MECHANISM DESIGN FOR COLLABORATIVE NORMAL MEAN ESTIMATION

STANFORD RAIN SEMINAR, APRIL 15, 2024

KIRTHEVASAN KANDASAMY
UNIVERSITY OF WISCONSIN-MADISON
BASED ON JOINT WORK WITH: YIDING CHEN, ALEX CLINTON, AND JERRY ZHU

MACHINE LEARNING IS UBIQUITOUS

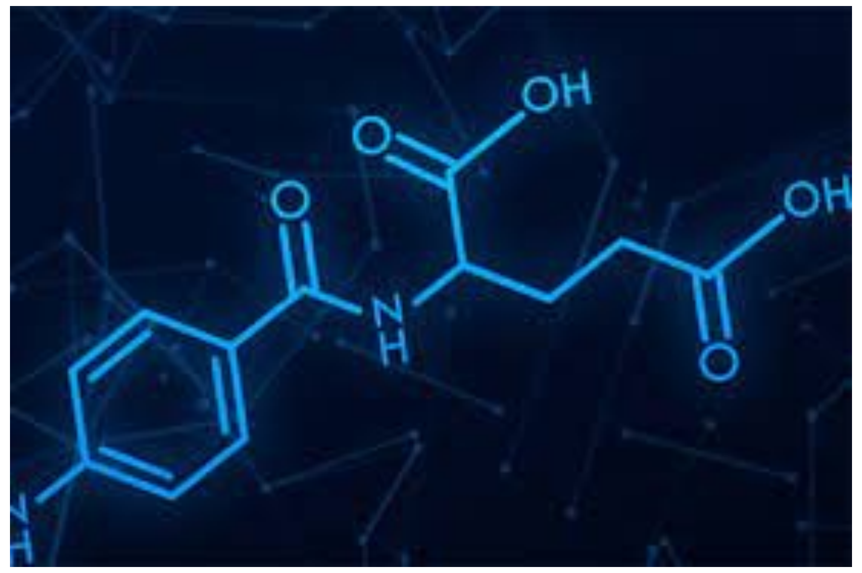
- Consumer facing businesses
- Industrial processes
- Scientific research
- Transport/logistics

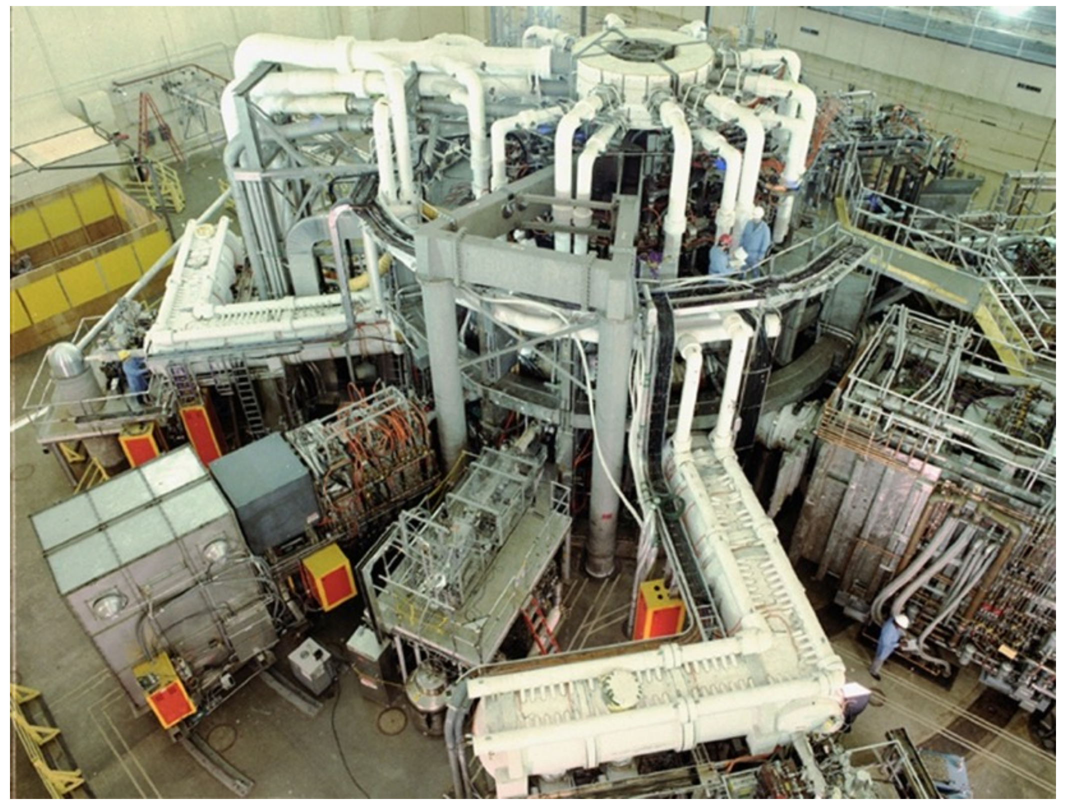
shopify

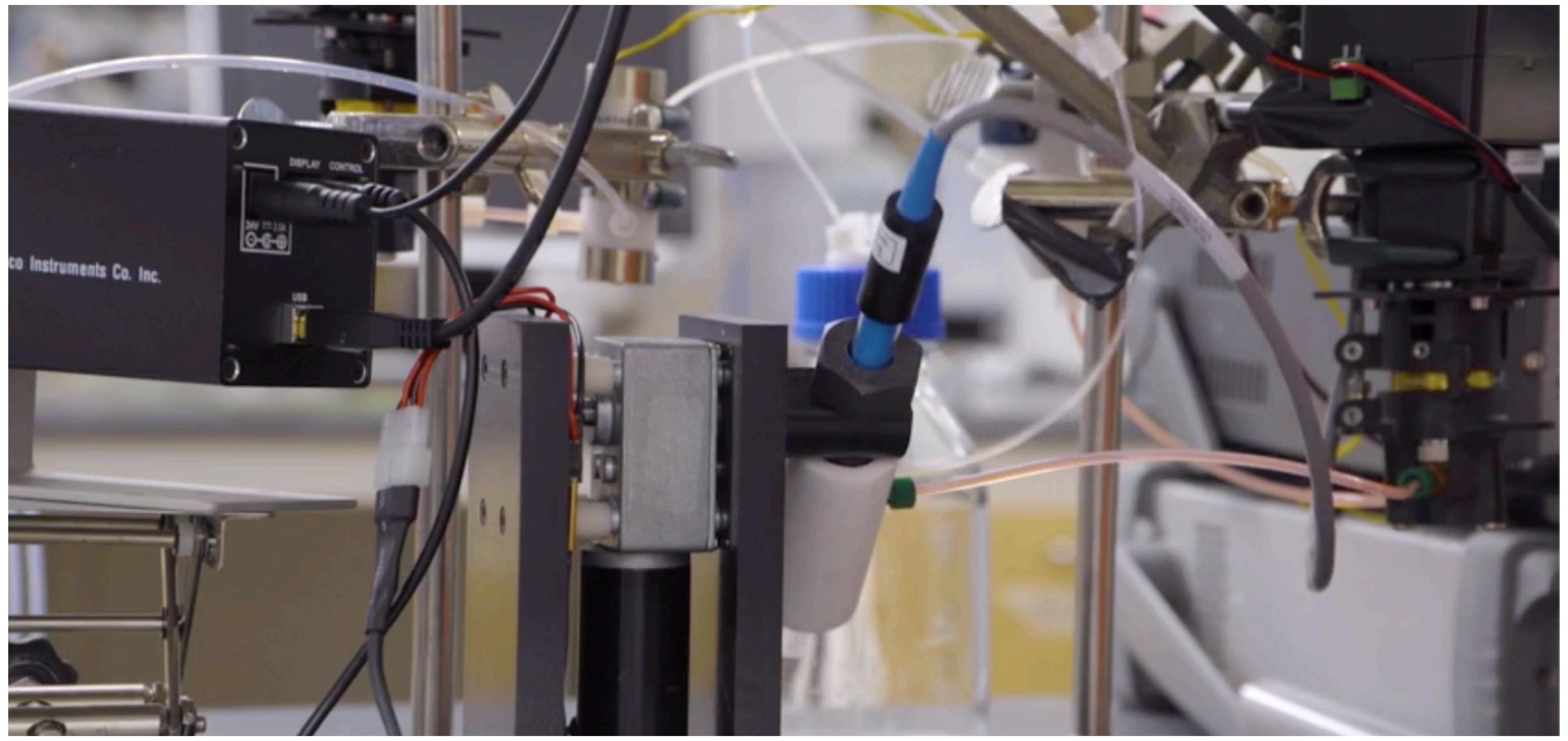
ebay

amazon

DOORDASH







DATA IS AN INVALUABLE RESOURCE

DATA IS AN INVALUABLE RESOURCE

- Data is the *new oil*.
- Data is the *new gold*.

The Economist, NY Times, Forbes, Wired, Deloitte, EY, Boston Consulting Group, and several more …

DATA IS AN INVALUABLE RESOURCE

- Data is the *new oil*.
- Data is the *new gold*.

The Economist, NY Times, Forbes, Wired, Deloitte, EY, Boston Consulting Group, and several more …

- But data is different to other types of resources
  - Data is ***costly*** to produce, but ***free*** to replicate.

A UTOPIAN GOAL

Everyone collects data, everyone shares their data with others.

- Cost incurred by one organization to produce data can benefit others.
- Better for the organizations, better for society at large.

SHARING WILL DEMOCRATIZE DATA

SHARING WILL DEMOCRATIZE DATA

Small organizations with little data:

A B C D E F

SHARING WILL DEMOCRATIZE DATA

Small organizations with little data:

A B C D E F

Large organization with lots of data:

SHARING WILL DEMOCRATIZE DATA

Small organizations with little data:

A B C D E F

Large organization with lots of data:

SHARING WILL DEMOCRATIZE DATA

Small organizations with little data:

A B C D E F

Large organization with lots of data:

By sharing data with each other, small organizations can compete with larger organizations.

OBSTACLES TO DATA SHARING

OBSTACLES TO DATA SHARING

Ethical/Legal

Privacy

Ownership of data

OBSTACLES TO DATA SHARING

Ethical/Legal

Privacy

Ownership of data

Security

Data breaches

Adversarial attacks

OBSTACLES TO DATA SHARING

Ethical/Legal	Security	Logistical
Privacy	Data breaches	Inter-operability
Ownership of data	Adversarial attacks	Communication costs

OBSTACLES TO DATA SHARING

Ethical/Legal

Privacy

Ownership of data

Security

Data breaches

Adversarial attacks

Logistical

Inter-operability

Communication costs

Incentives

Free-riding

Competition

Data monetization

Data valuation

OBSTACLES TO DATA SHARING

Ethical/Legal

Privacy

Ownership of data

Security

Data breaches

Adversarial attacks

Logistical

Inter-operability

Communication costs

Incentives

Free-riding

Competition

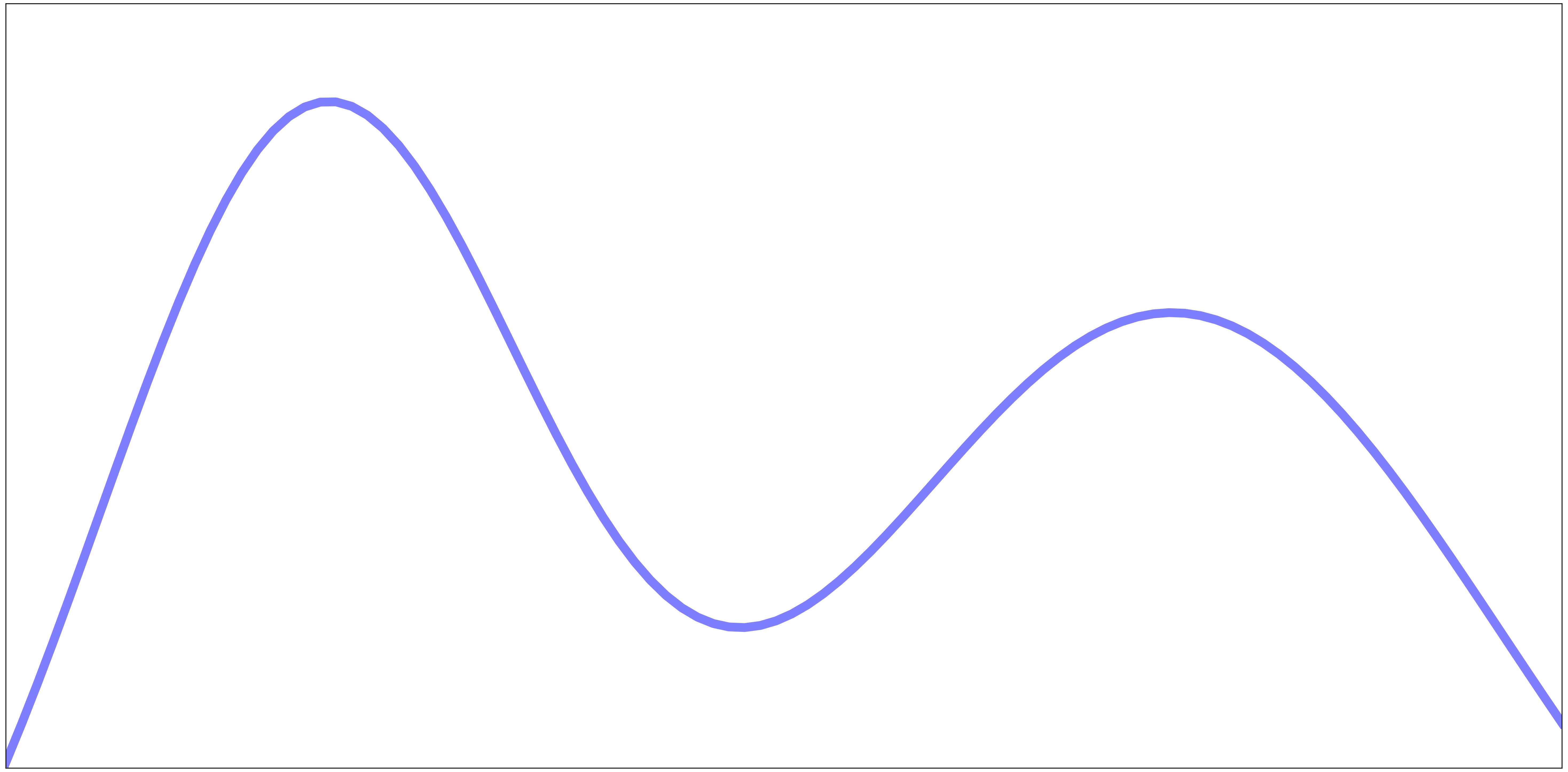
Data monetization

Data valuation

This talk → Free-riding

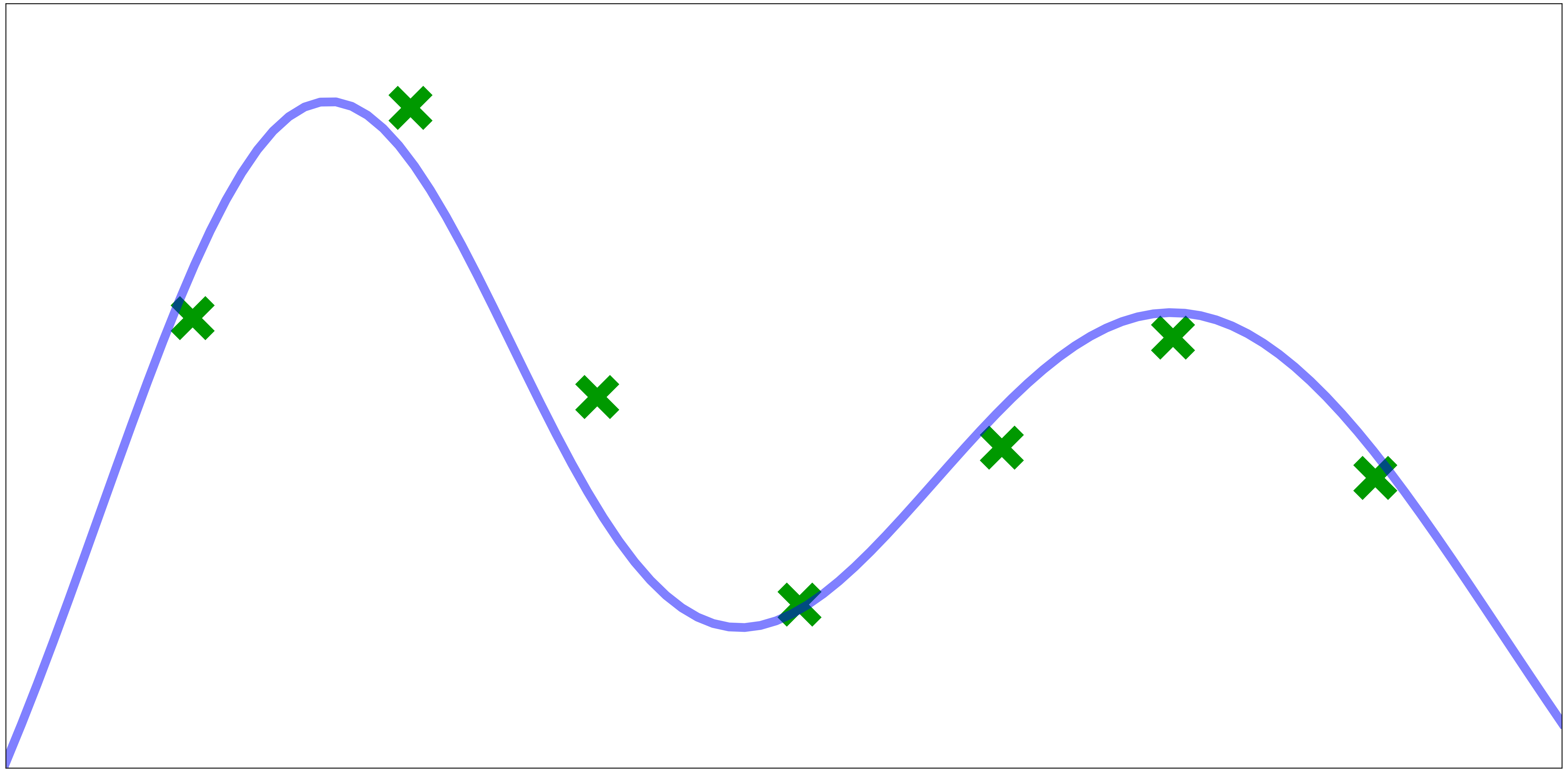
FREE-RIDING (#1/3)

agent's penalty = estimation error + cost of data collection



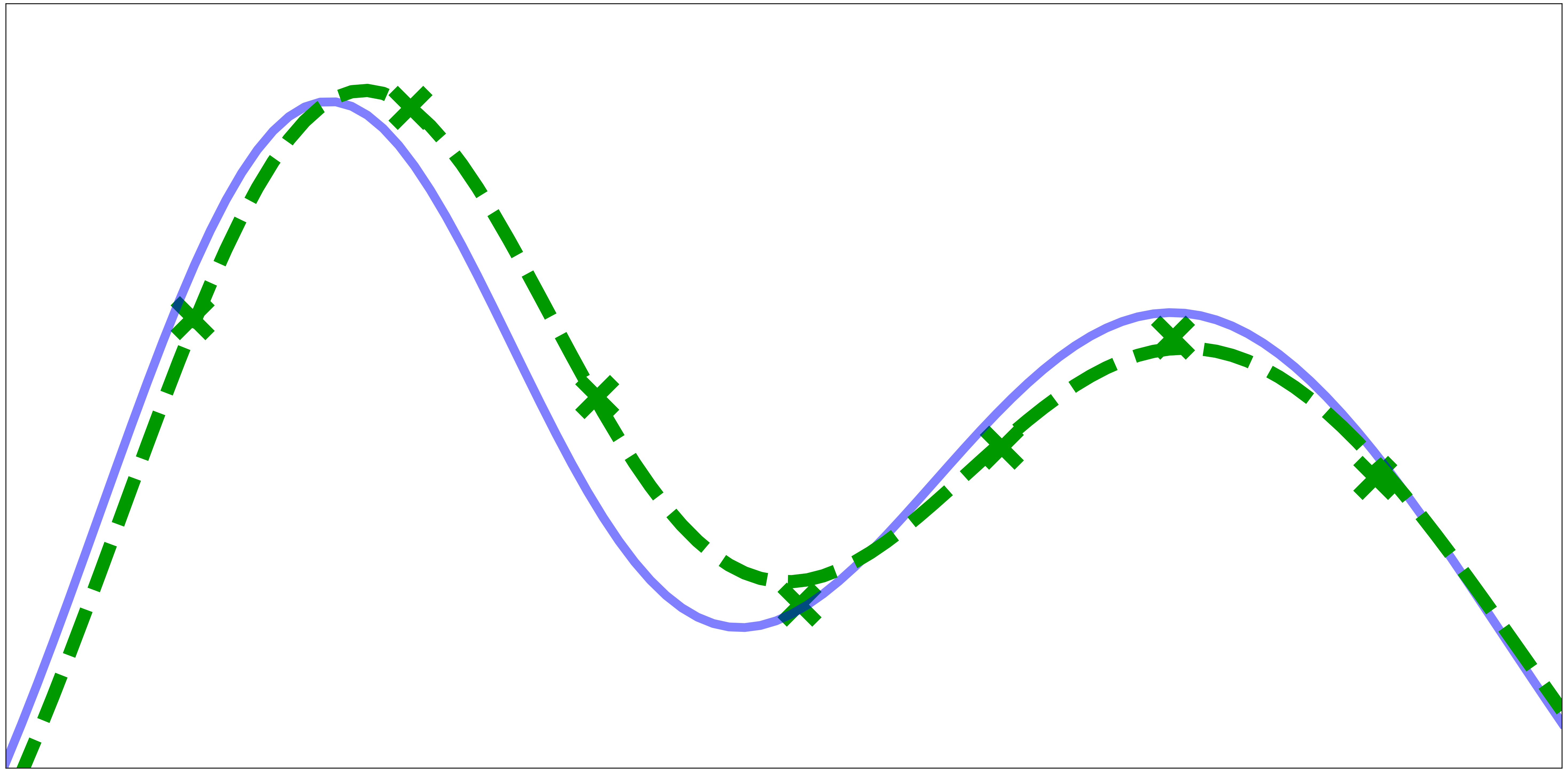
FREE-RIDING (#1/3)

