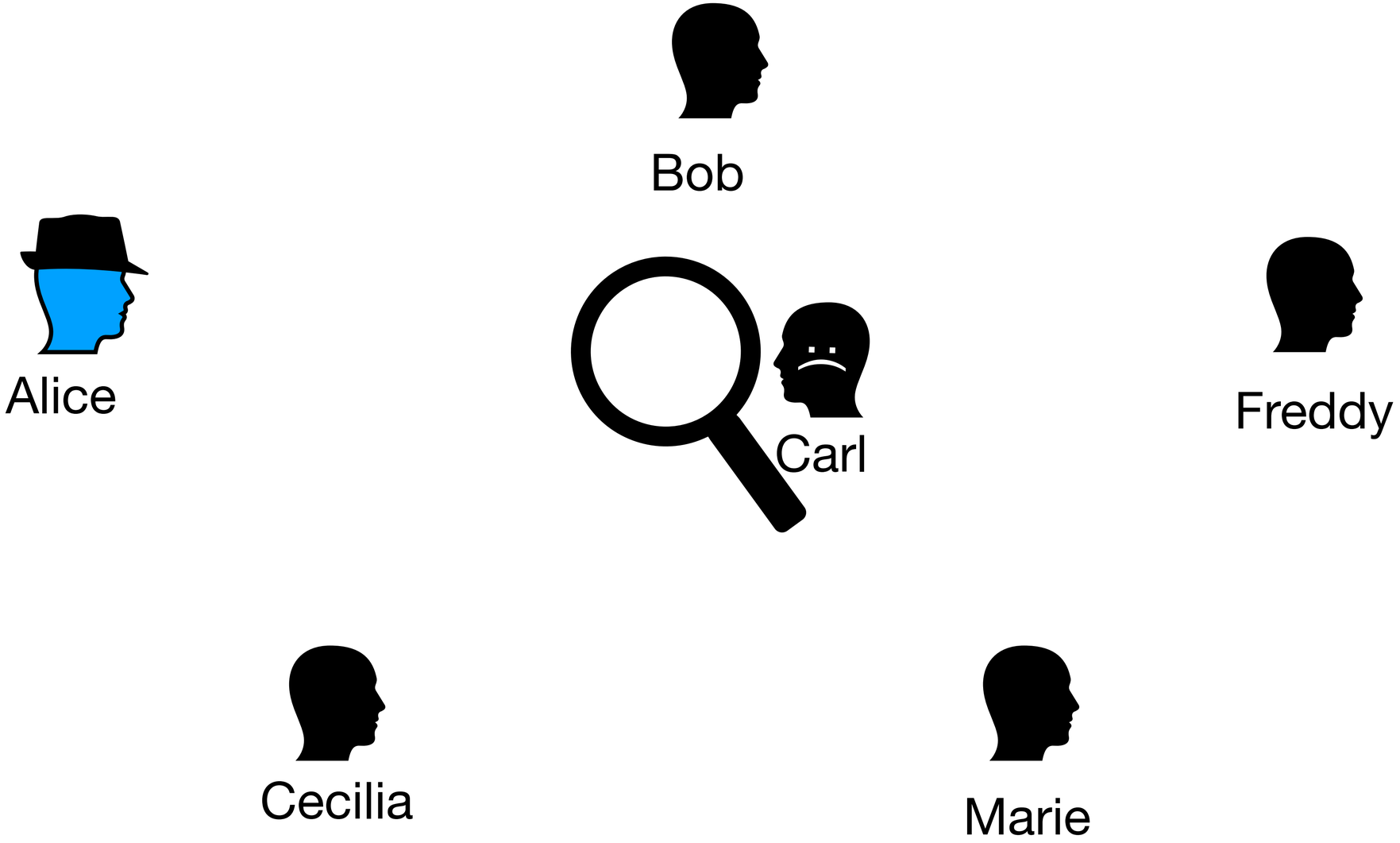
- Friends having public conversation over twitter or Facebook
- Everyone knows the know sent each message

Anonymous Transfer [Agricola, Couteau, Maier 22]



- One friend wants to transmit some secret message, unbeknownst to the others
- Without revealing identity

Anonymous Transfer [Agricola, Couteau, Maier 22]



Anyone can recover secret message (even an outsider) without discovering the sender

Motivation

- Ideal for facilitating whistleblowing

Motivation

- Ideal for facilitating whistleblowing
  - Whistleblowers act in an untrusted environment

Motivation

- Ideal for facilitating whistleblowing
  - Whistleblowers act in an untrusted environment
  - Face risk of punishment

Motivation

- Ideal for facilitating whistleblowing
  - Whistleblowers act in an untrusted environment
  - Face risk of punishment
  - Can we mitigate risk using cryptographic techniques?

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.
    - Other prior work on anonymous communication (e.g., Tor) require trusted parties (e.g., Tor relay nodes) to actively participate.

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.
    - Other prior work on anonymous communication (e.g., Tor) require trusted parties (e.g., Tor relay nodes) to actively participate.
- Technical: Model dummy messages as uniformly random strings. This is wlog since we can embed this in other distributions (e.g., conversation about bunnies). [HLv02, vH04, vHL05]

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.
    - Other prior work on anonymous communication (e.g., Tor) require trusted parties (e.g., Tor relay nodes) to actively participate.
- Technical: Model dummy messages as uniformly random strings. This is wlog since we can embed this in other distributions (e.g., conversation about bunnies). [HLv02, vH04, vHL05]
    - [ACM22] has two results

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.
    - Other prior work on anonymous communication (e.g., Tor) require trusted parties (e.g., Tor relay nodes) to actively participate.
- Technical: Model dummy messages as uniformly random strings. This is wlog since we can embed this in other distributions (e.g., conversation about bunnies). [HLv02, vH04, vHL05]
  - [ACM22] has two results
    - Positive result: a very weak form of AT is possible

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.
    - Other prior work on anonymous communication (e.g., Tor) require trusted parties (e.g., Tor relay nodes) to actively participate.
- Technical: Model dummy messages as uniformly random strings. This is wlog since we can embed this in other distributions (e.g., conversation about bunnies). [HLv02, vH04, vHL05]
  - [ACM22] has two results
    - Positive result: a very weak form of AT is possible
    - Negative result: a very strong form of AT is impossible

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.
    - Other prior work on anonymous communication (e.g., Tor) require trusted parties (e.g., Tor relay nodes) to actively participate.
- Technical: Model dummy messages as uniformly random strings. This is wlog since we can embed this in other distributions (e.g., conversation about bunnies). [HLv02, vH04, vHL05]
  - [ACM22] has two results
    - Positive result: a very weak form of AT is possible
    - Negative result: a very strong form of AT is impossible
    - Leaves a big unknown gap between them

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.
    - Other prior work on anonymous communication (e.g., Tor) require trusted parties (e.g., Tor relay nodes) to actively participate.
- Technical: Model dummy messages as uniformly random strings. This is wlog since we can embed this in other distributions (e.g., conversation about bunnies). [HLv02, vH04, vHL05]
  - [ACM22] has two results
    - Positive result: a very weak form of AT is possible
    - Negative result: a very strong form of AT is impossible
    - Leaves a big unknown gap between them
  - Our work closes the gap by extending the negative results

Prior Work

- Anonymous Transfer (AT) introduced by [ACM22].
    - Main novelty: no reliance on any trusted parties beyond non-senders generating dummy traffic.
    - Other prior work on anonymous communication (e.g., Tor) require trusted parties (e.g., Tor relay nodes) to actively participate.
- Technical: Model dummy messages as uniformly random strings. This is wlog since we can embed this in other distributions (e.g., conversation about bunnies). [HLv02, vH04, vHL05]
  - [ACM22] has two results
    - Positive result: a very weak form of AT is possible
    - Negative result: a very strong form of AT is impossible
    - Leaves a big unknown gap between them
  - Our work closes the gap by extending the negative results
      - Their very weak form of AT is the best we can hope for

Anonymous Transfer Specifics [ACM22]

- Focuses on c-round, 2 party AT (sender, non-sender)

Anonymous Transfer Specifics [ACM22]

- Focuses on c-round, 2 party AT (sender, non-sender)
  - Lower bounds imply ones of N-party [ACM22]

Anonymous Transfer Specifics [ACM22]

- Focuses on c-round, 2 party AT (sender, non-sender)
  - Lower bounds imply ones of N-party [ACM22]

- AT algorithms:

Anonymous Transfer Specifics [ACM22]

- Focuses on c-round, 2 party AT (sender, non-sender)
  - Lower bounds imply ones of N-party [ACM22]

- AT algorithms:
  - Trusted S*etup* → *crs*

Anonymous Transfer Specifics [ACM22]

- Focuses on c-round, 2 party AT (sender, non-sender)
  - Lower bounds imply ones of N-party [ACM22]

- AT algorithms:
  - Trusted S*etup* $\rightarrow$ *crs*
  - T*ransfer*(μ) $\rightarrow \pi$

Anonymous Transfer Specifics [ACM22]

- Focuses on c-round, 2 party AT (sender, non-sender)
  - Lower bounds imply ones of N-party [ACM22]

- AT algorithms:
  - Trusted S*etup* $\rightarrow$ *crs*
  - T*ransfer*(μ) $\rightarrow \pi$
  - R*econstruct*(π) $\rightarrow \mu'$

Anonymous Transfer Specifics [ACM22]

- Focuses on c-round, 2 party AT (sender, non-sender)
  - Lower bounds imply ones of N-party [ACM22]

- AT algorithms:
  - Trusted *Setup* $\rightarrow$ *crs*
  - *Transfer*(μ) $\rightarrow \pi$
  - *Reconstruct*(π) $\rightarrow \mu'$

Correctness

For all secret messages $\mu \in \{0,1\}^{\ell}$

$\Pr[\mu' \neq \mu]$ is negligible

Anonymous Transfer Specifics [ACM22]

- Focuses on c-round, 2 party AT (sender, non-sender)
  - Lower bounds imply ones of N-party [ACM22]

- AT algorithms:
  - Trusted *Setup* $\rightarrow$ *crs*
  - *Transfer*(μ) $\rightarrow \pi$
  - *Reconstruct*(π) $\rightarrow \mu'$

Correctness

For all secret messages $\mu \in \{0,1\}^\ell$

$\Pr[\mu' \neq \mu]$ is negligible

δ-anonymity "Distinguishing Advantage"

For all PPT D and all $\mu \in \{0,1\}^\ell$

$$|\Pr[D(\pi^A) = 1] - \Pr[D(\pi^B) = 1]| \leq \delta$$

Comparison: This work and [ACM22]

Comparison: This work and [ACM22]

[ACM22] Negative Result

Cannot get negligible anonymity δ against all poly-time adversaries

Comparison: This work and [ACM22]

[ACM22] Negative Result

Cannot get negligible anonymity δ against all poly-time adversaries

[ACM22] Positive Result

AT with anonymity $\delta = 1/c$ against *fine-grained adversaries* whose runtime is O(c) x honest parties. (strong assumptions)

Comparison: This work and [ACM22]

[ACM22] Negative Result

Cannot get negligible anonymity δ against all poly-time adversaries

[ACM22] Positive Result

AT with anonymity $\delta = 1/c$ against *fine-grained adversaries* whose runtime is O(c) x honest parties. (strong assumptions)

Big Gap: Can we get "decent" anonymity (say $\delta = .01$) against all poly adversaries? Can we get negligible anonymity against fine grained adversaries?