agent's penalty = estimation error + cost of data collection



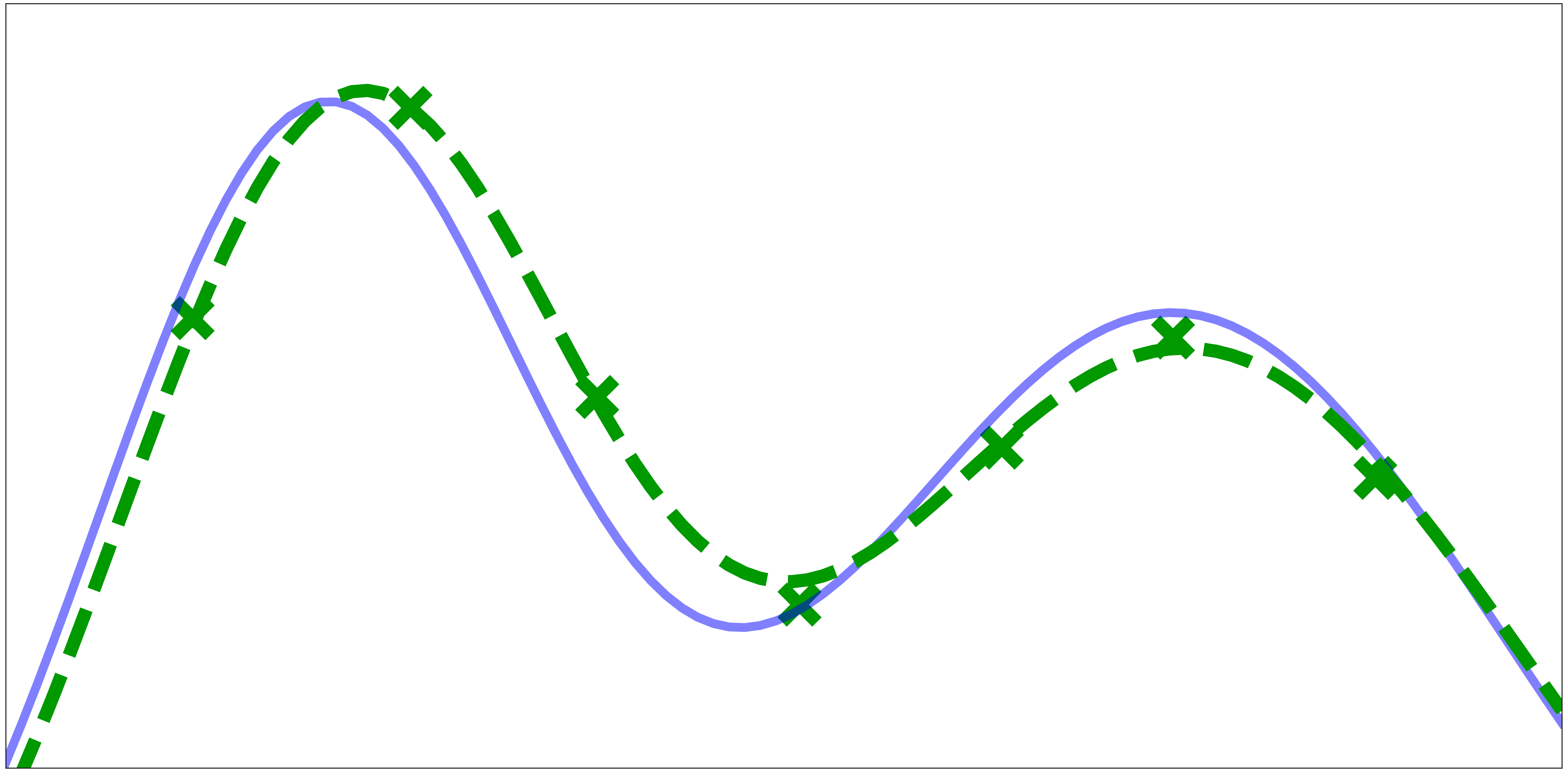
FREE-RIDING (#1/3)

agent's penalty = estimation error + cost of data collection



FREE-RIDING (#1/3)

agent's penalty = estimation error + cost of data collection

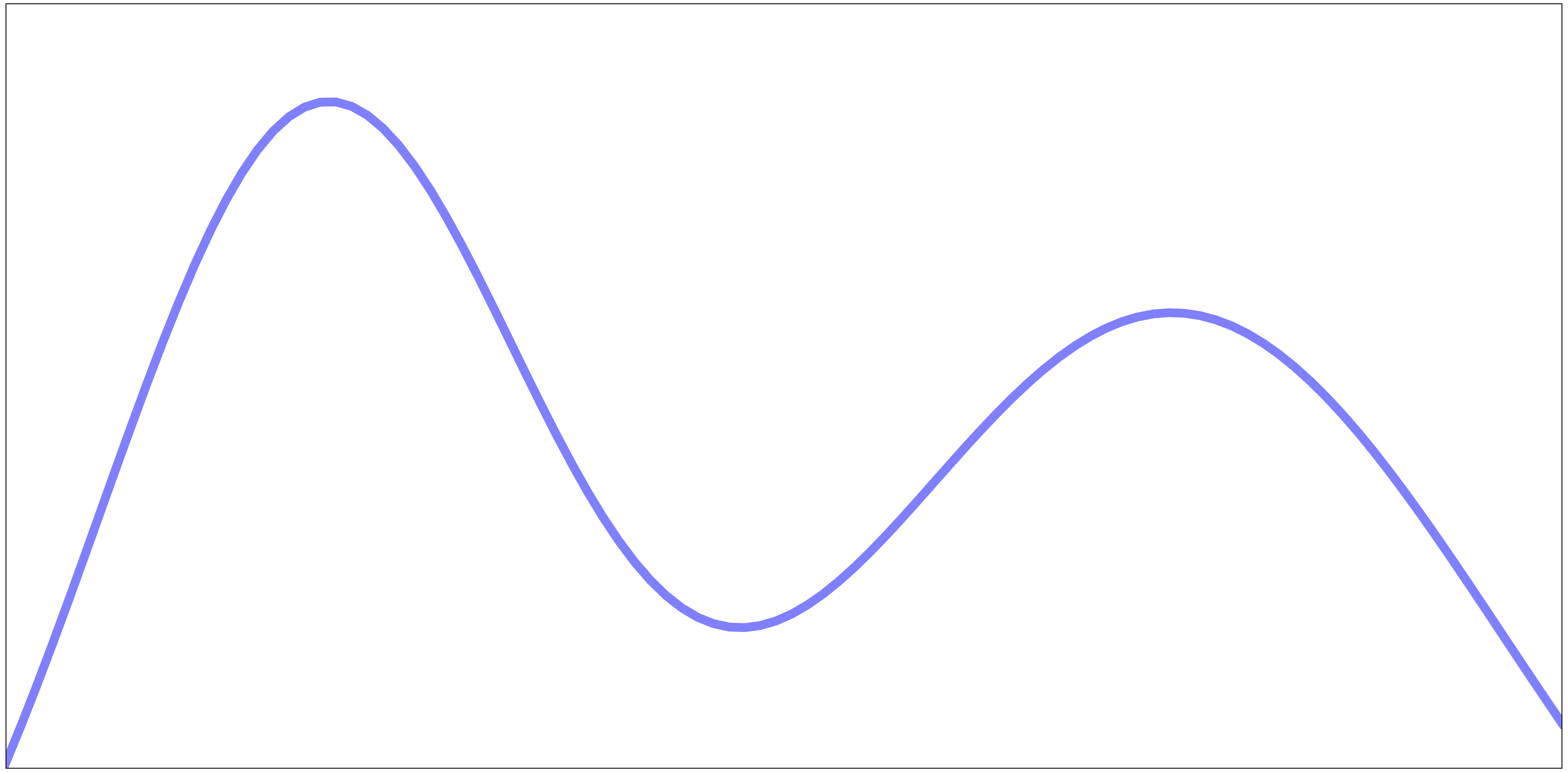


When **working on her own**, an agent will collect enough data until the cost offsets the (diminishing) increase in value from data.

FREE-RIDING (#2/3)

Multiple agents share data via a *naive* pool-and-share protocol:

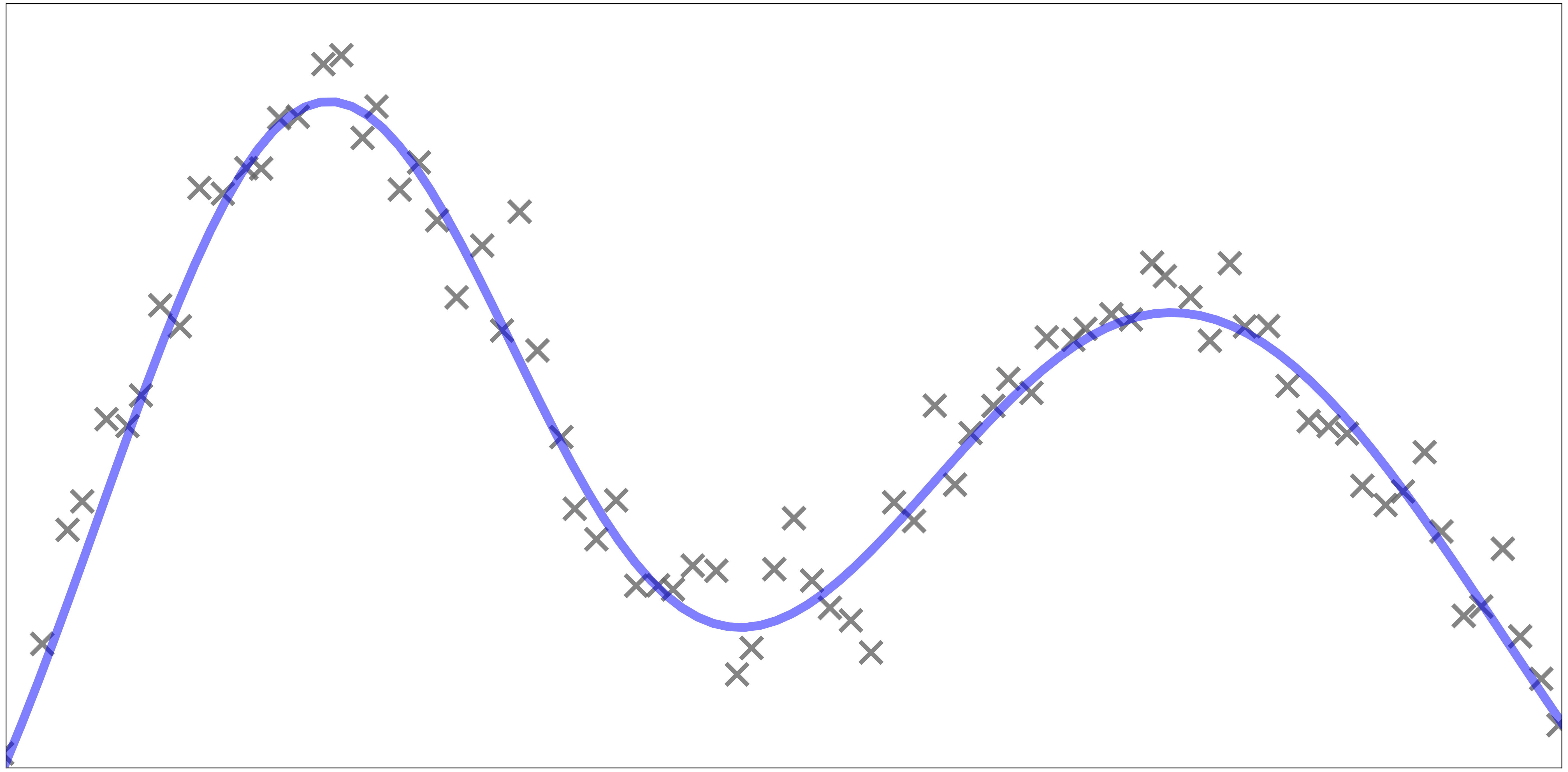
- Everyone collects data, everyone gets a copy of the others' data.



FREE-RIDING (#2/3)

Multiple agents share data via a *naive* pool-and-share protocol:

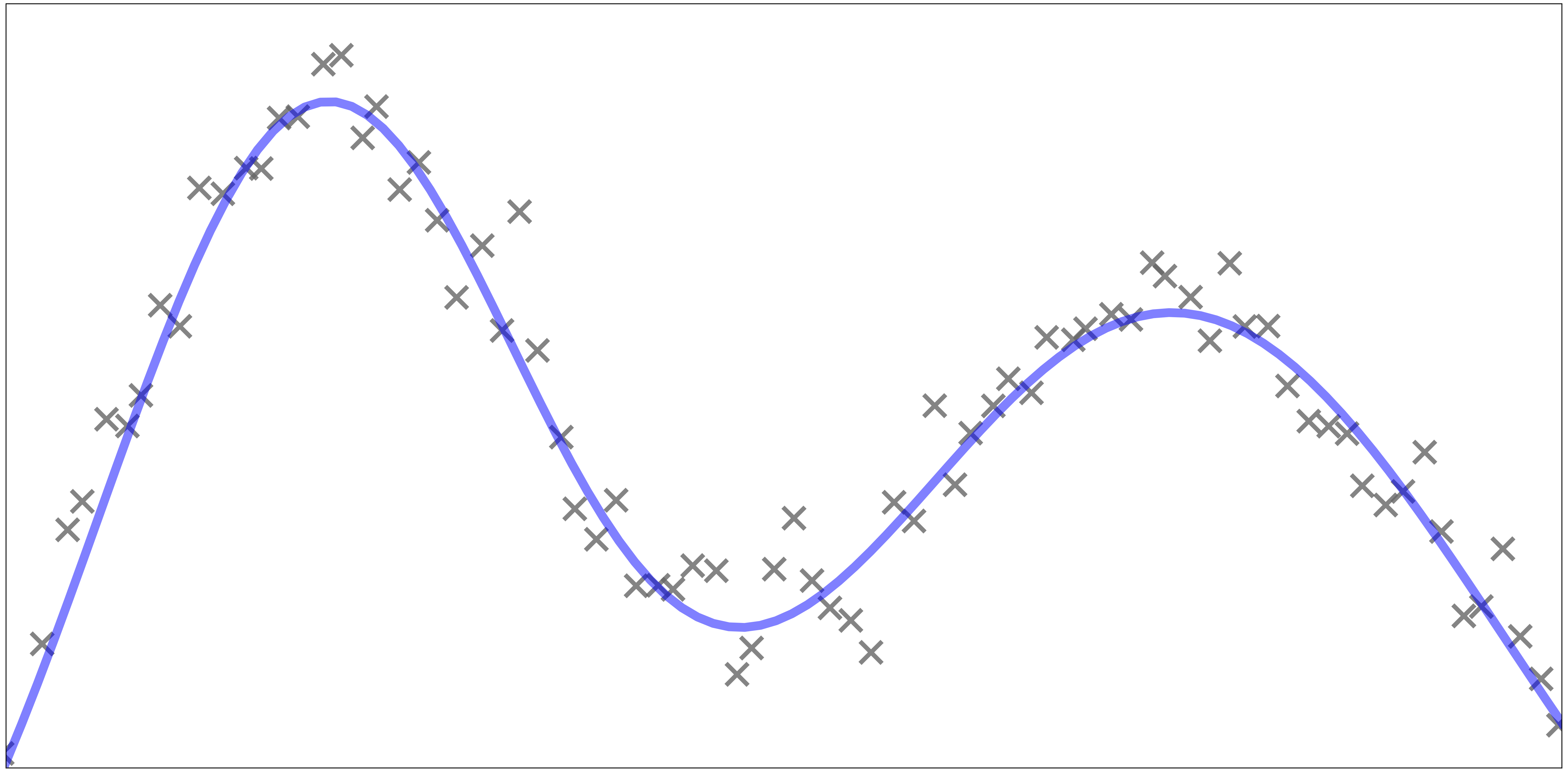
- Everyone collects data, everyone gets a copy of the others' data.



FREE-RIDING (#2/3)

Multiple agents share data via a *naive* pool-and-share protocol:

- Everyone collects data, everyone gets a copy of the others' data.

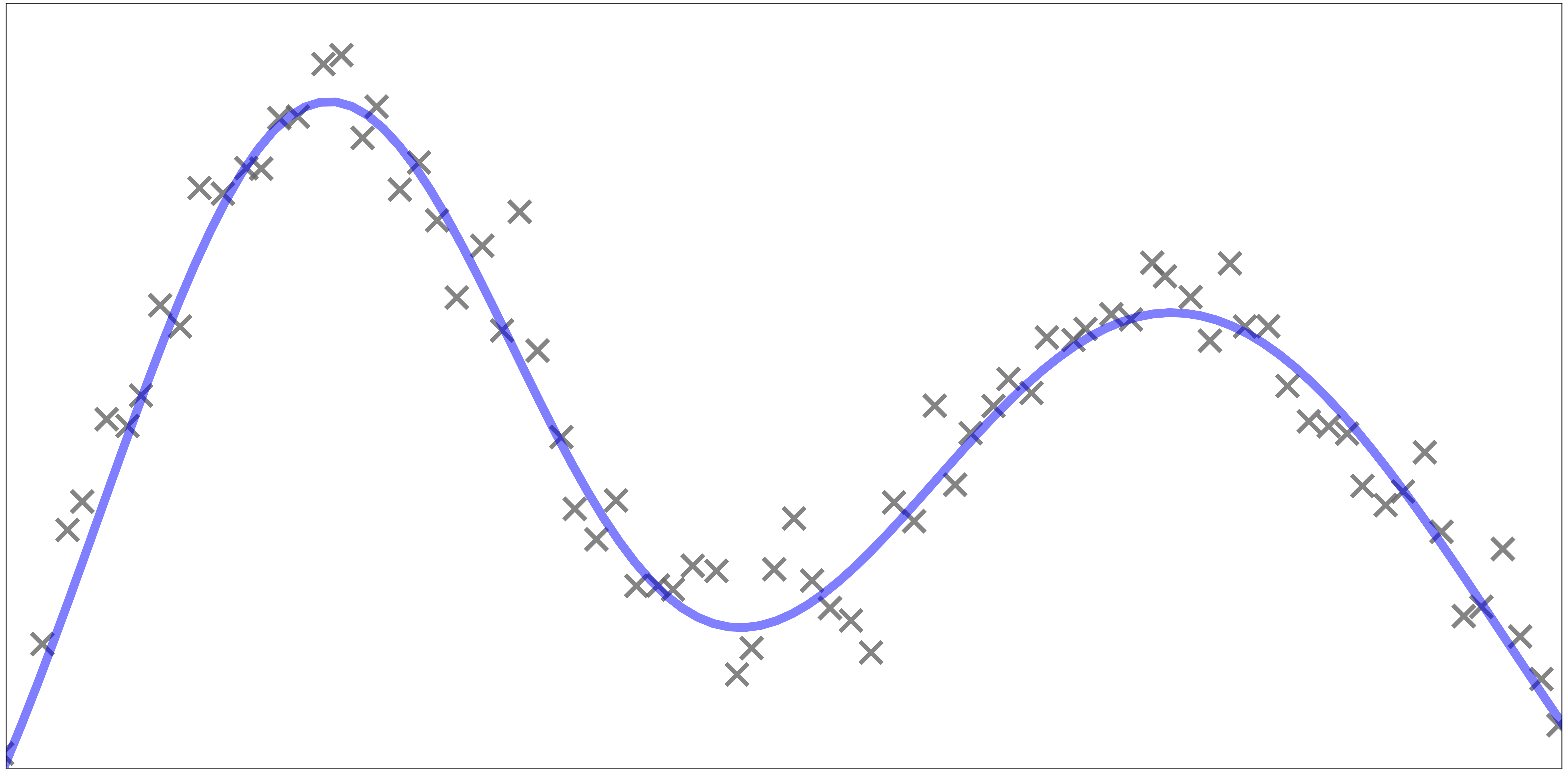


If others are already contributing large amounts of data, an agent has no incentive to collect/contribute data of her own.