Our results: NO

Comparison: This work and [ACM22]

[ACM22] Negative Result

Cannot get negligible anonymity δ against all poly-time adversaries

[ACM22] Positive Result

AT with anonymity $\delta = 1/c$ against *fine-grained adversaries* whose runtime is O(c) x honest parties. (strong assumptions)

Big Gap: Can we get "decent" anonymity (say $\delta = .01$) against all poly adversaries? Can we get negligible anonymity against fine grained adversaries?

Our results: NO

Our Negative Result 1

Cannot get security against all poly adversaries with any non-trivial anonymity $\delta<1$

Comparison: This work and [ACM22]

[ACM22] Negative Result

Cannot get negligible anonymity δ against all poly-time adversaries

[ACM22] Positive Result

AT with anonymity $\delta = 1/c$ against *fine-grained adversaries* whose runtime is O(c) x honest parties. (strong assumptions)

Big Gap: Can we get "decent" anonymity (say $\delta = .01$) against all poly adversaries? Can we get negligible anonymity against fine grained adversaries?

Our results: NO

Our Negative Result 1

Cannot get security against all poly adversaries with any non-trivial anonymity $\delta<1$

Our Negative Result 2

Cannot get negligible anonymity even against fine-grained adversaries

Our Main Contribution

Attack on anonymity of AT

Our Main Contribution

Attack on anonymity of AT

Goal: Given transcript π of the protocol, identify the sender

Our Main Contribution

Attack on anonymity of AT

Goal: Given transcript π of the protocol, identify the sender

Consider the notion of “progress” towards correctly recovering message

Our Main Contribution

Attack on anonymity of AT

Goal: Given transcript π of the protocol, identify the sender

Consider the notion of “progress” towards correctly recovering message

- “progress” of partial transcript $\pi[i]$

Our Main Contribution

Attack on anonymity of AT

Goal: Given transcript π of the protocol, identify the sender

Consider the notion of "progress" towards correctly recovering message

- "progress" of partial transcript $\pi[i]$

$$\pi = \{m_1, \ldots, m_i, m_{i+1}, \ldots, m_{|\pi|}\} \rightarrow \pi[i] = \{m_1, \ldots, m_i, r_{i+1}, \ldots, r_{|\pi|}\}$$

Our Main Contribution

Attack on anonymity of AT

Goal: Given transcript π of the protocol, identify the sender

Consider the notion of "progress" towards correctly recovering message

- "progress" of partial transcript $\pi[i]$

$$\pi = \{m_1, \ldots, m_i, m_{i+1}, \ldots, m_{|\pi|}\} \rightarrow \pi[i] = \{m_1, \ldots, m_i, r_{i+1}, \ldots, r_{|\pi|}\}$$

- The party who makes the most progress is the sender

Our Main Contribution

Attack on anonymity of AT

$p_i :=$ probability of correctly recovering message after the i-th message associated with $\pi[i]$

Our Main Contribution

Attack on anonymity of AT

$p_i :=$ probability of correctly recovering message after the i-th message associated with $\pi[i]$

Our Main Contribution

Attack on anonymity of AT

$p_i :=$ probability of correctly recovering message after the i-th message associated with $\pi[i]$

Our Main Contribution

Attack on anonymity of AT

$p_i :=$ probability of correctly recovering message after the i-th message associated with $\pi[i]$

- Assign progress from $p_{i-1} \rightarrow p_i$ to A

Our Main Contribution

Attack on anonymity of AT

$p_i :=$ probability of correctly recovering message after the i-th message associated with $\pi[i]$

- Assign progress from $p_{i-1} \rightarrow p_i$ to A
- Main insight: Non-sender messages do not (on expectation) change p_i

Our Main Contribution

Attack on anonymity of AT

Our Main Contribution

Attack on anonymity of AT

Blueprint: Estimate each party's contribution

Our Main Contribution

Attack on anonymity of AT

Blueprint: Estimate each party's contribution

Argue:

Our Main Contribution

Attack on anonymity of AT

Blueprint: Estimate each party's contribution

Argue:

1. The contribution of the non-sender is small

Our Main Contribution

Attack on anonymity of AT

Blueprint: Estimate each party's contribution

Argue:

1. The contribution of the non-sender is small

2. Total contribution is large so, the party who contributed the most must be the sender

Covert Cheating Game

- Abstract blueprint into the Cover Cheating Game
- Between two player  and 

Covert Cheating Game

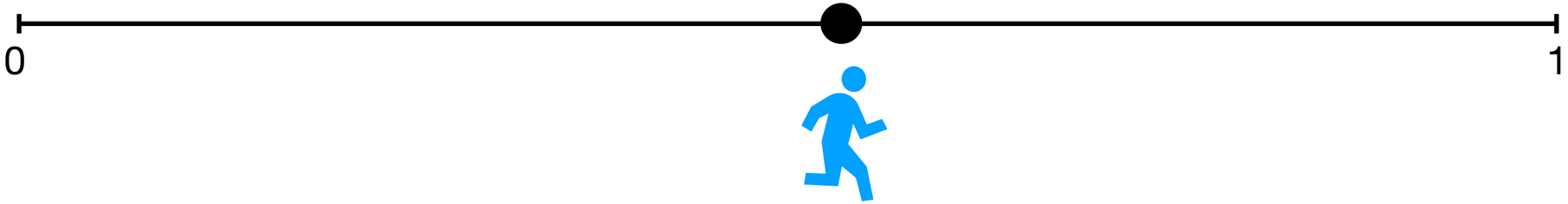
- Abstract blueprint into the Cover Cheating Game
- Between two player  and 

Covert Cheating Game

- Abstract blueprint into the Cover Cheating Game
- Between two player  and 

Covert Cheating Game

- Abstract blueprint into the Cover Cheating Game
- Between two player  and 



Covert Cheating Game

- Abstract blueprint into the Cover Cheating Game
- Between two player  and 

Covert Cheating Game

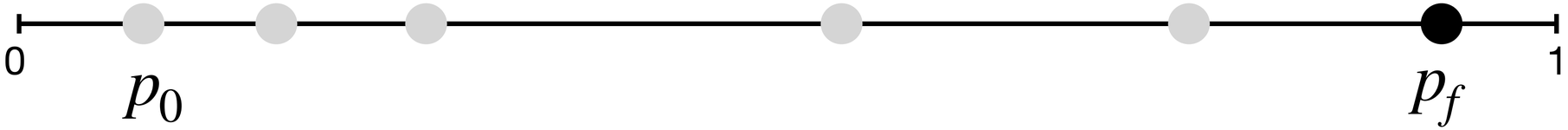
- Abstract blueprint into the Cover Cheating Game
- Between two player  and 

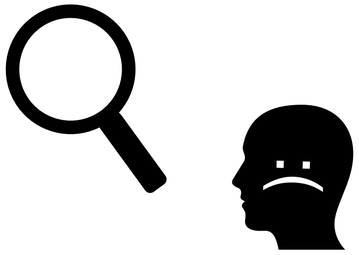
Covert Cheating Game

- Abstract blueprint into the Cover Cheating Game
- Between two player  and 

Covert Cheating Game

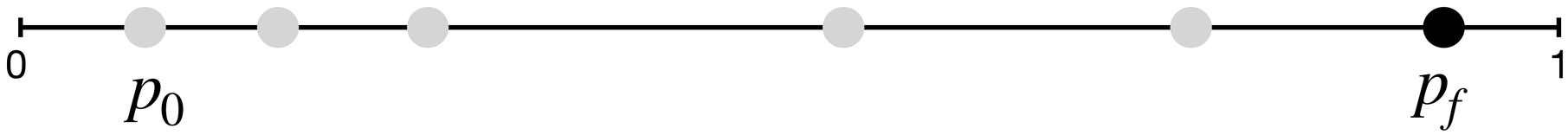
- Abstract blueprint into the Cover Cheating Game
- Between two player  and 





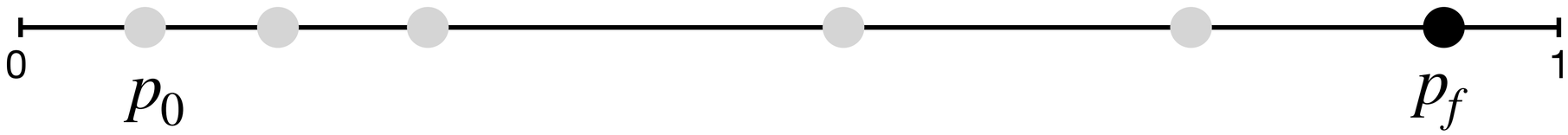
Carl

Generic Attack with Large Advantage



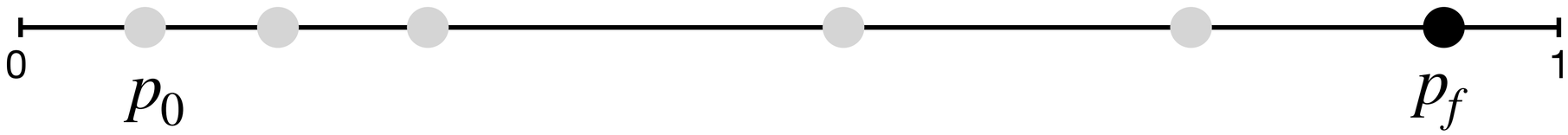
Generic Attack with Large Advantage

- Uses the fact that non-biased player cannot change state much



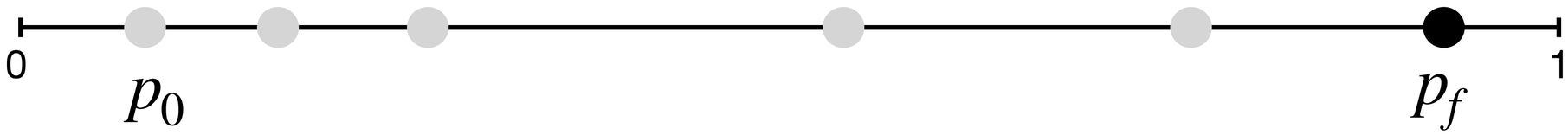
Generic Attack with Large Advantage

- Uses the fact that non-biased player cannot change state much
  - If p_{i-1} is close to zero then p_i can't be very different [by Markov]