FREE-RIDING (#3/3)

A seemingly plausible work-around (but does not work):

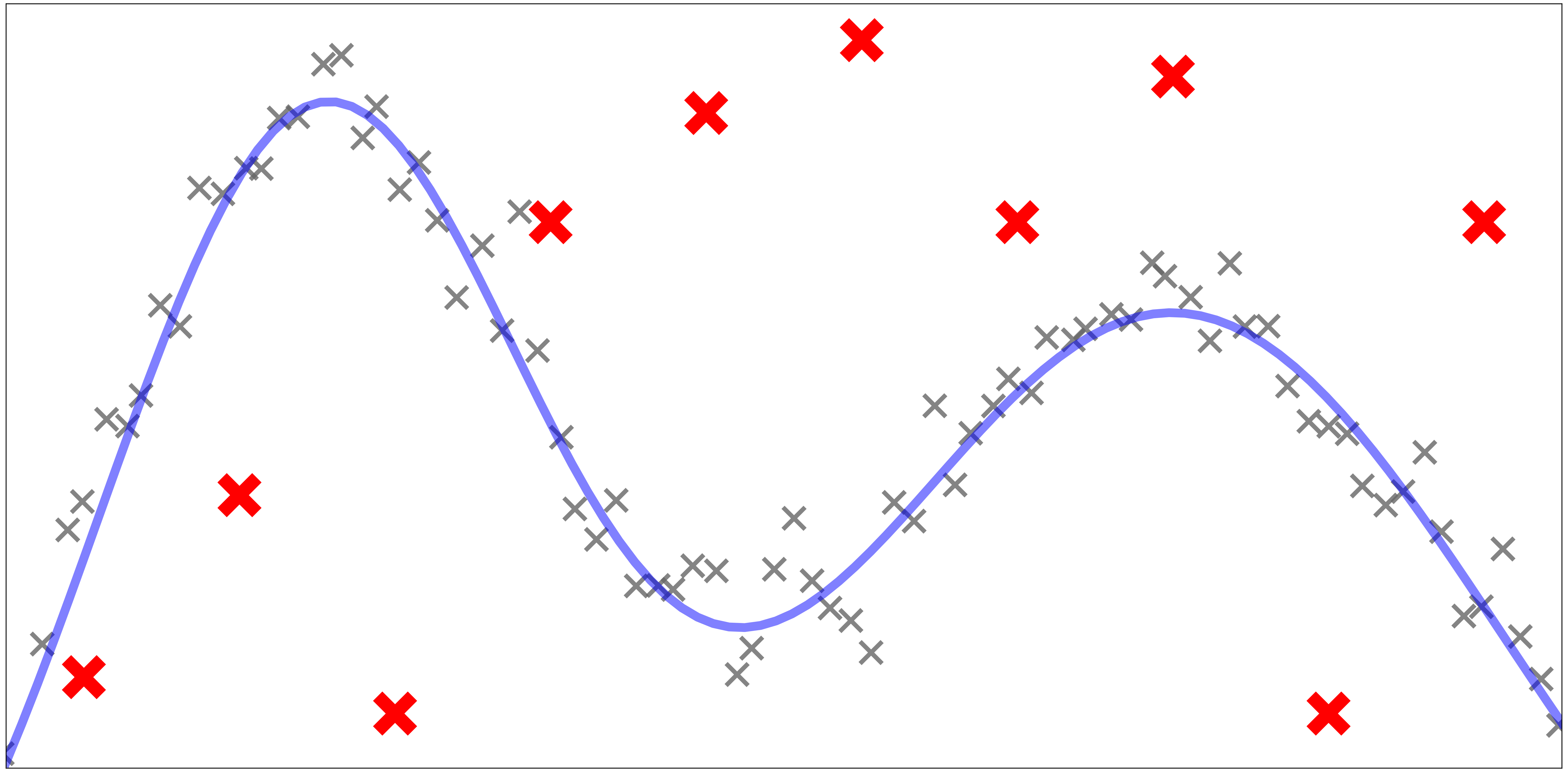
Pool-and-share but only if the agent contributes sufficient data



FREE-RIDING (#3/3)

A seemingly plausible work-around (but does not work):

Pool-and-share but only if the agent contributes sufficient data

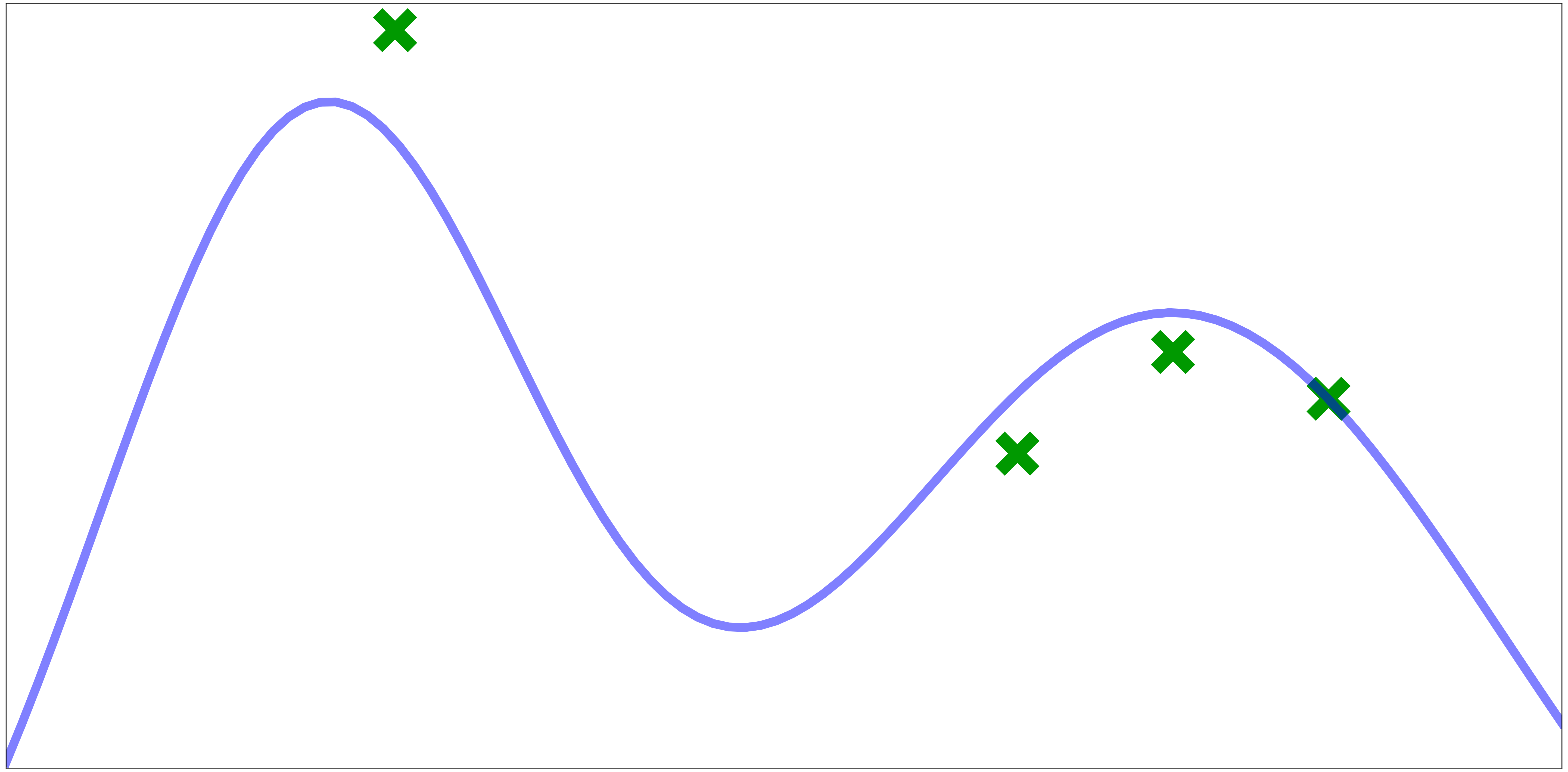


- Agent can submit fabricated data and then discard it when learning.

FREE-RIDING (#3/3)

A seemingly plausible work-around (but does not work):

Pool-and-share but only if the agent contributes sufficient data

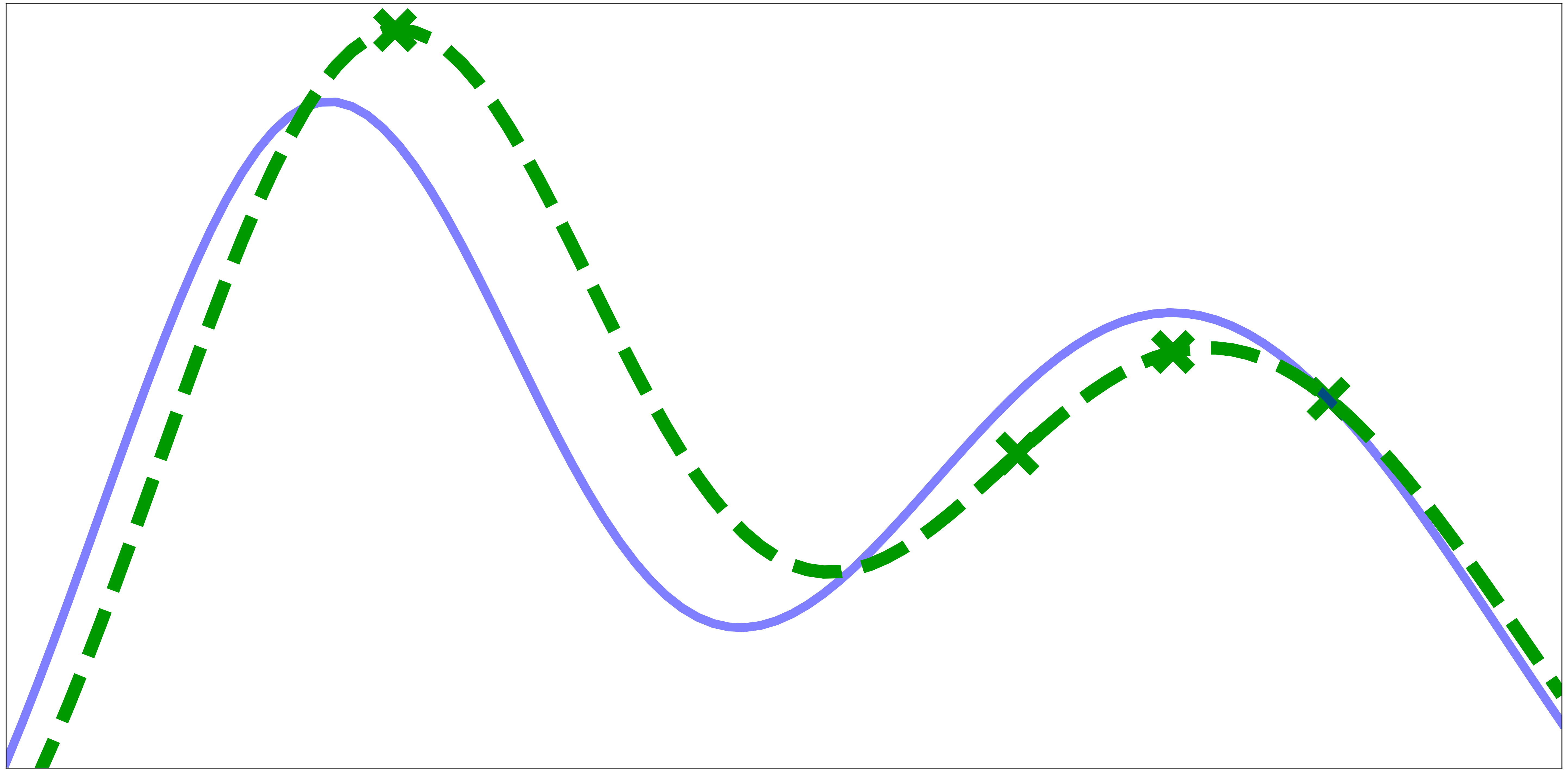


- Agent can submit fabricated data and then discard it when learning.

FREE-RIDING (#3/3)

A seemingly plausible work-around (but does not work):

Pool-and-share but only if the agent contributes sufficient data

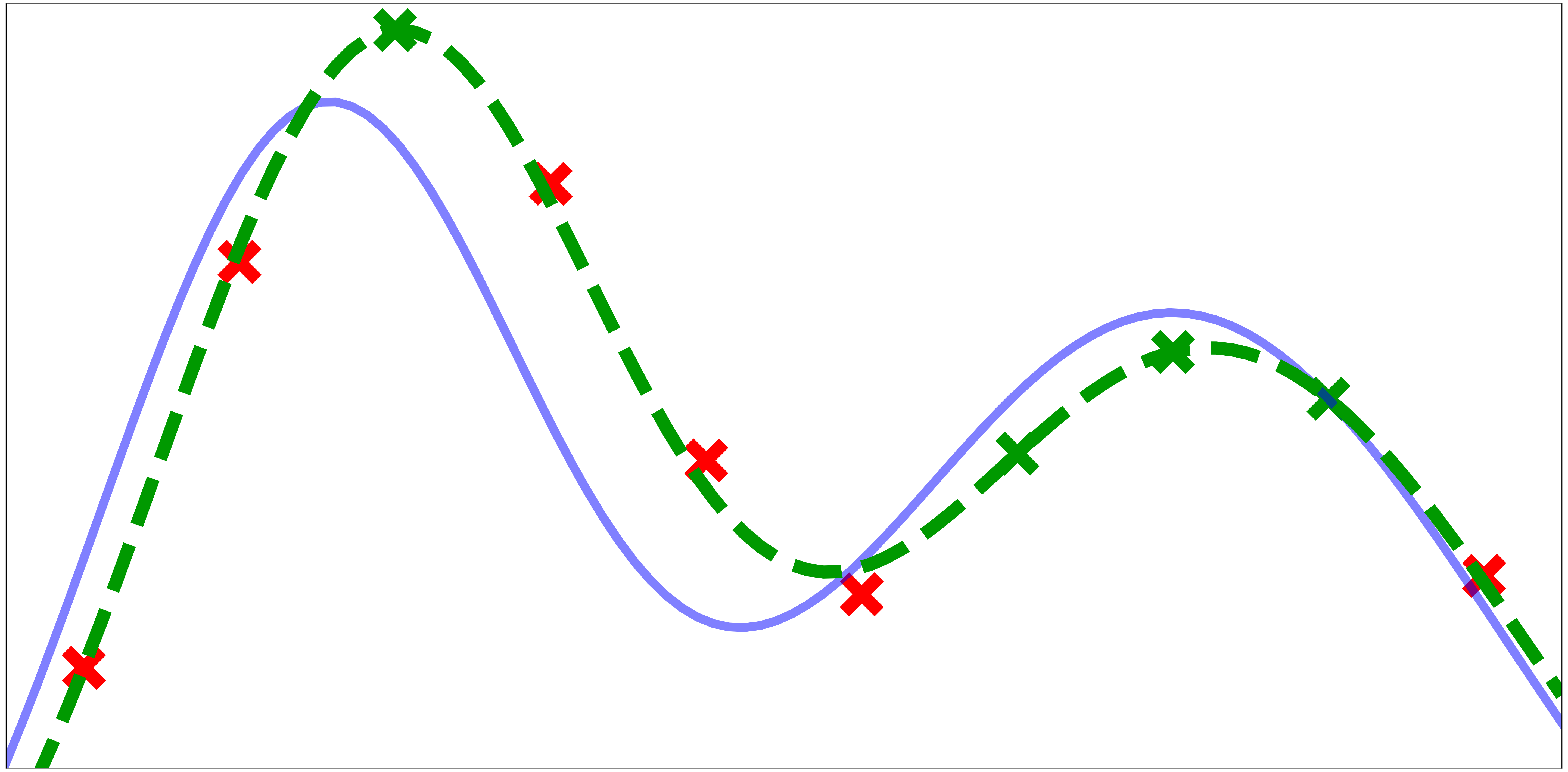


- Agent can submit fabricated data and then discard it when learning.

FREE-RIDING (#3/3)

A seemingly plausible work-around (but does not work):

Pool-and-share but only if the agent contributes sufficient data

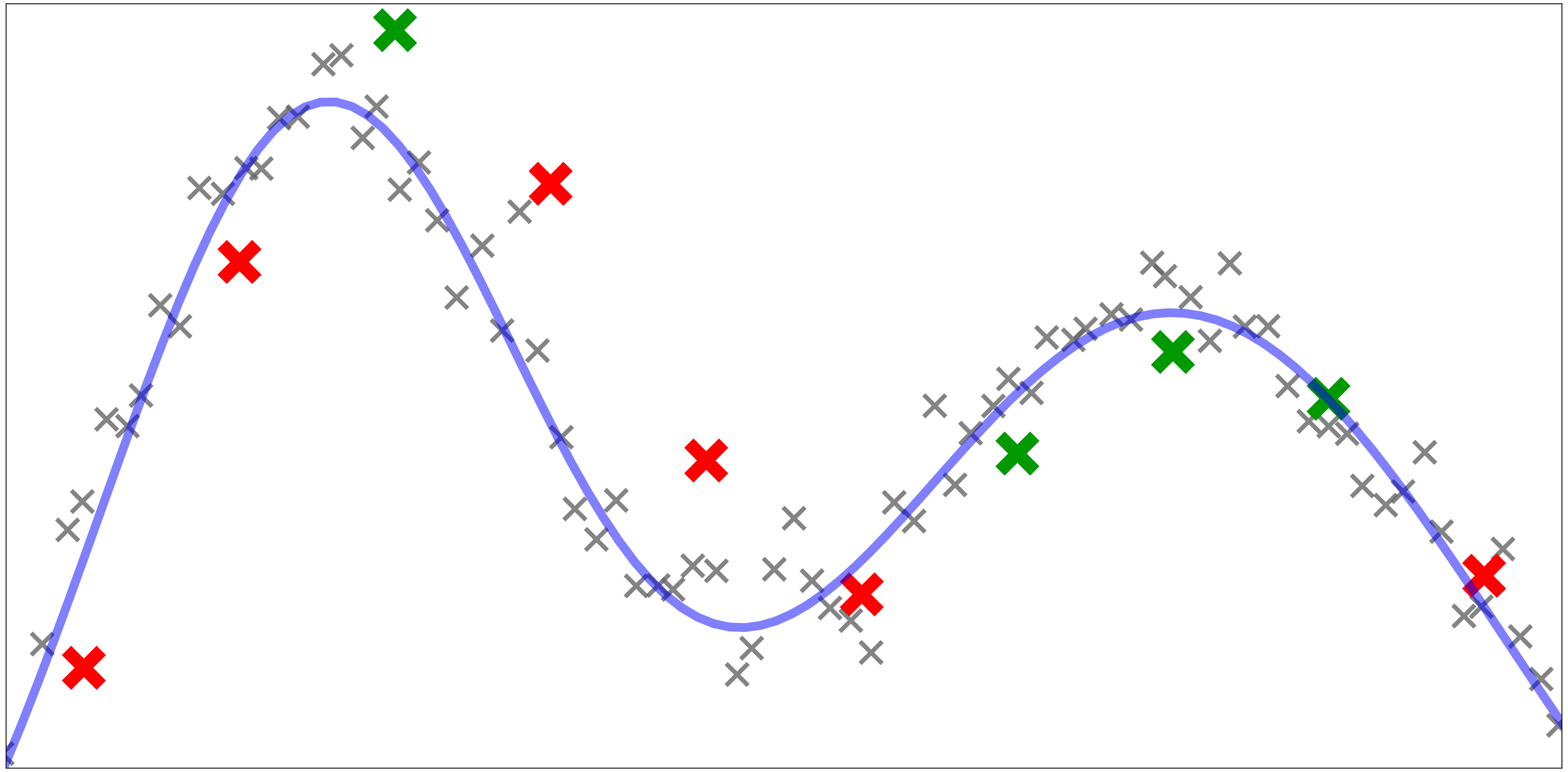


- Agent can submit fabricated data and then discard it when learning.
- Agent may fabricate based on a small sample she has collected, so it may not always be easy to detect.

FREE-RIDING (#3/3)

A seemingly plausible work-around (but does not work):

Pool-and-share but only if the agent contributes sufficient data



- Agent can submit fabricated data and then discard it when learning.
- Agent may fabricate based on a small sample she has collected, so it may not always be easy to detect.