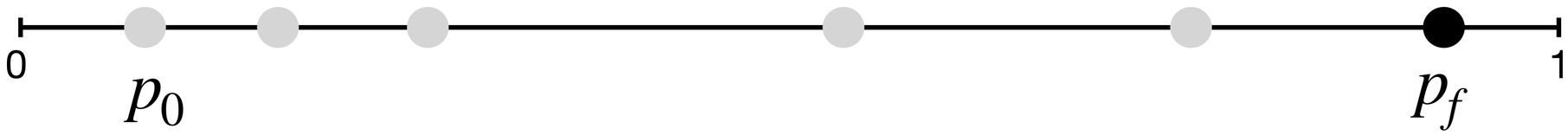
Generic Attack with Large Advantage

- Uses the fact that non-biased player cannot change state much
  - If p_{i-1} is close to zero then p_i can't be very different [by Markov]
- Task: Weigh progress made close to zero higher



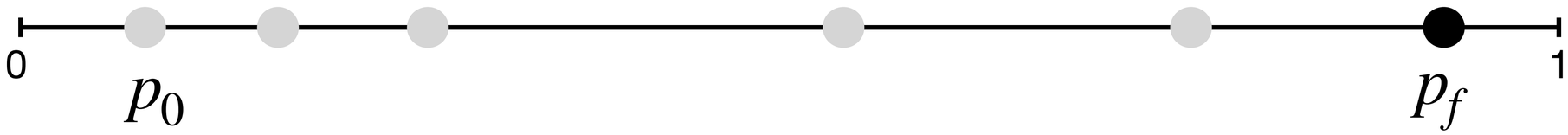
Generic Attack with Large Advantage

- Uses the fact that non-biased player cannot change state much
  - If p_{i-1} is close to zero then p_i can't be very different [by Markov]
- Task: Weigh progress made close to zero higher
  - Larger progress made close to zero is made by the sender



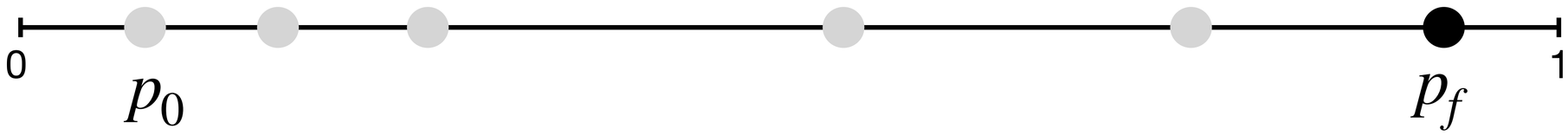
Generic Attack with Large Advantage

- Uses the fact that non-biased player cannot change state much
  - If p_{i-1} is close to zero then p_i can't be very different [by Markov]
- Task: Weigh progress made close to zero higher
  - Larger progress made close to zero is made by the sender
- Consider multiplicative progress



Generic Attack with Large Advantage

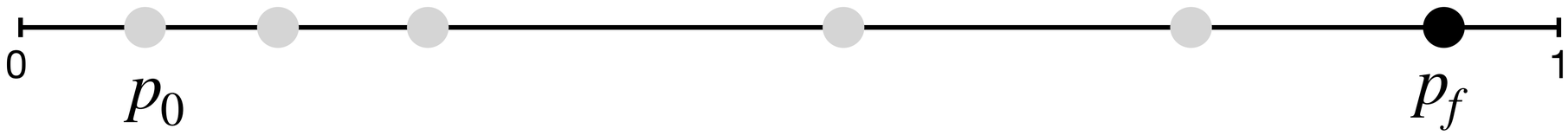
- Uses the fact that non-biased player cannot change state much
  - If p_{i-1} is close to zero then p_i can't be very different [by Markov]
- Task: Weigh progress made close to zero higher
  - Larger progress made close to zero is made by the sender
- Consider multiplicative progress
  - Progress from p_{i-1} to p_i is represented by



Generic Attack with Large Advantage

- Uses the fact that non-biased player cannot change state much
  - If p_{i-1} is close to zero then p_i can't be very different [by Markov]
- Task: Weigh progress made close to zero higher
  - Larger progress made close to zero is made by the sender
- Consider multiplicative progress
  - Progress from p_{i-1} to p_i is represented by

$$r_i = \frac{p_i}{p_{i-1}}$$

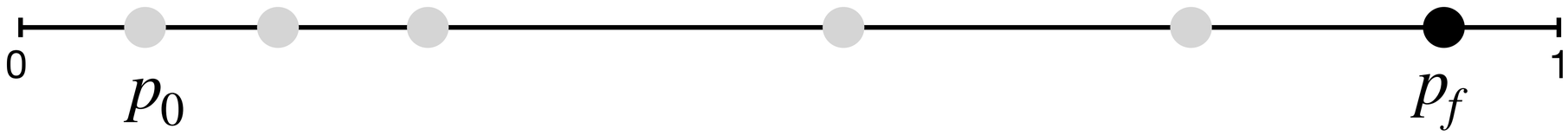


Generic Attack with Large Advantage

- Uses the fact that non-biased player cannot change state much
  - If p_{i-1} is close to zero then p_i can't be very different [by Markov]
- Task: Weigh progress made close to zero higher
  - Larger progress made close to zero is made by the sender
- Consider multiplicative progress
  - Progress from p_{i-1} to p_i is represented by

$$r_i = \frac{p_i}{p_{i-1}}$$

  - Total progress is



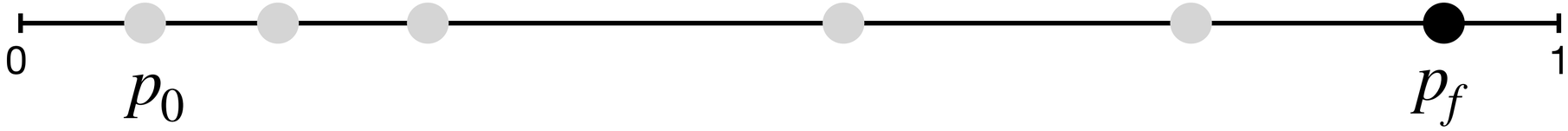
Generic Attack with Large Advantage

- Uses the fact that non-biased player cannot change state much
  - If p_{i-1} is close to zero then p_i can't be very different [by Markov]
- Task: Weigh progress made close to zero higher
  - Larger progress made close to zero is made by the sender
- Consider multiplicative progress
  - Progress from p_{i-1} to p_i is represented by

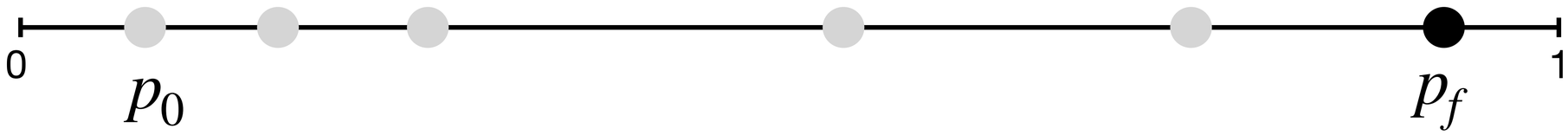
$$r_i = \frac{p_i}{p_{i-1}}$$

  - Total progress is

$$\Pi_i r_i = \Pi_i \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

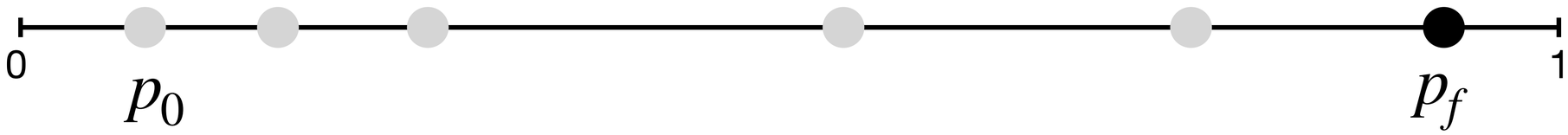


Generic Attack with Large Advantage



Generic Attack with Large Advantage

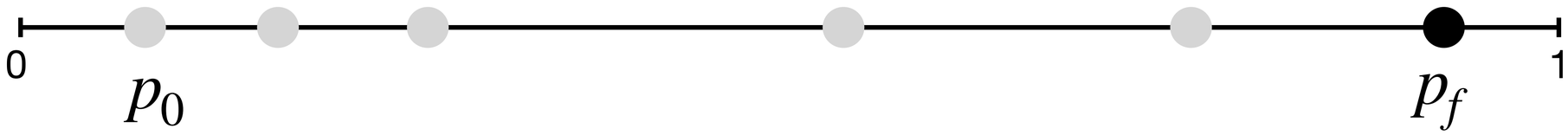
- Total progress is



Generic Attack with Large Advantage

- Total progress is

$$\Pi_i r_i = \Pi_i \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

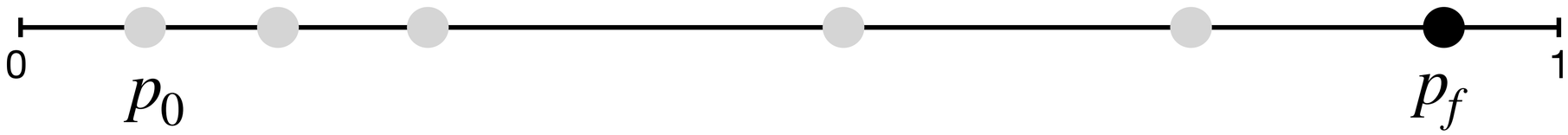


Generic Attack with Large Advantage

- Total progress is

$$\Pi_i r_i = \Pi_i \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

- One player must have progress



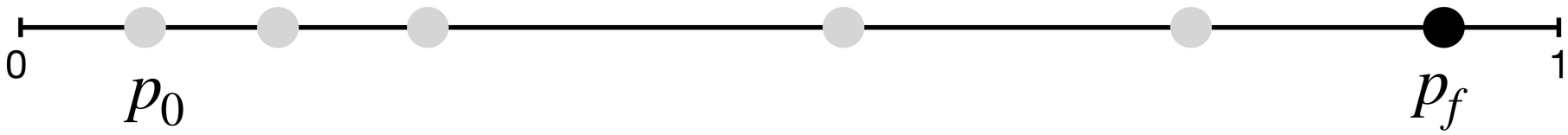
Generic Attack with Large Advantage

- Total progress is

$$\Pi_i r_i = \Pi_i \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

- One player must have progress

$$\geq \sqrt{\frac{p_f}{p_0}}$$



Generic Attack with Large Advantage

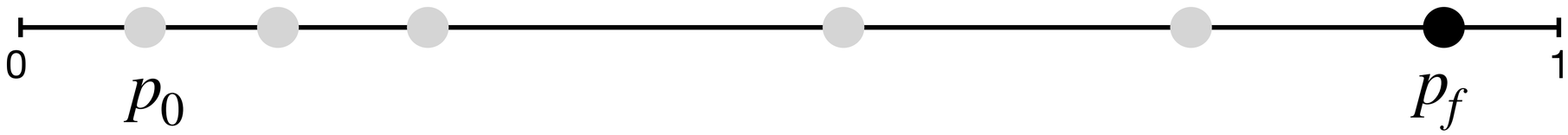
- Total progress is

$$\Pi_i r_i = \Pi_i \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