OBSTACLES TO DATA SHARING

Ethical/Legal

Privacy

Ownership of data

Security

Data breaches

Adversarial attacks

Logistical

Inter-operability

Communication costs

Incentives

Free-riding

Competition

Data monetization

Data valuation

BUT THERE IS A DEMAND FOR DATA SHARING IN THE REAL WORLD

Data sharing platforms/consortia

PubChem

PubMed®

ChEMBL

DELTA SHARING

An open standard for secure data sharing

Marketplaces for data and ML models

aws AWS Data Exchange

Ads Data Hub

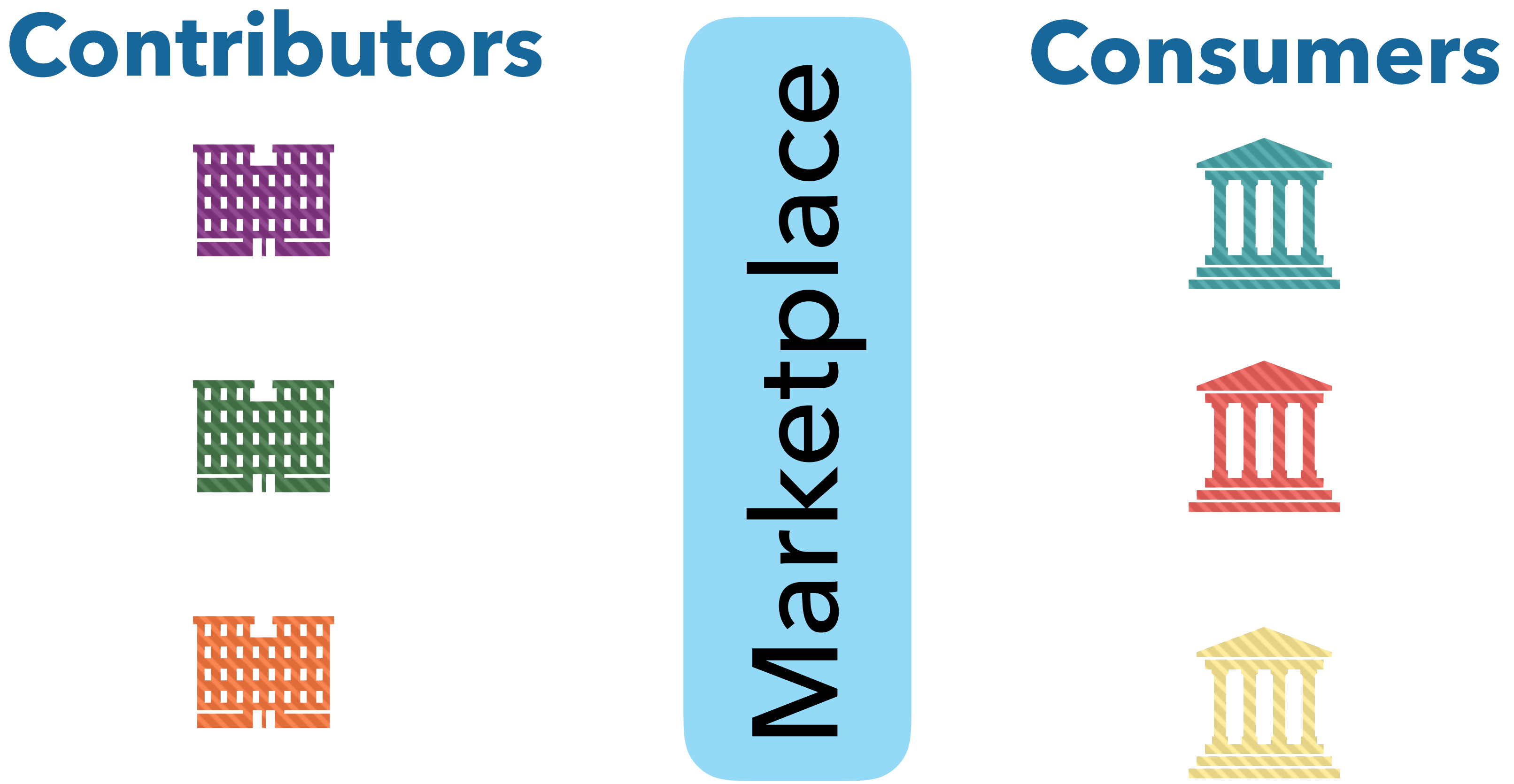
CITRINE INFORMATICS

Azure Data Share

OUR APPROACHES

Mechanisms for data sharing and federated learning

Data marketplaces

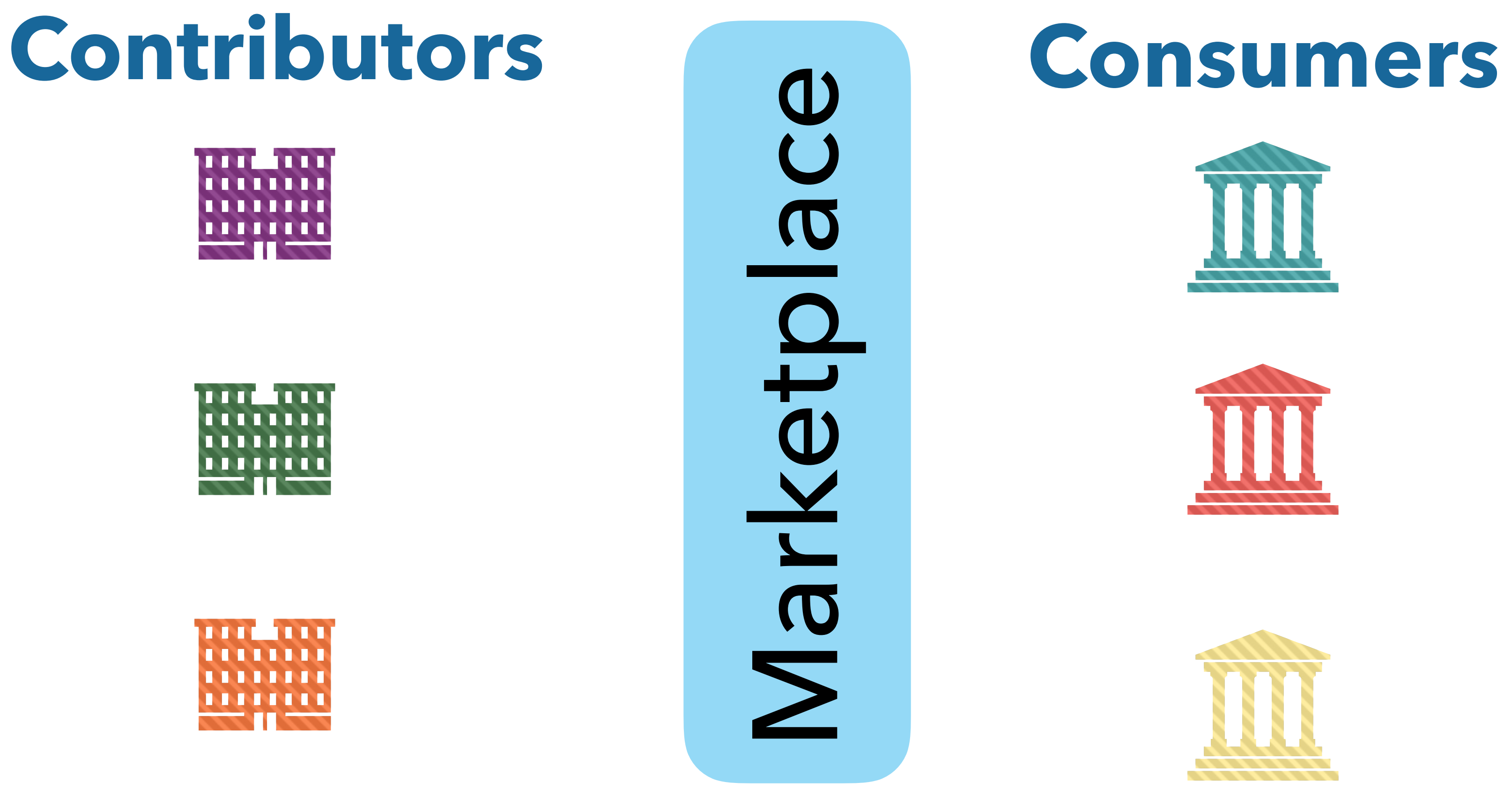


OUR APPROACHES

Mechanisms for data sharing and federated learning

Goal: Incentivize agents to collect as much data and share it honestly.

Data marketplaces

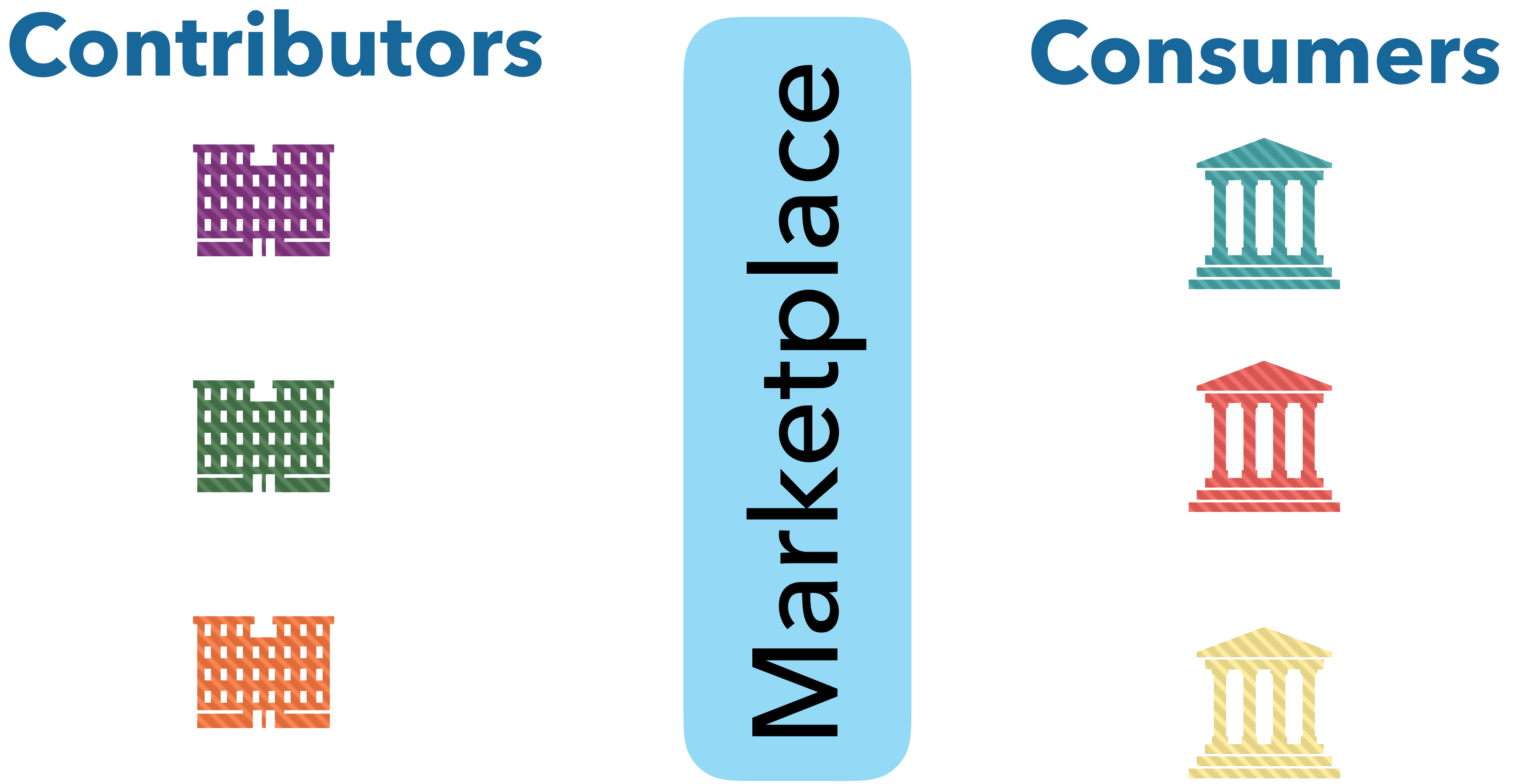


Mechanisms for data sharing and federated learning

Goal: Incentivize agents to collect as much data and <u>share it honestly</u>.

- Do not simply pool and share data!
- Cross-check for quality of the data contributed.

Data marketplaces

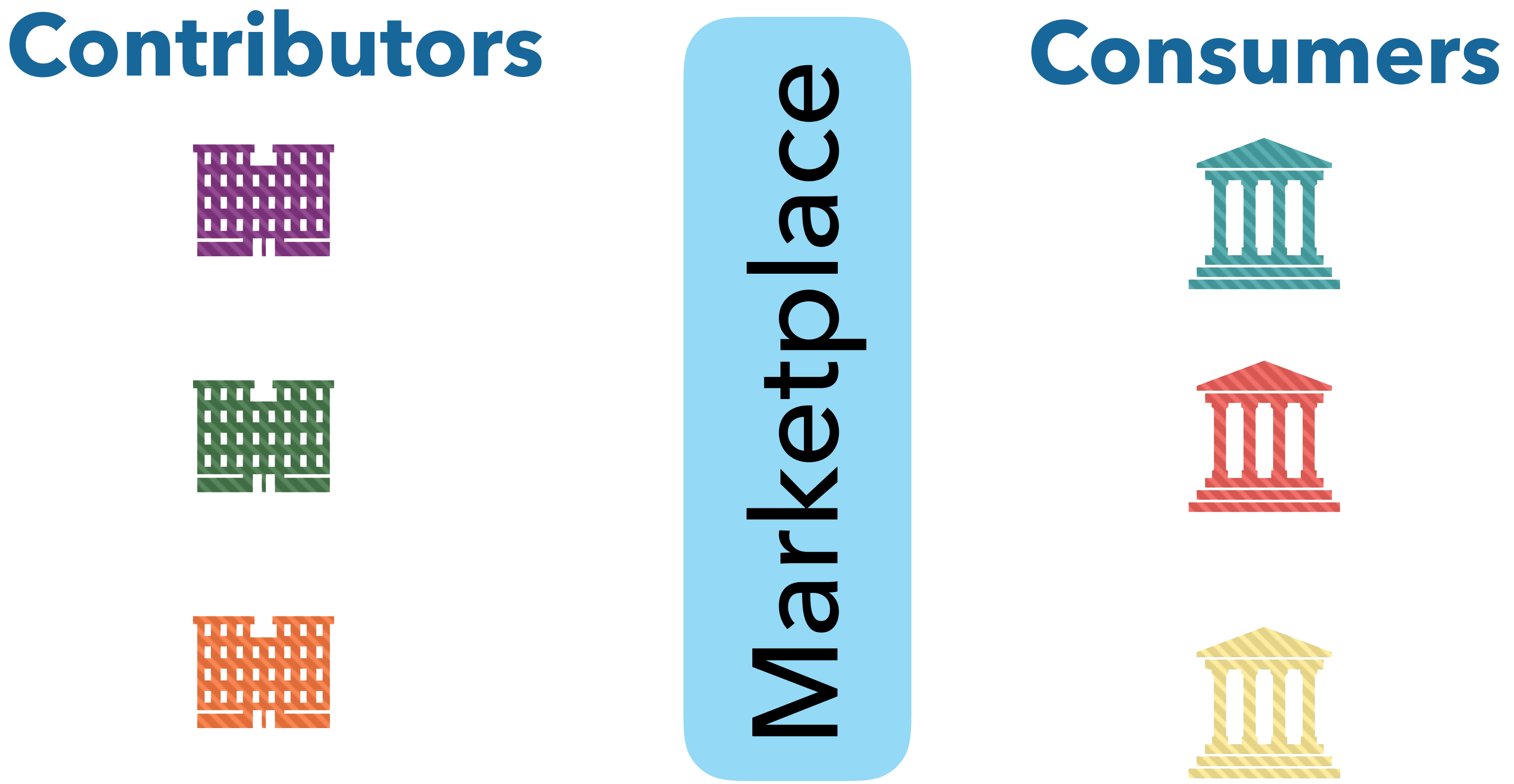


Mechanisms for data sharing and federated learning

Goal: Incentivize agents to collect as much data and share it honestly.

- Do not simply pool and share data!
- Cross-check for quality of the data contributed.
- More/better data contributed ⟹ more/better data received.

Data marketplaces

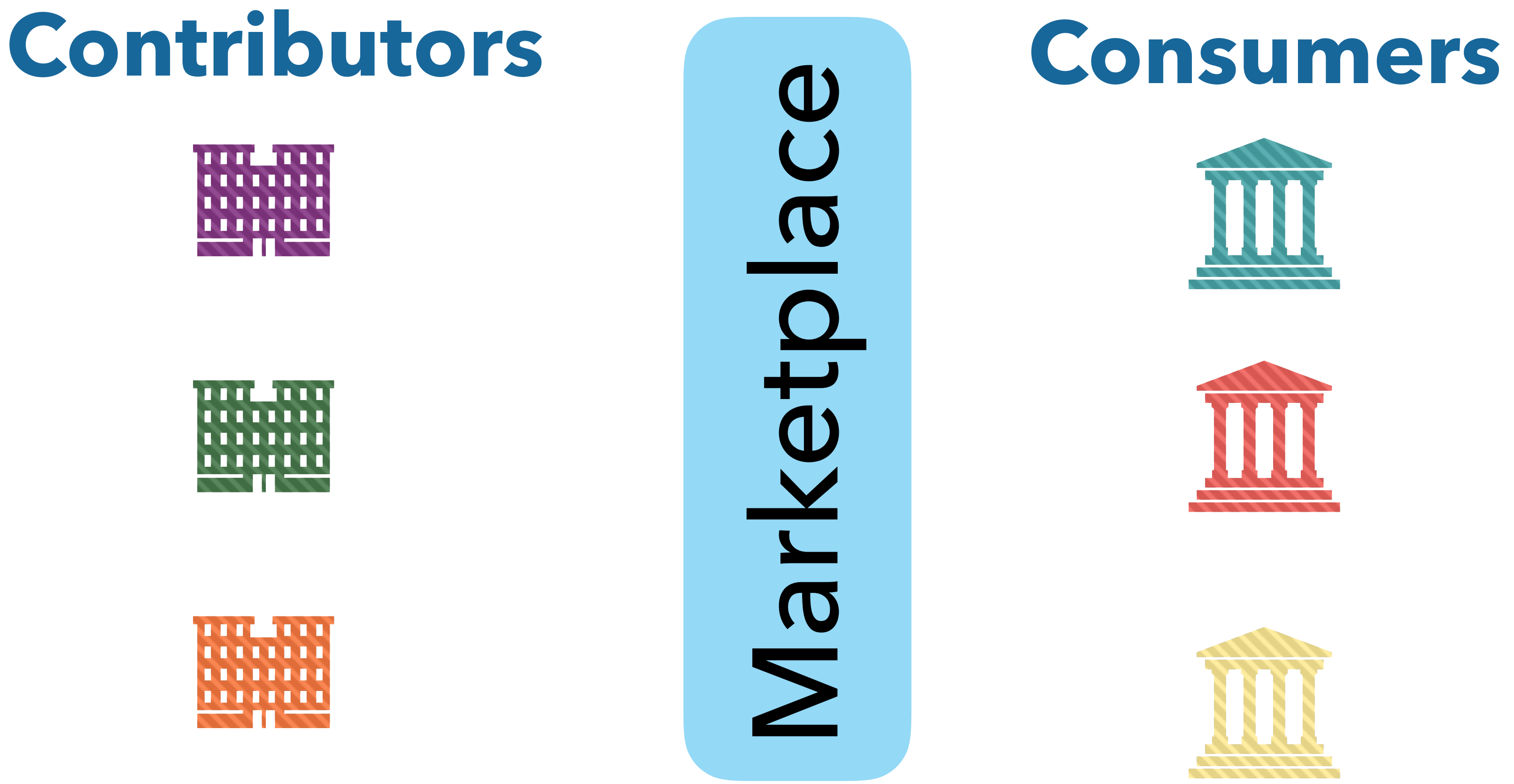


Mechanisms for data sharing and federated learning

Goal: Incentivize agents to collect as much data and share it honestly.

- Do not simply pool and share data!
- Cross-check for quality of the data contributed.
- More/better data contributed $\Longrightarrow$ more/better data received.

Data marketplaces



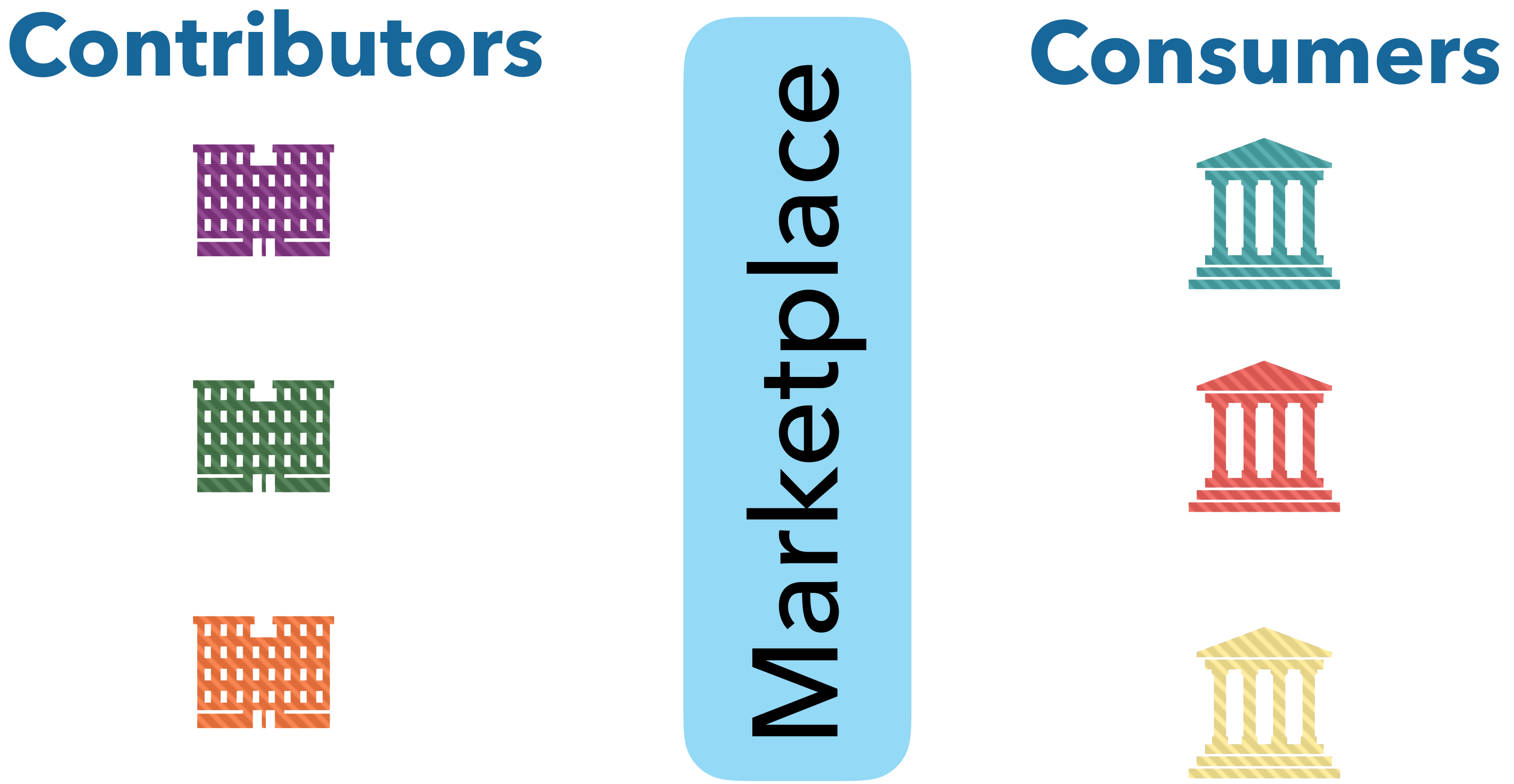
Goal: Incentivize contributors to honestly contribute lots of data. Fairly reward them for effort via payments from consumers.

Mechanisms for data sharing and federated learning

Goal: Incentivize agents to collect as much data and share it honestly.

- Do not simply pool and share data!
- Cross-check for quality of the data contributed.
- More/better data contributed $\Longrightarrow$ more/better data received.

Data marketplaces



Goal: Incentivize contributors to honestly contribute lots of data. Fairly reward them for effort via payments from consumers.

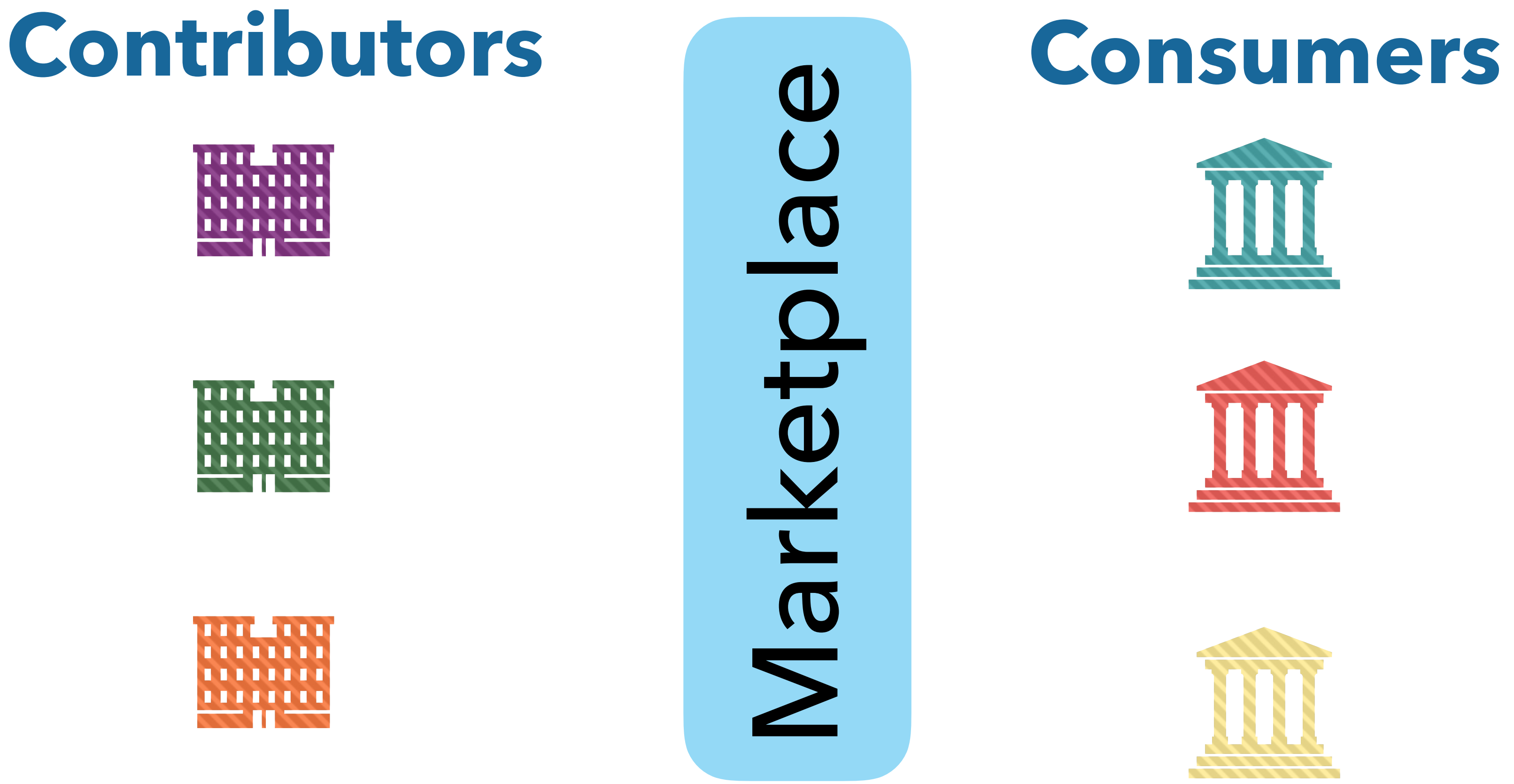
- A mediator checks for the quality of the data from contributors.

Mechanisms for data sharing and federated learning

Goal: Incentivize agents to collect as much data and share it honestly.

- Do not simply pool and share data!
- Cross-check for quality of the data contributed.
- More/better data contributed $\Longrightarrow$ more/better data received.

Data marketplaces



Goal: Incentivize contributors to honestly contribute lots of data. Fairly reward them for effort via payments from consumers.

- A mediator checks for the quality of the data from contributors.
- Higher quality data $\Longrightarrow$ higher revenue for data contributors.

PRIOR WORK

Mechanisms for data sharing and federated learning

Sim, Zhang, Chan, Low 2020
Xu, Lyu, Ma et al 2021
Blum, Haghtalab, Phillips, Shao 2021
Karimireddy, Guo, Jordan 2022
Fraboni, Vidal, Lorenzi 2021
Lin, Du, Liu 2019
Ding, Fang, Huang 2020
Liu, Tian, Chen et al 2022

Data marketplaces

Cai, Daskalakis, Papadimitriou 2015
Agarwal, Dahleh, Sarkar, 2019
Agarwal, Dahleh, Horel, Rui, 2020
Jia, Dao, Wang et al, 2019
Wang, Rausch, Zhang et al 2020

Key difference:

- All these works assume agents will always truthfully submit the data they have, i.e without fabrication/alteration.

OUTLINE

1. **Mechanism design for collaborative normal mean estimation**

   (Chen, Zhu, Kandasamy, *NeurIPS 2023*)

   - **Intuitions, overview of results**
   - **Problem formalism**
   - **Mechanism and theoretical analysis**

2. **Extensions** **(Clinton, Chen, Zhu, Kandasamy, *Ongoing work*)**

   - **Multiple distributions with asymmetric data collection capabilities**
   - **Collaborative supervised learning and experiment design**

OUTLINE

1. **Mechanism design for collaborative normal mean estimation**

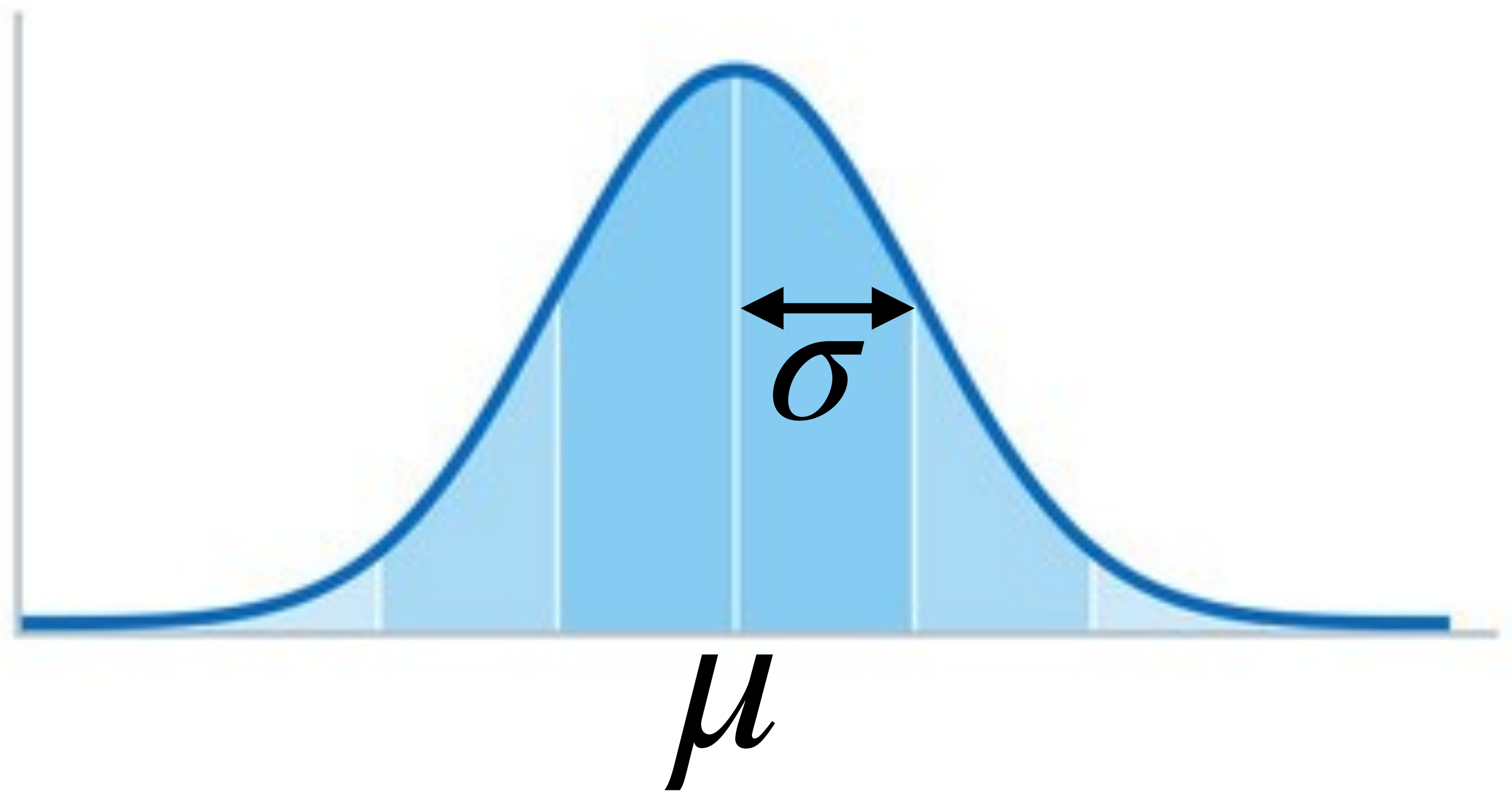
   (Chen, Zhu, Kandasamy, *NeurIPS 2023*)

   - **Intuitions, overview of results**
   - Problem formalism
   - Mechanism and theoretical analysis

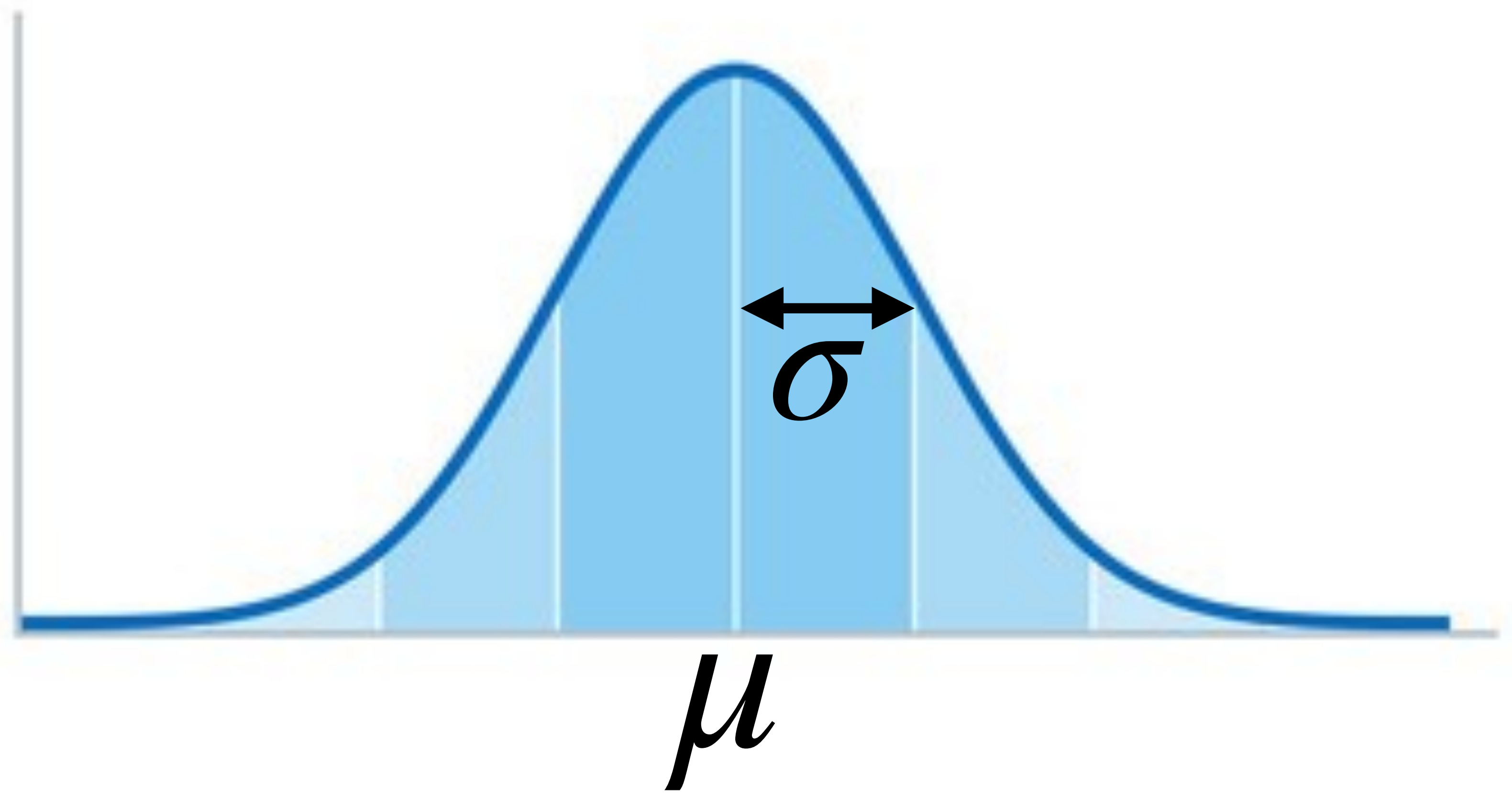
2. Extensions (Clinton, Chen, Zhu, Kandasamy, *Ongoing work*)

   - Multiple distributions with asymmetric data collection capabilities
   - Collaborative supervised learning and experiment design

NORMAL MEAN ESTIMATION



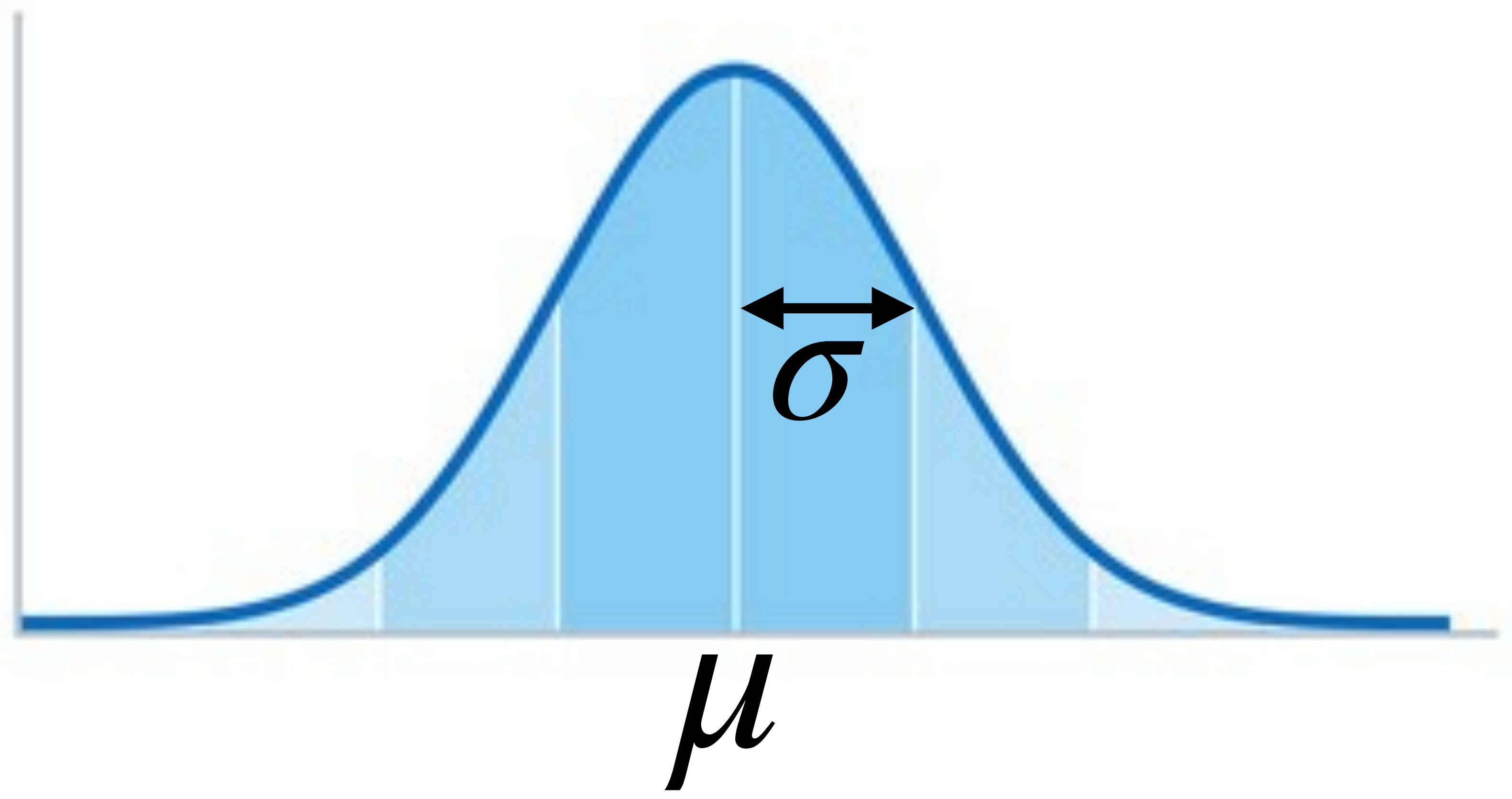
- Estimate the mean μ of a normal distribution with *known* variance σ^2.



NORMAL MEAN ESTIMATION

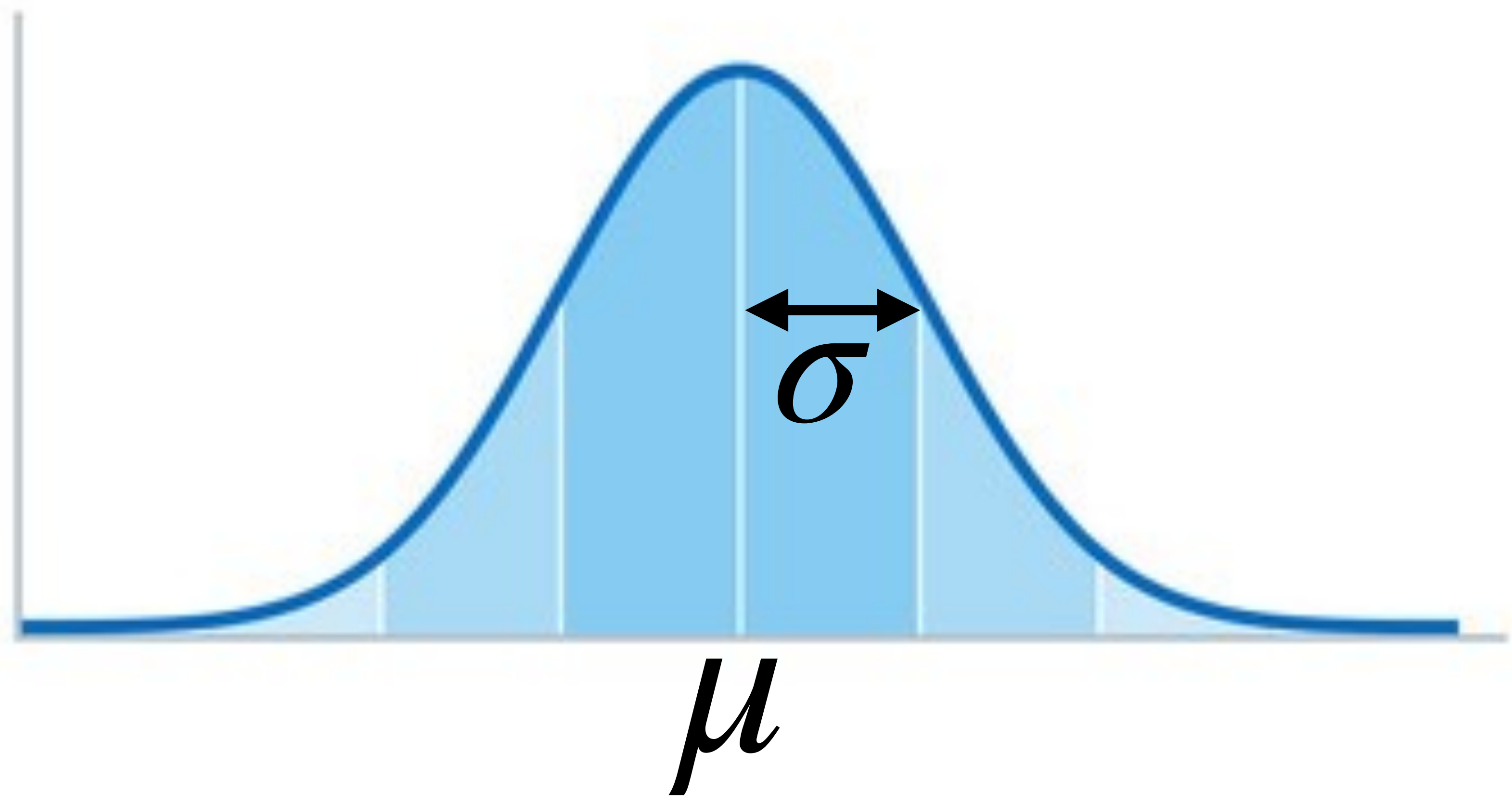
- Estimate the mean μ of a normal distribution with *known* variance σ^2.

- An agent can collect samples at *known* unit cost c.