- One player must have progress

$$\geq \sqrt{\frac{p_f}{p_0}}$$

- Let N be the set of indices for non-biased player, then



Generic Attack with Large Advantage

- Total progress is

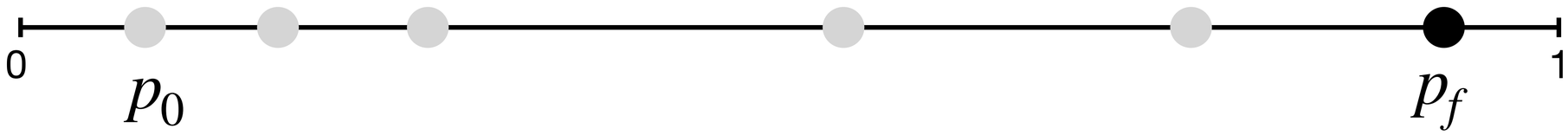
$$\Pi_i r_i = \Pi_i \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

- One player must have progress

$$\geq \sqrt{\frac{p_f}{p_0}}$$

- Let N be the set of indices for non-biased player, then

$$E[\Pi_{i \in N} r_i] = 1$$



Generic Attack with Large Advantage

- Total progress is

$$\Pi_i r_i = \Pi_i \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

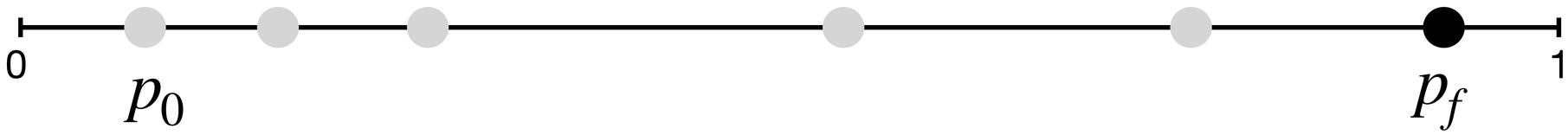
- One player must have progress

$$\geq \sqrt{\frac{p_f}{p_0}}$$

- Let N be the set of indices for non-biased player, then

$$E[\Pi_{i \in N} r_i] = 1$$

and by Markov the probability that $\Pi_{i \in N} r_i \geq \sqrt{\frac{p_f}{p_0}}$ is less than or equal to $\sqrt{\frac{p_0}{p_f}}$



Generic Attack with Large Advantage

- Total progress is

$$\Pi_{i\in[2c]} r_i = \Pi_{i\in[2c]} \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

- One player must have progress

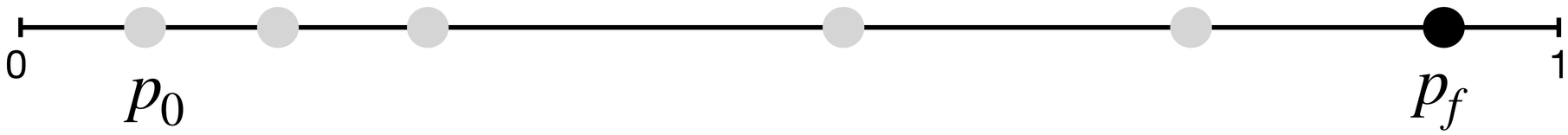
$$\geq \sqrt{\frac{p_f}{p_0}}$$

- Let N be the set of indices for non-biased player, then

$$E[\Pi_{i\in N} r_i] = 1$$

and by Markov

$$\Pr[\Pi_{i\in N} r_i \geq \sqrt{\frac{p_f}{p_0}}] \leq \sqrt{\frac{p_0}{p_f}}$$



Generic Attack with Large Advantage

- Total progress is

$$\Pi_{i\in[2c]} r_i = \Pi_{i\in[2c]} \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

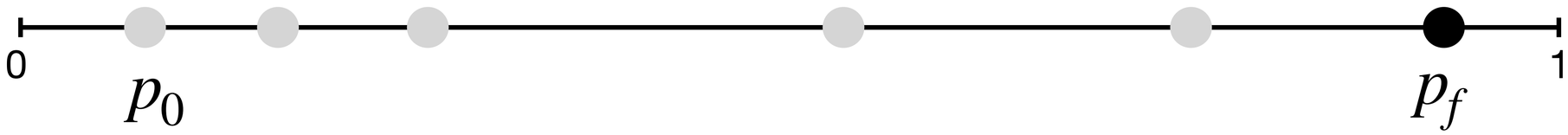
- One player must have progress

Strategy (high-level):