NORMAL MEAN ESTIMATION

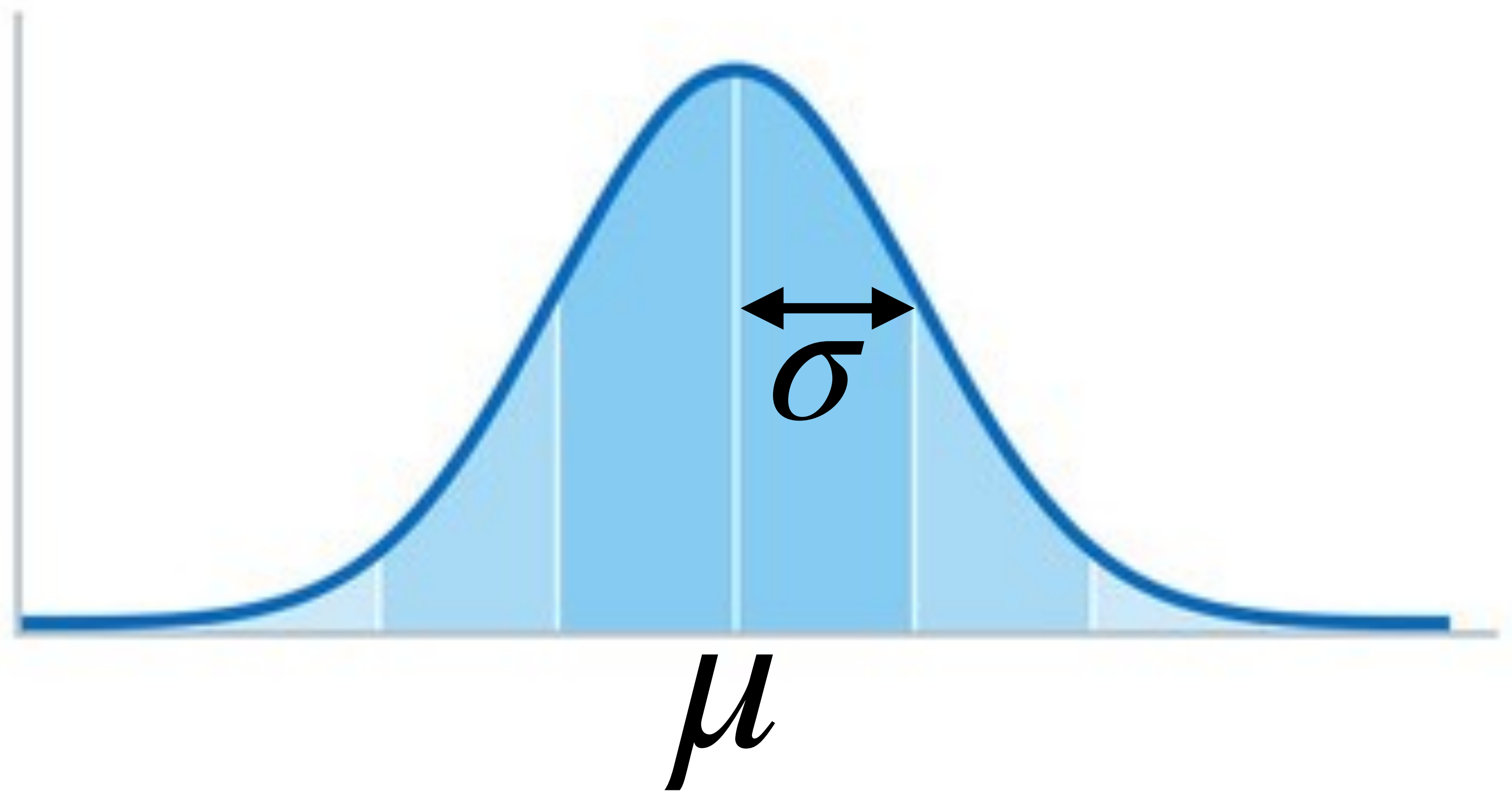
- Estimate the mean μ of a normal distribution with *known* variance σ^2.
- An agent can collect samples at *known* unit cost c.
- Each agent wishes to minimize



$$\text{penalty} = \text{estimation error} + \text{data collection cost}$$

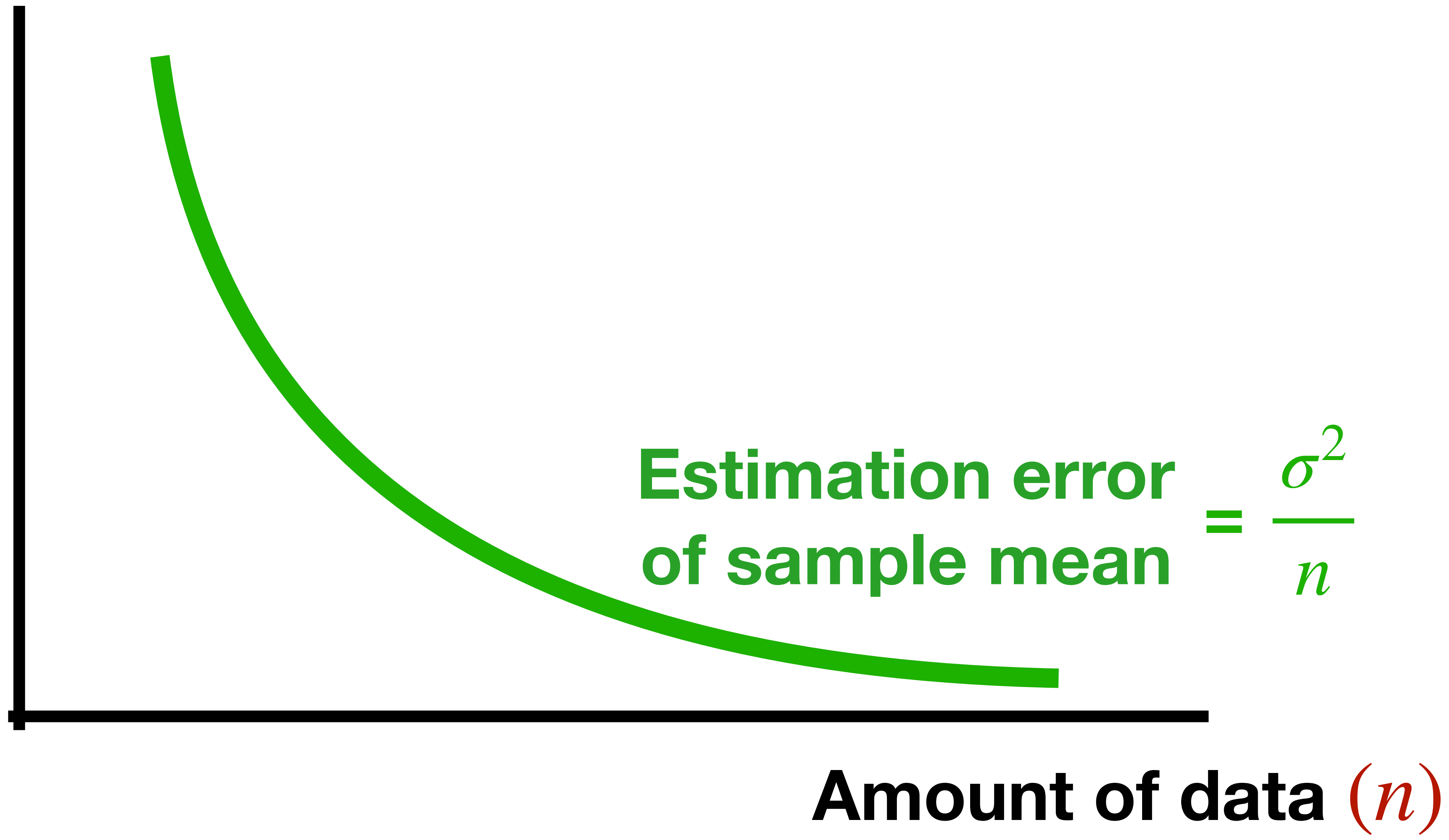
NORMAL MEAN ESTIMATION

- Estimate the mean μ of a normal distribution with *known* variance σ^2.
- An agent can collect samples at *known* unit cost c.
- Each agent wishes to minimize



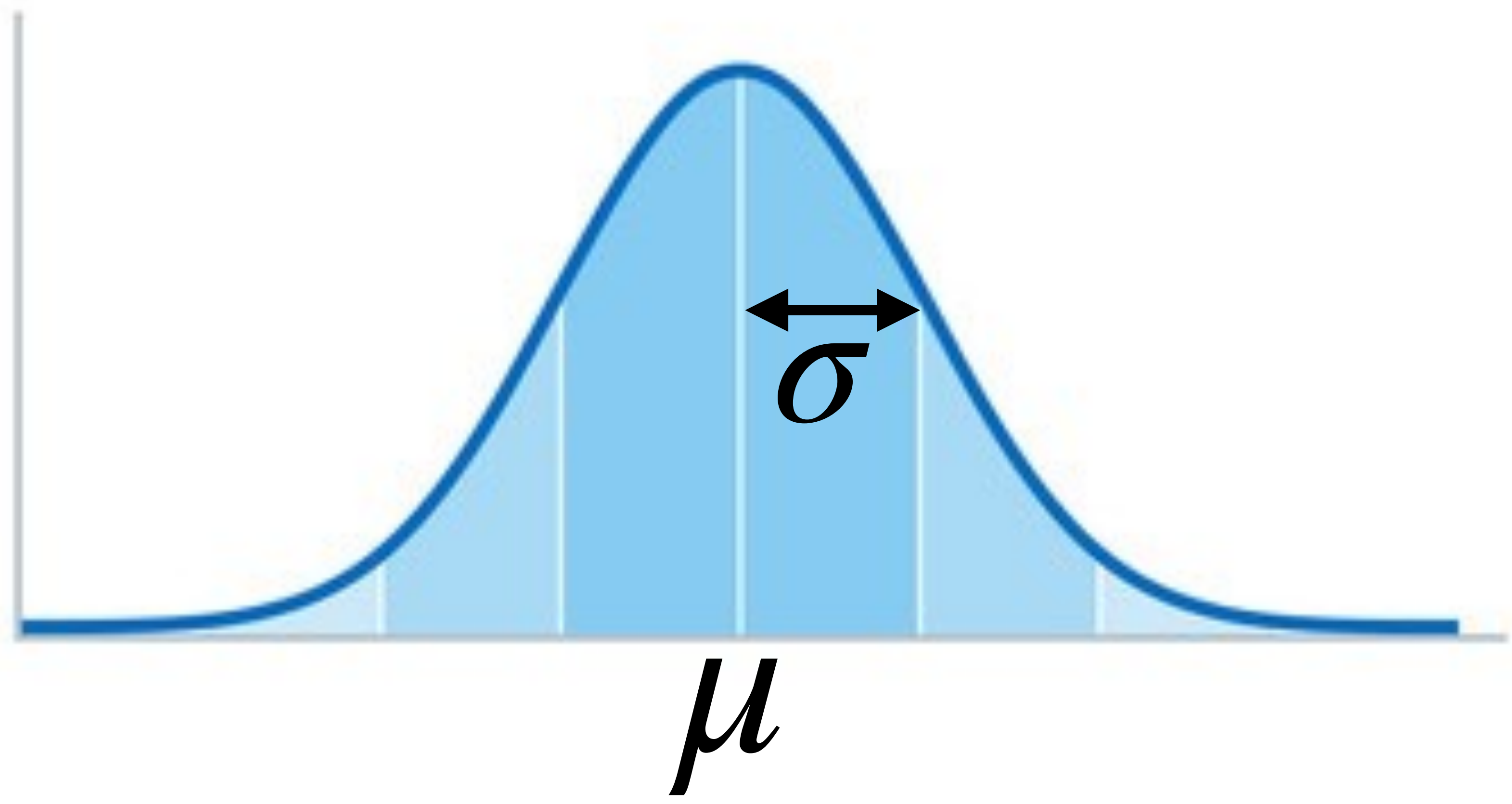
$$\text{penalty} = \text{estimation error} + \text{data collection cost}$$

$$= \frac{\sigma^2}{n} + cn$$

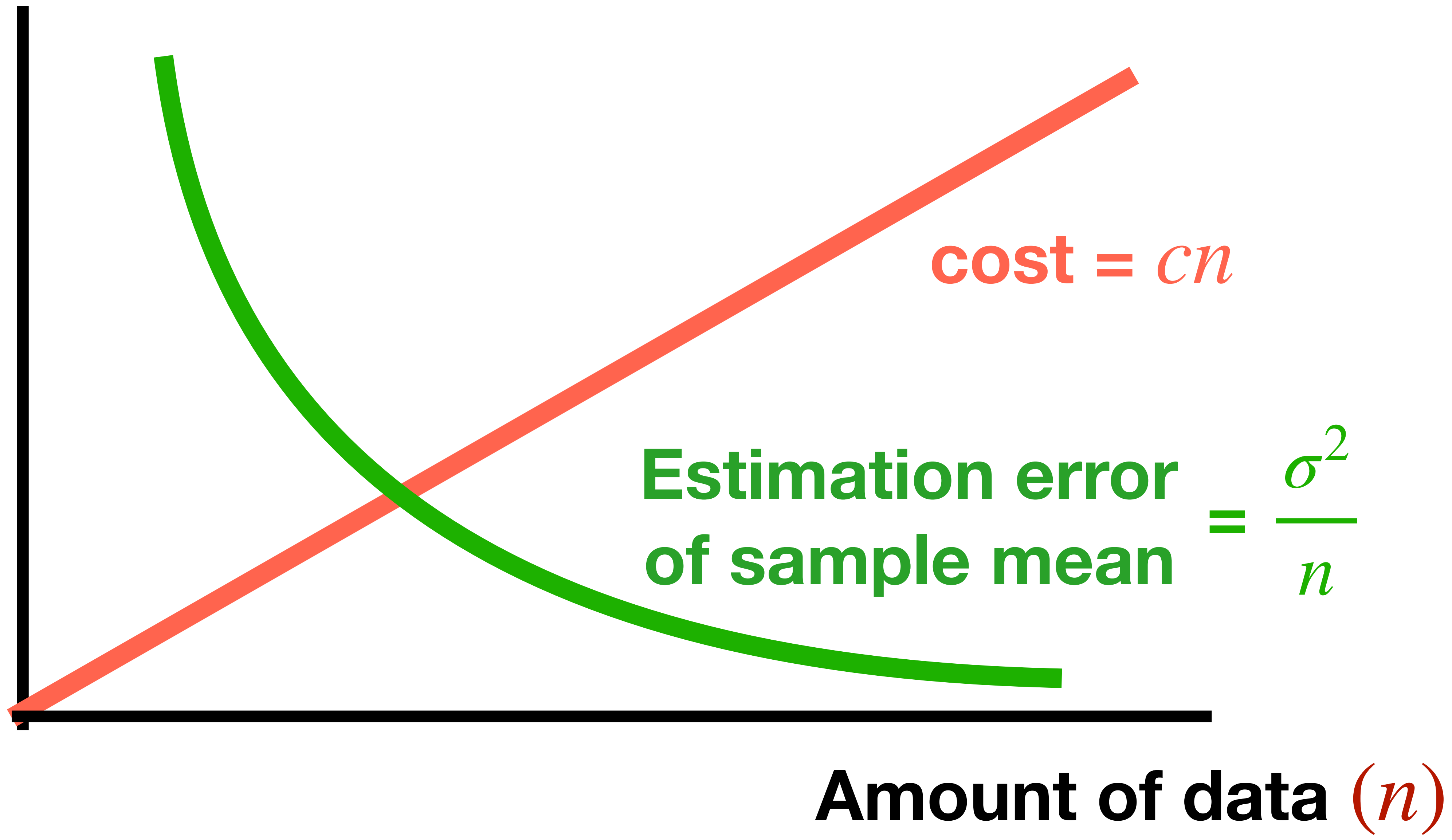


NORMAL MEAN ESTIMATION

- Estimate the mean μ of a normal distribution with *known* variance σ^2.
- An agent can collect samples at *known* unit cost c.
- Each agent wishes to minimize

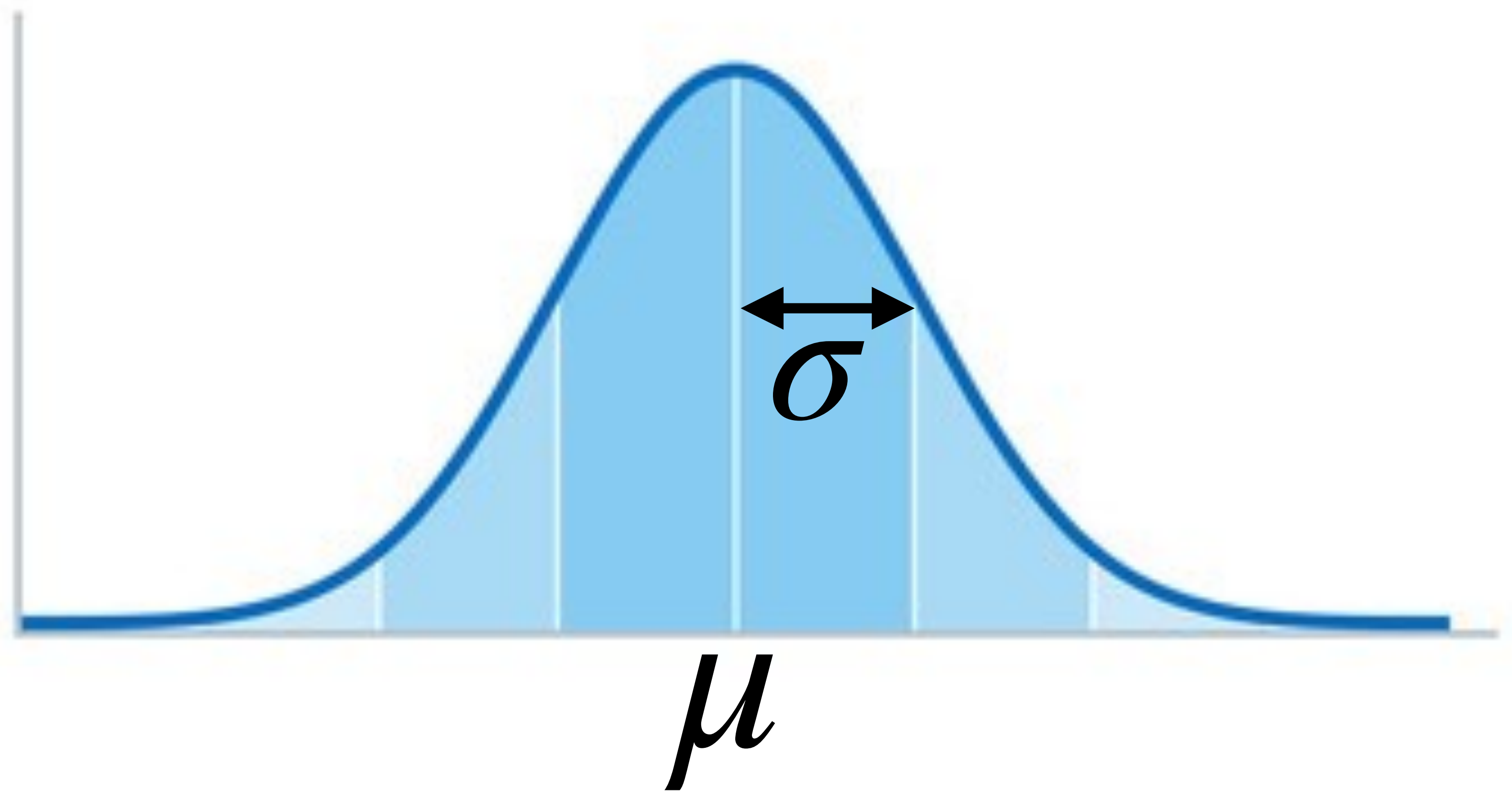


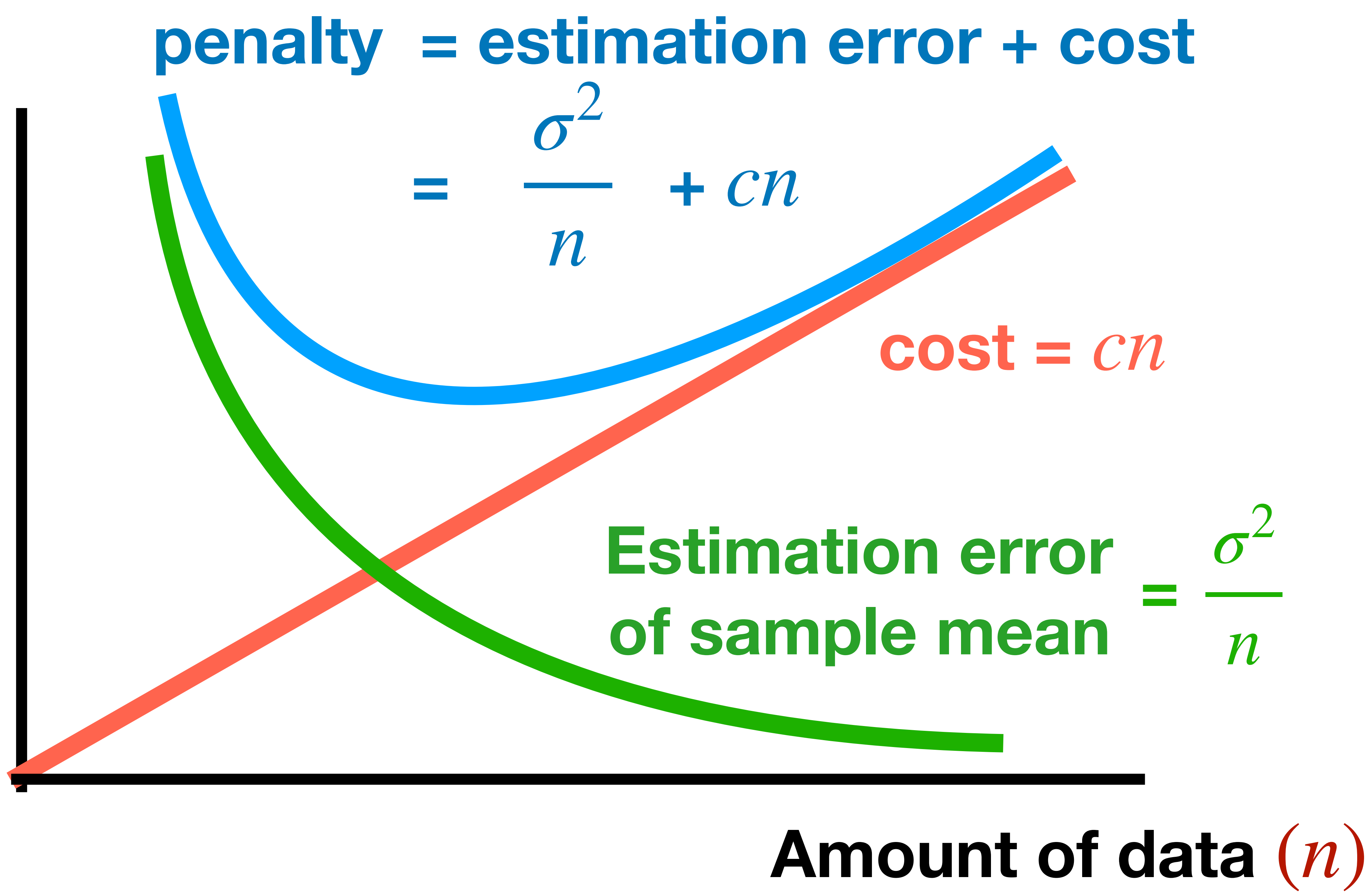
$$\text{penalty} = \text{estimation error} + \text{data collection cost}$$
$$= \frac{\sigma^2}{n} + cn$$



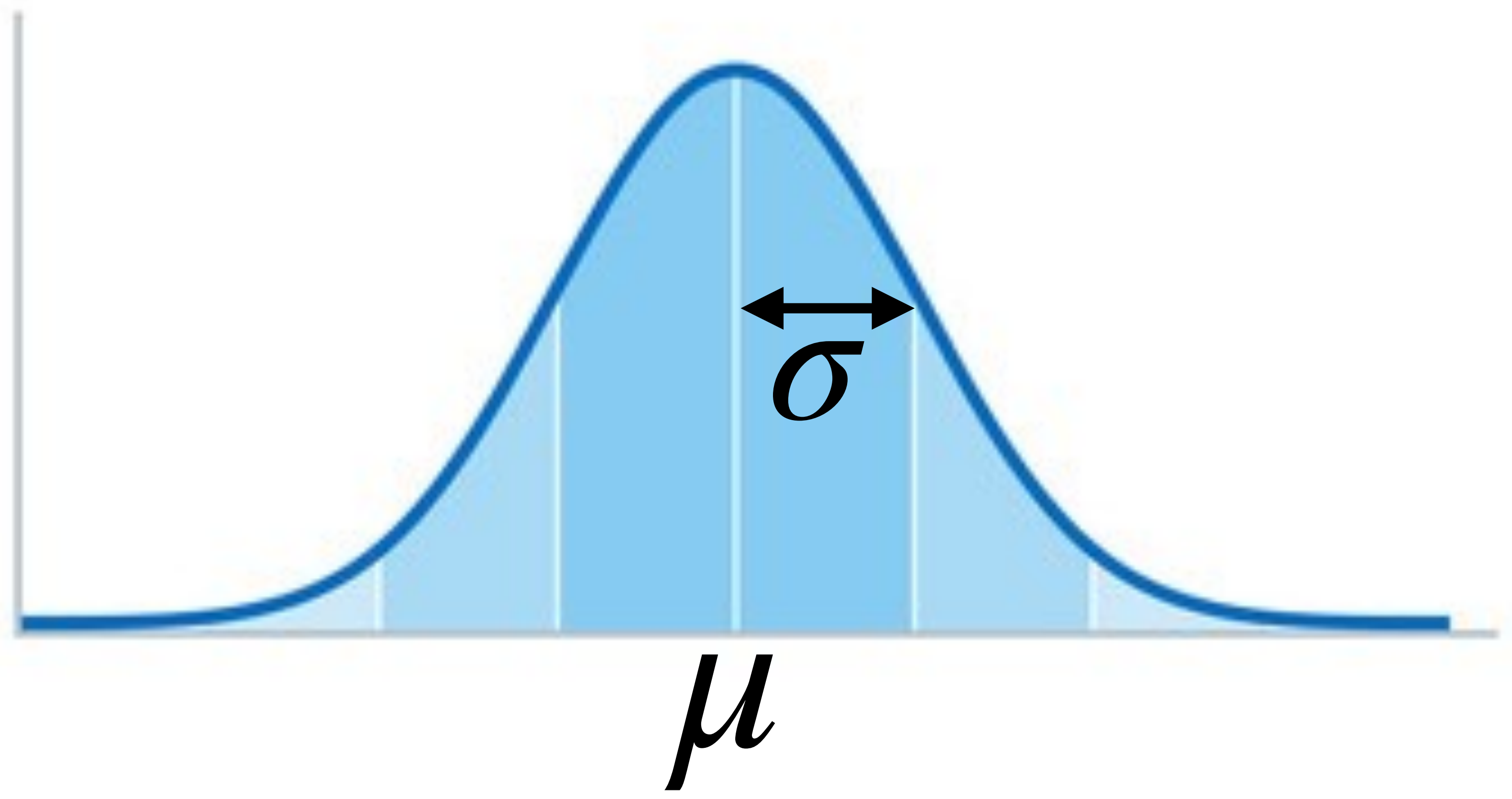
- Estimate the mean μ of a normal distribution with *known* variance σ^2.
- An agent can collect samples at *known* unit cost c.
- Each agent wishes to minimize

$$\text{penalty} = \text{estimation error} + \text{data collection cost}$$
$$= \frac{\sigma^2}{n} + cn$$





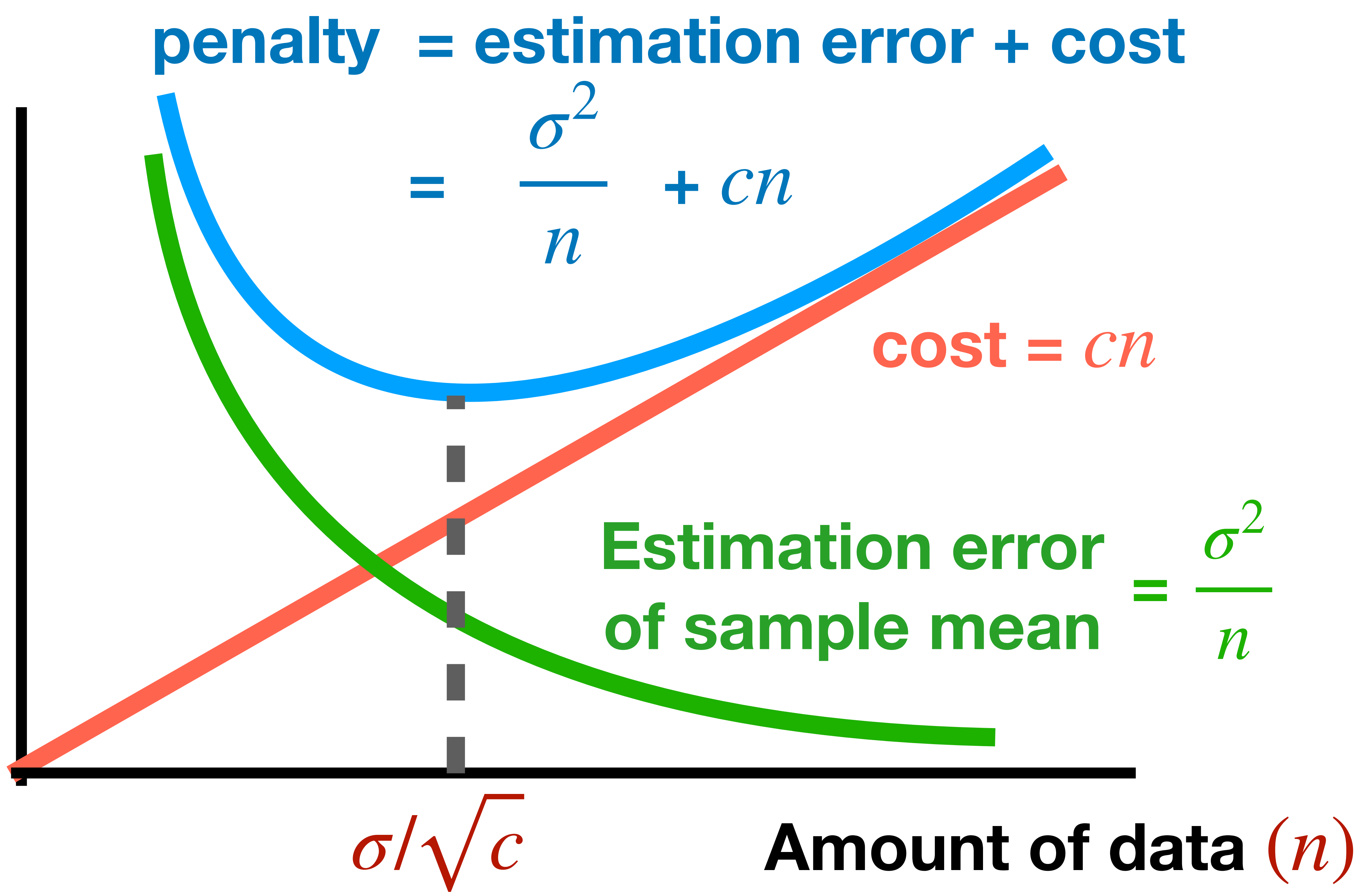
- Estimate the mean μ of a normal distribution with *known* variance σ^2.



- An agent can collect samples at *known* unit cost c.
- Each agent wishes to minimize

$$\text{penalty} = \text{estimation error} + \text{data collection cost}$$
$$= \frac{\sigma^2}{n} + cn$$

- When *working on her own*, agent will collect $\sigma/\sqrt{c}$ points to minimize penalty.



SEVERAL AGENTS WORKING TOGETHER

SEVERAL AGENTS WORKING TOGETHER

- Now consider m agents collecting and sharing their data.

SEVERAL AGENTS WORKING TOGETHER

- Now consider m agents collecting and sharing their data.
- *Social penalty* of all m agents if they collectively collect n_{tot} points.

$$\text{social penalty} = \text{ estimation error of all agents } + \text{data collection cost} = m \times \frac{\sigma^2}{n_{\text{tot}}} + c n_{\text{tot}}$$

SEVERAL AGENTS WORKING TOGETHER

- Now consider m agents collecting and sharing their data.

- *Social penalty* of all m agents if they collectively collect n_{tot} points.

$$\text{social penalty} = \text{ estimation error of all agents } + \text{data collection cost} = m \times \frac{\sigma^2}{n_{\text{tot}}} + c n_{\text{tot}}$$

- To minimize social penalty, they should collect $n^\star_{\text{tot}} = \frac{\sigma\sqrt{m}}{\sqrt{c}}$ points.

SEVERAL AGENTS WORKING TOGETHER

- Now consider m agents collecting and sharing their data.
- *Social penalty* of all m agents if they collectively collect n_{tot} points.

$$\text{social penalty} = \text{ estimation error of all agents } + \text{data collection cost} = m \times \frac{\sigma^2}{n_{\text{tot}}} + c n_{\text{tot}}$$

- To minimize social penalty, they should collect $n_{\text{tot}}^\star = \frac{\sigma\sqrt{m}}{\sqrt{c}}$ points.
  - Each agent needs to collect only $n^\star = \frac{\sigma}{\sqrt{mc}}$ points

    Only $\times\ 1/\sqrt{m}$ when compared to working on her own ($\sigma/\sqrt{c}$ points).

SEVERAL AGENTS WORKING TOGETHER

- Now consider m agents collecting and sharing their data.
- *Social penalty* of all m agents if they collectively collect n_{tot} points.

$$\text{social penalty} = \text{ estimation error of all agents } + \text{data collection cost} = m \times \frac{\sigma^2}{n_{\text{tot}}} + cn_{\text{tot}}$$

- To minimize social penalty, they should collect $n_{\text{tot}}^\star = \frac{\sigma\sqrt{m}}{\sqrt{c}}$ points.
  - Each agent needs to collect only $n^\star = \frac{\sigma}{\sqrt{mc}}$ points

    Only $\times\, 1/\sqrt{m}$ when compared to working on her own ($\sigma/\sqrt{c}$ points).
  - But she has $\times \sqrt{m}$ data.

SEVERAL AGENTS WORKING TOGETHER

	Amount of data she needs to collect (n_i)	Amount of data available to her (n_{tot})	Penalty $\frac{\sigma^2}{n_{\text{tot}}} + cn_i$
Working on her own			
Working together			

SEVERAL AGENTS WORKING TOGETHER

	Amount of data she needs to collect (n_i)	Amount of data available to her (n_{tot})	Penalty $\frac{\sigma^2}{n_{tot}} + cn_i$
Working on her own	$\frac{\sigma}{\sqrt{c}}$		
Working together	$\frac{\sigma}{\sqrt{cm}}$		

SEVERAL AGENTS WORKING TOGETHER

	Amount of data she needs to collect (n_i)	Amount of data available to her (n_{tot})	Penalty $\frac{\sigma^2}{n_{\text{tot}}} + cn_i$
Working on her own	$\frac{\sigma}{\sqrt{c}}$	$\frac{\sigma}{\sqrt{c}}$	
Working together	$\frac{\sigma}{\sqrt{cm}}$	$\frac{\sigma\sqrt{m}}{\sqrt{c}}$	

SEVERAL AGENTS WORKING TOGETHER

	Amount of data she needs to collect (n_i)	Amount of data available to her (n_{tot})	Penalty $\frac{\sigma^2}{n_{\text{tot}}} + cn_i$
Working on her own	$\frac{\sigma}{\sqrt{c}}$	$\frac{\sigma}{\sqrt{c}}$	$2\sigma\sqrt{c}$
Working together	$\frac{\sigma}{\sqrt{cm}}$	$\frac{\sigma\sqrt{m}}{\sqrt{c}}$	$\frac{2\sigma\sqrt{c}}{\sqrt{m}}$

SEVERAL AGENTS WORKING TOGETHER

	Amount of data she needs to collect (n_i)	Amount of data available to her (n_{tot})	Penalty $\frac{\sigma^2}{n_{\text{tot}}} + cn_i$
Working on her own	$\frac{\sigma}{\sqrt{c}}$	$\frac{\sigma}{\sqrt{c}}$	$2\sigma\sqrt{c}$
Working together	$\frac{\sigma}{\sqrt{cm}}$	$\frac{\sigma\sqrt{m}}{\sqrt{c}}$	$\frac{2\sigma\sqrt{c}}{\sqrt{m}}$

Agents can reduce data collection costs, and improve estimation error by sharing data with others.

DESIGNING A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

- Naive mechanism 1: “pool and share”

- Naive mechanism 1: "pool and share"
  - Selfish agents will *free-ride:* not contributing any data herself, but using data that the others have contributed.

$$\text{penalty} = \frac{\sigma^2}{n_{\text{tot}}} + c \times n_i$$

- Naive mechanism 1: "pool and share"
  - Selfish agents will *free-ride:* not contributing any data herself, but using data that the others have contributed.

$$\text{penalty} = \frac{\sigma^2}{n_{\text{tot}}} + c \times n_i = \frac{\sigma^2}{(m-1) \times \frac{\sigma}{\sqrt{mc}}} + c \times 0$$

- Naive mechanism 1: "pool and share"
  - Selfish agents will *free-ride:* not contributing any data herself, but using data that the others have contributed.

$$\text{penalty} = \frac{\sigma^2}{n_{\text{tot}}} + c \times n_i = \frac{\sigma^2}{(m-1) \times \frac{\sigma}{\sqrt{mc}}} + c \times 0 \approx \frac{\sigma\sqrt{c}}{\sqrt{m}}$$

- Naive mechanism 1: "pool and share"
  - Selfish agents will *free-ride:* not contributing any data herself, but using data that the others have contributed.

$$\text{penalty} = \frac{\sigma^2}{n_{\text{tot}}} + c \times n_i = \frac{\sigma^2}{(m-1) \times \frac{\sigma}{\sqrt{mc}}} + c \times 0 \approx \frac{\sigma\sqrt{c}}{\sqrt{m}} = \frac{1}{2} \times \underbrace{\frac{2\sigma\sqrt{c}}{\sqrt{m}}}_{\text{penalty for a well-behaved agent}}$$

- Naive mechanism 1: "pool and share"
  - Selfish agents will *free-ride:* not contributing any data herself, but using data that the others have contributed.

$$\text{penalty} = \frac{\sigma^2}{n_{\text{tot}}} + c \times n_i = \frac{\sigma^2}{(m-1) \times \frac{\sigma}{\sqrt{mc}}} + c \times 0 \approx \frac{\sigma\sqrt{c}}{\sqrt{m}} = \frac{1}{2} \times \underbrace{\frac{2\sigma\sqrt{c}}{\sqrt{m}}}_{\text{penalty for a well-behaved agent}}$$

- Naive mechanism 2: "pool and share, but only if you contribute enough data"

- Naive mechanism 1: "pool and share"
  - Selfish agents will *free-ride:* not contributing any data herself, but using data that the others have contributed.

$$\text{penalty} = \frac{\sigma^2}{n_{\text{tot}}} + c \times n_i = \frac{\sigma^2}{(m-1) \times \frac{\sigma}{\sqrt{mc}}} + c \times 0 \approx \frac{\sigma\sqrt{c}}{\sqrt{m}} = \frac{1}{2} \times \underbrace{\frac{2\sigma\sqrt{c}}{\sqrt{m}}}_{\text{penalty for a well-behaved agent}}$$

- Naive mechanism 2: "pool and share, but only if you contribute enough data"
  - Agents can fabricate and then discard after receiving others' data.

OVERVIEW OF OUR MECHANISM

OVERVIEW OF OUR MECHANISM

Each agent i will:

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit n_i' points $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\}$.

    # Agents may collect any number of points, and lie (e.g withhold, fabricate) about what they collect.

OVERVIEW OF OUR MECHANISM

Each agent i will:

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit n_i' points $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\}$.

  # Agents may collect any number of points, and lie (e.g withhold, fabricate) about what they collect.

The mechanism:

- To each agent, allocates a noisy version A_i of the others' data. The noise is proportional to how much the agent's submission Y_i differs from the others' submissions $\{Y_j\}_{j \neq i}$.

Each agent i will:

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit n_i' points $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\}$.

Agents may collect any number of points, and lie (e.g withhold, fabricate) about what they collect.

The mechanism:

- To each agent, allocates a noisy version A_i of the others' data. The noise is proportional to how much the agent's submission Y_i differs from the others' submissions $\{Y_j\}_{j \neq i}$.

Each agent i will:

- Estimate μ using all the information they have (X_i, Y_i, A_i).

Each agent i will:

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit n_i' points $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\}$.

Agents may collect any number of points, and lie (e.g withhold, fabricate) about what they collect.

The mechanism:

- To each agent, allocates a noisy version A_i of the others' data. The noise is proportional to how much the agent's submission Y_i differs from the others' submissions $\{Y_j\}_{j \neq i}$.

Each agent i will:

- Estimate μ using all the information they have (X_i, Y_i, A_i).
  - *We design a (minimax) optimal estimator to enforce truthful reporting.*

OVERVIEW OF OUR RESULTS

OVERVIEW OF OUR RESULTS

This mechanism is

OVERVIEW OF OUR RESULTS

This mechanism is

- **Nash incentive-compatible:** Provided that other agents are well-behaved, the best strategy for an agent is to,

OVERVIEW OF OUR RESULTS

This mechanism is

- **Nash incentive-compatible:** Provided that other agents are well-behaved, the best strategy for an agent is to,
  - Collect a sufficient amount ($n^\star = \sigma/\sqrt{mc}$) of data.

OVERVIEW OF OUR RESULTS

This mechanism is

- **Nash incentive-compatible:** Provided that other agents are well-behaved, the best strategy for an agent is to,
  - Collect a sufficient amount ($n^\star = \sigma/\sqrt{mc}$) of data.
  - Submit it truthfully.

OVERVIEW OF OUR RESULTS

This mechanism is

- **Nash incentive-compatible:** Provided that other agents are well-behaved, the best strategy for an agent is to,
    - Collect a sufficient amount ($n^\star = \sigma/\sqrt{mc}$) of data.
    - Submit it truthfully.
    - Use the recommended minimax-optimal estimator.

OVERVIEW OF OUR RESULTS

This mechanism is

- **Nash incentive-compatible:** Provided that other agents are well-behaved, the best strategy for an agent is to,
  - Collect a sufficient amount ($n^\star = \sigma/\sqrt{mc}$) of data.
  - Submit it truthfully.
  - Use the recommended minimax-optimal estimator.
- **Individually rational:** Provided that others are well-behaved, an agent does not do worse than the best she could do on her own.

OVERVIEW OF OUR RESULTS

This mechanism is

- **Nash incentive-compatible:** Provided that other agents are well-behaved, the best strategy for an agent is to,
  - Collect a sufficient amount ($n^\star = \sigma/\sqrt{mc}$) of data.
  - Submit it truthfully.
  - Use the recommended minimax-optimal estimator.
- **Individually rational:** Provided that others are well-behaved, an agent does not do worse than the best she could do on her own.
- **Approximately efficient:** Social penalty at the Nash strategies is at most a factor 2 of the global minimum.

OUTLINE

1. **Mechanism design for collaborative normal mean estimation**

   (Chen, Zhu, Kandasamy, *NeurIPS 2023*)

   - Intuitions, overview of results
   - **Problem formalism**
   - Mechanism and theoretical analysis

2. Extensions (Clinton, Chen, Zhu, Kandasamy, *Ongoing work*)

   - Multiple distributions with asymmetric data collection capabilities
   - Collaborative supervised learning and experiment design

FORMALISM 1/4: MECHANISM

FORMALISM 1/4: MECHANISM

A mechanism M receives a dataset from each agent, and returns an *allocation* A_i to each agent i.

FORMALISM 1/4: MECHANISM

A mechanism M receives a dataset from each agent, and returns an *allocation* A_i to each agent i.

- The mechanism designer can choose a space of allocations $\mathscr{A}$.

FORMALISM 1/4: MECHANISM

A mechanism M receives a dataset from each agent, and returns an *allocation* A_i to each agent i.

- The mechanism designer can choose a space of allocations $\mathscr{A}$.

  E.g. A larger dataset, $\mathscr{A} = \bigcup_{k \geq 0} \mathbb{R}^k$

FORMALISM 1/4: MECHANISM

A mechanism M receives a dataset from each agent, and returns an *allocation* A_i to each agent i.

- The mechanism designer can choose a space of allocations $\mathscr{A}$.

  E.g. A larger dataset, $\mathscr{A} = \bigcup_{k \geq 0} \mathbb{R}^k$

We can write the space of mechanisms $\mathscr{M}$ as,

FORMALISM 1/4: MECHANISM

A mechanism M receives a dataset from each agent, and returns an *allocation* A_i to each agent i.

- The mechanism designer can choose a space of allocations $\mathscr{A}$.