- Let N be the

and by Mark

$$\Pr[\Pi_{i\in N} r_i \geq \sqrt{\frac{p_f}{p_0}}] \leq \sqrt{\frac{p_0}{p_f}}$$



Generic Attack with Large Advantage

- Total progress is

$$\Pi_{i\in[2c]} r_i = \Pi_{i\in[2c]} \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

- One player must have progress

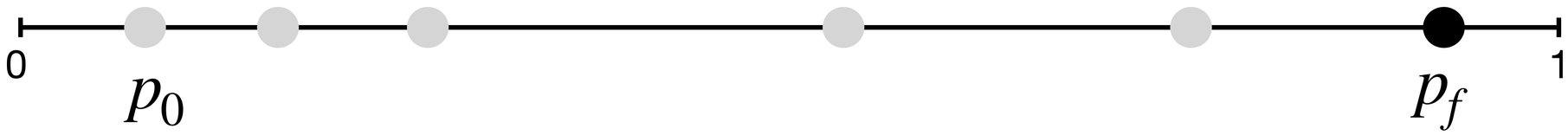
Strategy (high-level):

1. **Estimate each p_i**

- Let N be the

and by Mark

$$\Pr[\Pi_{i\in N} r_i \geq \sqrt{\frac{p_f}{p_0}}] \leq \sqrt{\frac{p_0}{p_f}}$$



Generic Attack with Large Advantage

- Total progress is

$$\Pi_{i\in[2c]} r_i = \Pi_{i\in[2c]} \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

- One player must have progress

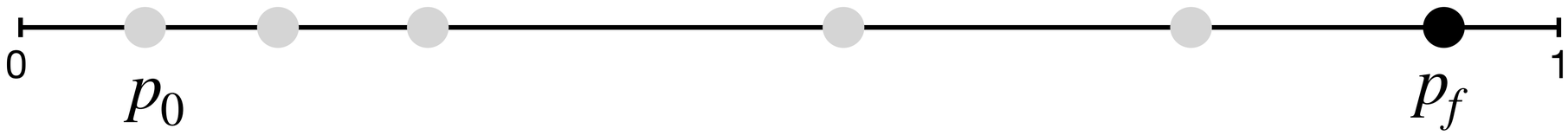
Strategy (high-level):

1. **Estimate each p_i**
2. **Compute contribution of each player**

- Let N be the

and by Mark

$$\Pr[\Pi_{i\in N} r_i \geq \sqrt{\frac{p_f}{p_0}}] \leq \sqrt{\frac{p_0}{p_f}}$$



Generic Attack with Large Advantage

- Total progress is

$$\Pi_{i\in[2c]} r_i = \Pi_{i\in[2c]} \frac{p_i}{p_{i-1}} = \frac{p_f}{p_0}$$

- One player must have progress

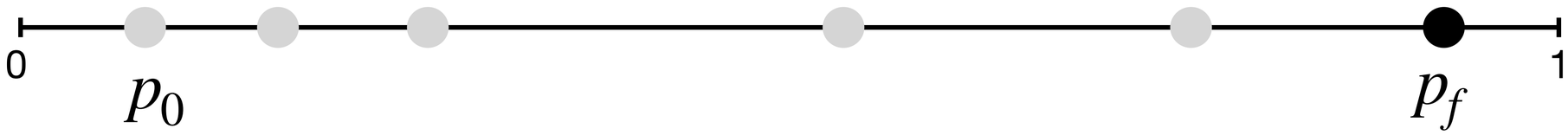
- Let N be the

and by Mark

Strategy (high-level):

1. Estimate each p_i
2. Compute contribution of each player
3. Declare biaser to be player with contribution $\geq \sqrt{\frac{p_f}{p_0}}$

$$\Pr[\Pi_{i\in N} r_i \geq \sqrt{\frac{p_f}{p_0}}] \leq \sqrt{\frac{p_0}{p_f}}$$



Summary

Summary

- [ACM22] has positive and negative results, but a large gap between them

Summary

- [ACM22] has positive and negative results, but a large gap between them
- Our work closes the gap by extending the negative results

Summary

- [ACM22] has positive and negative results, but a large gap between them
- Our work closes the gap by extending the negative results
  - Cannot get security against all poly adversaries with any non-trivial anonymity $\delta<1$

Summary

- [ACM22] has positive and negative results, but a large gap between them
- Our work closes the gap by extending the negative results
  - Cannot get security against all poly adversaries with any non-trivial anonymity $\delta<1$
  - Cannot get negligible anonymity even against fine-grained adversaries

Summary

- [ACM22] has positive and negative results, but a large gap between them
- Our work closes the gap by extending the negative results
  - Cannot get security against all poly adversaries with any non-trivial anonymity $\delta<1$
  - Cannot get negligible anonymity even against fine-grained adversaries
    - Their positive result is the best we can get

Open Questions

Open Questions

- [ACM22] feasibility result relies on ideal obfuscation

Open Questions

- [ACM22] feasibility result relies on ideal obfuscation
  - Construct under standard assumption

Open Questions

- [ACM22] feasibility result relies on ideal obfuscation
  - Construct under standard assumption
- Covert Cheating Attack runs in fairly large polynomial time

Open Questions

- [ACM22] feasibility result relies on ideal obfuscation
  - Construct under standard assumption
- Covert Cheating Attack runs in fairly large polynomial time
  - Improve the runtime of the attack

Thanks!