  E.g. A larger dataset, $\mathscr{A} = \bigcup_{k \geq 0} \mathbb{R}^k$

We can write the space of mechanisms $\mathscr{M}$ as,

$$\mathscr{M} = \left\{ M = (\mathscr{A}, b); \quad \mathscr{A} \subset \text{universal set}, \quad b : \left(\bigcup_{n \geq 0} \mathbb{R}^n \right)^m \to \mathscr{A}^m \right\}$$

A mechanism M receives a dataset from each agent, and returns an *allocation* A_i to each agent i.

- The mechanism designer can choose a space of allocations $\mathscr{A}$.

  E.g. A larger dataset, $\mathscr{A} = \bigcup_{k\geq 0} \mathbb{R}^k$

Datasets received from the m agents.

We can write the space of mechanisms $\mathscr{M}$ as,

$$\mathscr{M} = \left\{ M = (\mathscr{A}, b); \quad \mathscr{A} \subset \text{universal set}, \quad b : \left(\bigcup_{n\geq 0} \mathbb{R}^n \right)^m \to \mathscr{A}^m \right\}$$

FORMALISM 2/4: AGENT'S STRATEGY SPACE

FORMALISM 2/4: AGENT'S STRATEGY SPACE

After the mechanism is published an agent will

FORMALISM 2/4: AGENT'S STRATEGY SPACE

After the mechanism is published an agent will

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$.

FORMALISM 2/4: AGENT'S STRATEGY SPACE

After the mechanism is published an agent will

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$.
- Submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.
  - f_i maps the dataset collected to possibly altered dataset (e.g fabrication, withholding etc), of a potentially different size.

FORMALISM 2/4: AGENT'S STRATEGY SPACE

After the mechanism is published an agent will

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$.
- Submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.
  - f_i maps the dataset collected to possibly altered dataset (e.g fabrication, withholding etc), of a potentially different size.
- On receiving her allocation A_i, she will estimate μ via an estimator $h_i(X_i, Y_i, A_i)$.

FORMALISM 2/4: AGENT'S STRATEGY SPACE

After the mechanism is published an agent will

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$.
- Submit $Y_i = \{y_{i,1}, \ldots, y_{i,n'_i}\} = f_i(X_i)$.
  - f_i maps the dataset collected to possibly altered dataset (e.g fabrication, withholding etc), of a potentially different size.
- On receiving her allocation A_i, she will estimate μ via an estimator $h_i(X_i, Y_i, A_i)$.
  - An agent need not use the "straightforward" (e.g sample mean) estimator.

FORMALISM 2/4: AGENT'S STRATEGY SPACE

After the mechanism is published an agent will

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$.
- Submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.
  - f_i maps the dataset collected to possibly altered dataset (e.g fabrication, withholding etc), of a potentially different size.
- On receiving her allocation A_i, she will estimate μ via an estimator $h_i(X_i, Y_i, A_i)$.
  - An agent need not use the "straightforward" (e.g sample mean) estimator.

An agent's strategy $s_i = (n_i, f_i, h_i) \in \mathbb{N} \times \mathcal{F} \times \mathcal{H}$

FORMALISM 2/4: AGENT'S STRATEGY SPACE

After the mechanism is published an agent will

- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$.
- Submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.
  - f_i maps the dataset collected to possibly altered dataset (e.g fabrication, withholding etc), of a potentially different size.
- On receiving her allocation A_i, she will estimate μ via an estimator $h_i(X_i, Y_i, A_i)$.
  - An agent need not use the "straightforward" (e.g sample mean) estimator.

An agent's strategy $s_i = (n_i, f_i, h_i) \in \mathbb{N} \times \mathcal{F} \times \mathcal{H}$

$$\mathcal{F} = \text{submission functions} = \left\{ f : \bigcup_{n \geq 0} \mathbb{R}^n \to \bigcup_{n \geq 0} \mathbb{R}^n \right\}, \qquad \mathcal{H} = \text{estimators} = \left\{ h : \bigcup_{n \geq 0} \mathbb{R}^n \times \bigcup_{n \geq 0} \mathbb{R}^n \times \mathcal{A} \to \mathbb{R} \right\}$$

FORMALISM 3/4: AGENT'S PENALTY (NEGATIVE UTILITY)

FORMALISM 3/4: AGENT'S PENALTY (NEGATIVE UTILITY)

An agent's penalty p_i in a mechanism M under a strategy profile $s = (s_1, \ldots, s_m)$,

FORMALISM 3/4: AGENT'S PENALTY (NEGATIVE UTILITY)

An agent's penalty p_i in a mechanism M under a strategy profile $s = (s_1, \ldots, s_m)$,

$$p_i(M, s) = \text{estimation error} + \text{data collection cost}$$

$$= \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i(X_i, f_i(X_i), A_i) - \mu \right)^2 \right] + cn_i$$

FORMALISM 3/4: AGENT'S PENALTY (NEGATIVE UTILITY)

An agent's penalty p_i in a mechanism M under a strategy profile $s = (s_1, \ldots, s_m)$,

$$
\begin{aligned}
p_i(M, s) &= \text{estimation error} + \text{data collection cost} \\
&= \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i(X_i, f_i(X_i), A_i) - \mu \right)^2 \right] + c n_i
\end{aligned}
$$

We take a $\sup_{\mu \in \mathbb{R}} \ldots$ since μ is unknown. Makes the problem well-defined.

FORMALISM 3/4: AGENT'S PENALTY (NEGATIVE UTILITY)

An agent's penalty p_i in a mechanism M under a strategy profile $s = (s_1, \ldots, s_m)$,

$$p_i(M, s) = \text{estimation error} + \text{data collection cost}$$

$$= \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i(X_i, f_i(X_i), A_i) - \mu \right)^2 \right] + c n_i$$

We take a $\sup_{\mu \in \mathbb{R}} \ldots$ since μ is unknown. Makes the problem well-defined.

- Otherwise, consider setting $n_i = 0$ and $h_i(\cdot, \cdot, \cdot) = \mu'$ for some $\mu' \in \mathbb{R}$.

FORMALISM 3/4: AGENT'S PENALTY (NEGATIVE UTILITY)

An agent's penalty p_i in a mechanism M under a strategy profile $s = (s_1, \ldots, s_m)$,

$$p_i(M, s) = \text{estimation error} + \text{data collection cost}$$

$$= \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i(X_i, f_i(X_i), A_i) - \mu \right)^2 \right] + c n_i$$

We take a $\sup_{\mu \in \mathbb{R}} \ldots$ since μ is unknown. Makes the problem well-defined.

- Otherwise, consider setting $n_i = 0$ and $h_i(\cdot, \cdot, \cdot) = \mu'$ for some $\mu' \in \mathbb{R}$.
- When the true mean is $\mu = \mu'$, this strategy achieves zero penalty!

FORMALISM 3/4: AGENT'S PENALTY (NEGATIVE UTILITY)

An agent's penalty p_i in a mechanism M under a strategy profile $s = (s_1, \ldots, s_m)$,

$$p_i(M, s) = \text{estimation error} + \text{data collection cost}$$

$$= \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i(X_i, f_i(X_i), A_i) - \mu \right)^2 \right] + cn_i$$

We take a $\sup_{\mu \in \mathbb{R}} \ldots$ since μ is unknown. Makes the problem well-defined.

- Otherwise, consider setting $n_i = 0$ and $h_i(\cdot, \cdot, \cdot) = \mu'$ for some $\mu' \in \mathbb{R}$.
- When the true mean is $\mu = \mu'$, this strategy achieves zero penalty!
- But this works only if agent knows $\mu = \mu'$ a priori.

FORMALISM 4/4: DESIDERATA FOR A MECHANISM

FORMALISM 4/4: DESIDERATA FOR A MECHANISM

A mechanism will also publish a *recommended strategy profile* $s^\star = \{s_i^\star\}_{i \in [n]}$.

FORMALISM 4/4: DESIDERATA FOR A MECHANISM

A mechanism will also publish a *recommended strategy profile* $s^\star = \{s_i^\star\}_{i \in [n]}$.

Desiderata:

FORMALISM 4/4: DESIDERATA FOR A MECHANISM

A mechanism will also publish a *recommended strategy profile* $s^\star = \{s_i^\star\}_{i\in[n]}$.

Desiderata:

1. **Nash incentive-compatible (NIC):** $s^\star$ is a Nash equilibrium, i.e
   $p_i(M, (s_i^\star, s_{-i}^\star)) \leq p_i(M, (s_i, s_{-i}^\star))$ for all agents i and all other strategies s_i.

FORMALISM 4/4: DESIDERATA FOR A MECHANISM

A mechanism will also publish a *recommended strategy profile* $s^\star = \{s_i^\star\}_{i \in [n]}$.

Desiderata:

1. **Nash incentive-compatible (NIC):** $s^\star$ is a Nash equilibrium, i.e $p_i(M, (s_i^\star, s_{-i}^\star)) \leq p_i(M, (s_i, s_{-i}^\star))$ for all agents i and all other strategies s_i.
2. **Individually rational (IR):** An agent's penalty at $s^\star$ is no worse than the lowest penalty she could achieve on her own, i.e $p_i(M, s^\star) \leq 2\sigma/\sqrt{c}$.

FORMALISM 4/4: DESIDERATA FOR A MECHANISM

A mechanism will also publish a *recommended strategy profile* $s^\star = \{s_i^\star\}_{i \in [n]}$.

Desiderata:

1. **Nash incentive-compatible (NIC):** $s^\star$ is a Nash equilibrium, i.e $p_i(M, (s_i^\star, s_{-i}^\star)) \leq p_i(M, (s_i, s_{-i}^\star))$ for all agents i and all other strategies s_i.
2. **Individually rational (IR):** An agent's penalty at $s^\star$ is no worse than the lowest penalty she could achieve on her own, i.e $p_i(M, s^\star) \leq 2\sigma/\sqrt{c}$.
3. **Approximately efficient:** The social penalty $P(M, s^\star) = \sum_i p_i(M, s^\star)$ is at most a constant factor of the global minimum, i.e

$$P(M, s^\star) \leq \mathcal{O}(1) \cdot \min_{M', s'} p(M', s')$$

FORMALISM 4/4: DESIDERATA FOR A MECHANISM

A mechanism will also publish a *recommended strategy profile* $s^\star = \{s_i^\star\}_{i \in [n]}$.

Desiderata:

1. **Nash incentive-compatible (NIC):** $s^\star$ is a Nash equilibrium, i.e $p_i(M, (s_i^\star, s_{-i}^\star)) \leq p_i(M, (s_i, s_{-i}^\star))$ for all agents i and all other strategies s_i.
2. **Individually rational (IR):** An agent's penalty at $s^\star$ is no worse than the lowest penalty she could achieve on her own, i.e $p_i(M, s^\star) \leq 2\sigma/\sqrt{c}$.
3. **Approximately efficient:** The social penalty $P(M, s^\star) = \sum_i p_i(M, s^\star)$ is at most a constant factor of the global minimum, i.e

$$P(M, s^\star) \leq \mathcal{O}(1) \cdot \min_{M', s'} p(M', s')$$

min without NIC, IR constraints

FORMALISM 4/4: DESIDERATA FOR A MECHANISM

A mechanism will also publish a *recommended strategy profile* $s^\star = \{s_i^\star\}_{i\in[n]}$.

Desiderata:

1. **Nash incentive-compatible (NIC):** $s^\star$ is a Nash equilibrium, i.e
   $p_i(M, (s_i^\star, s_{-i}^\star)) \le p_i(M, (s_i, s_{-i}^\star))$ for all agents i and all other strategies s_i.
2. **Individually rational (IR):** An agent's penalty at $s^\star$ is no worse than the lowest penalty she could achieve on her own, i.e $p_i(M, s^\star) \le 2\sigma/\sqrt{c}$.
3. **Approximately efficient:** The social penalty $P(M, s^\star) = \sum_i p_i(M, s^\star)$ is at most a constant factor of the global minimum, i.e

$$P(M, s^\star) \le \mathcal{O}(1) \cdot \underbrace{\min_{M', s'} p(M', s')}_{= \; 2\sigma\sqrt{mc}}$$

min without NIC, IR constraints

$= 2\sigma\sqrt{mc}$ (pool-and-share)

OUTLINE

1. **Mechanism design for collaborative normal mean estimation**

(Chen, Zhu, Kandasamy, *NeurIPS 2023*)

- Intuitions, overview of results
- Problem formalism
- **Mechanism and theoretical analysis**

2. Extensions (Clinton, Chen, Zhu, Kandasamy, *Ongoing work*)

- Multiple distributions with asymmetric data collection capabilities
- Collaborative supervised learning and experiment design

A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$

A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

Mechanism

A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

Mechanism

- For each agent i:

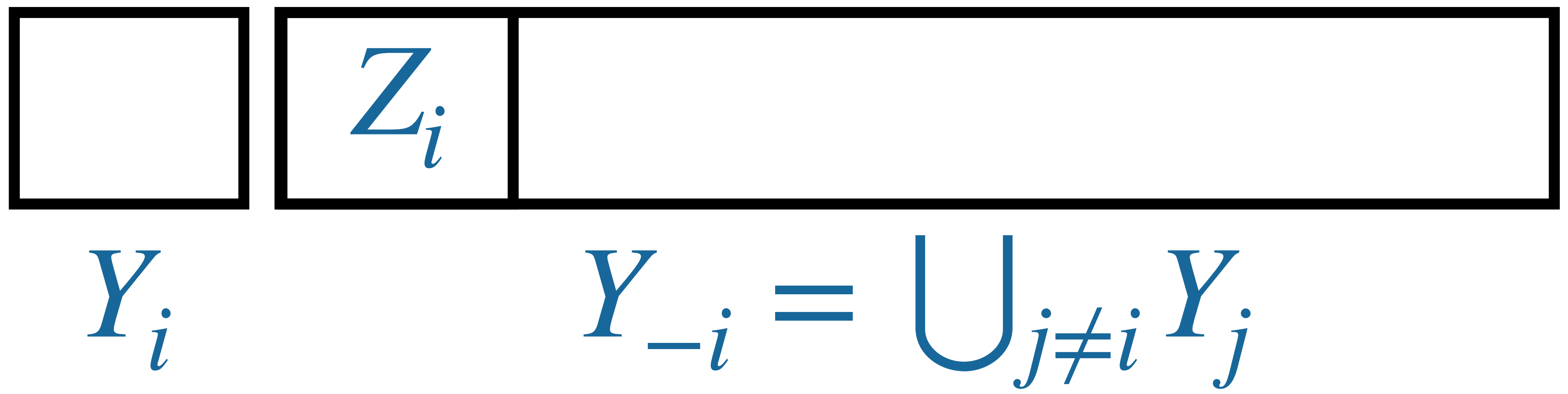
A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

Mechanism

- For each agent i:
  - $Z_i \leftarrow$ randomly sample $\sigma/\sqrt{cm}$ points from others' submissions Y_{-i}.



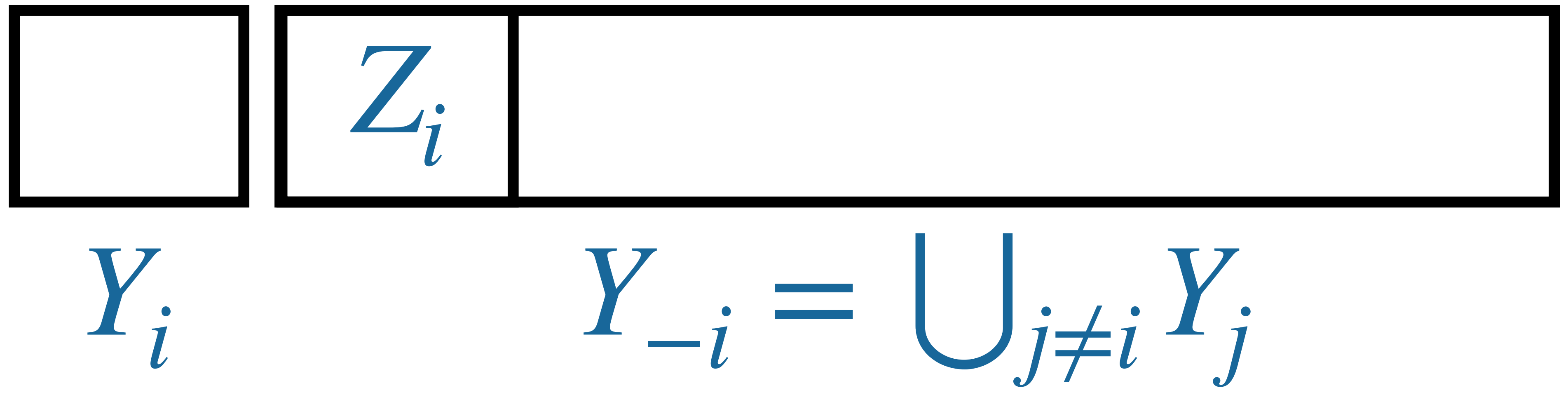
A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

Mechanism

- For each agent i:
    - $Z_i \leftarrow$ randomly sample $\sigma/\sqrt{cm}$ points from others' submissions Y_{-i}.
    - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$ # Variance proportional to difference



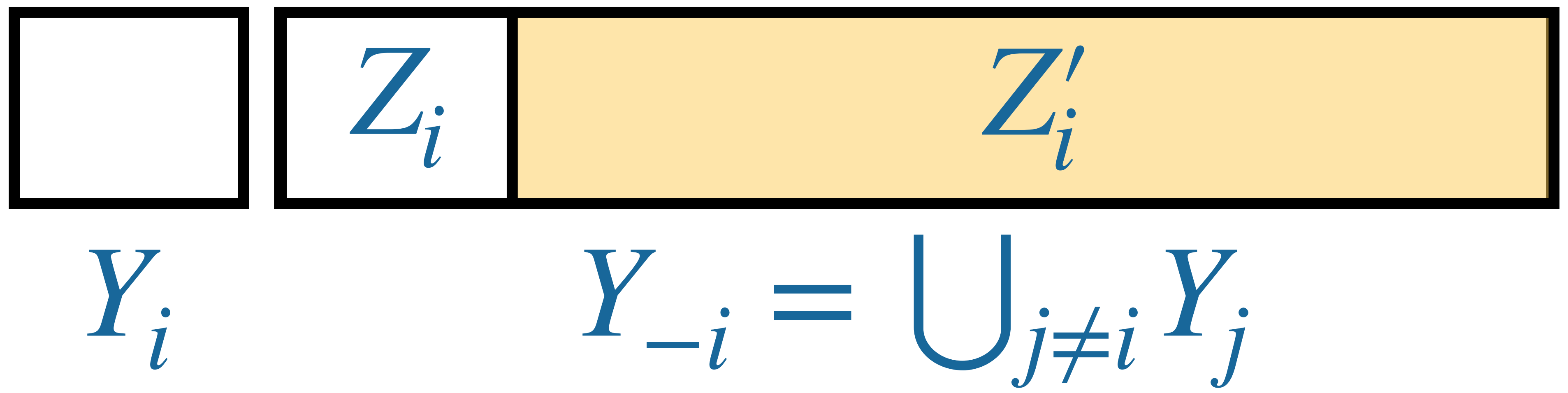
A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

Mechanism

- For each agent i:
  - $Z_i \leftarrow$ randomly sample $\sigma/\sqrt{cm}$ points from others' submissions Y_{-i}.
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$ # Variance proportional to difference
  - $Z_i' \leftarrow \left\{ z + \epsilon_z, \quad \text{for all } z \in Y_{-i} \backslash Z_i, \quad \text{where } \epsilon_z \sim \mathcal{N}(0, \eta_i^2) \right\}$.



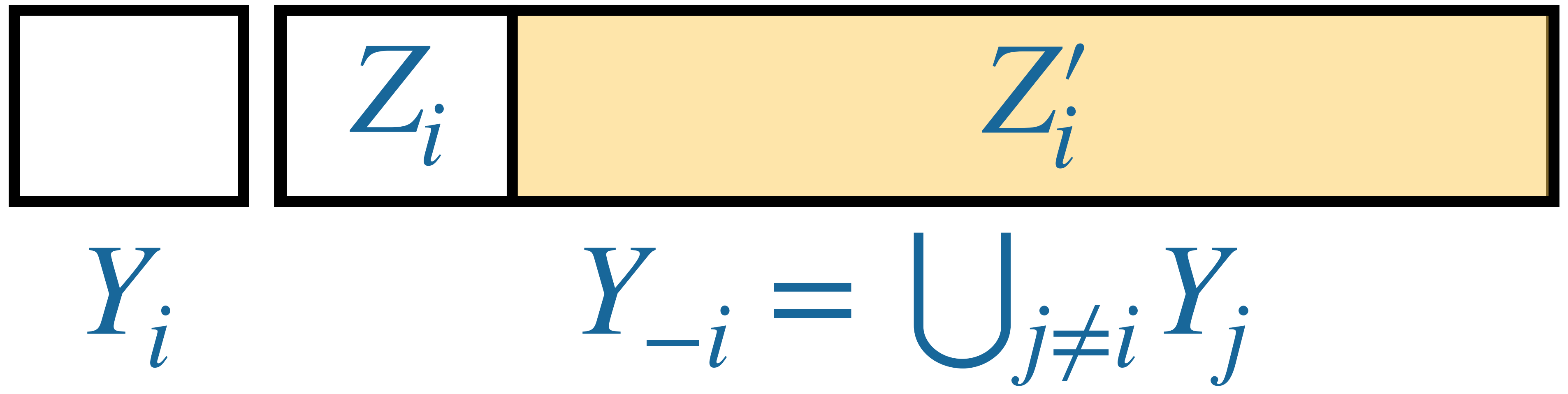
A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

Mechanism

- For each agent i:
  - $Z_i \leftarrow$ randomly sample $\sigma/\sqrt{cm}$ points from others' submissions Y_{-i}.
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$ # Variance proportional to difference
  - $Z_i' \leftarrow \left\{ z + \epsilon_z, \quad \text{for all } z \in Y_{-i} \backslash Z_i, \quad \text{where } \epsilon_z \sim \mathcal{N}(0, \eta_i^2) \right\}$.
  - Set allocation to each agent, $A_i \leftarrow (Z_i, Z_i', \eta_i^2)$.



A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

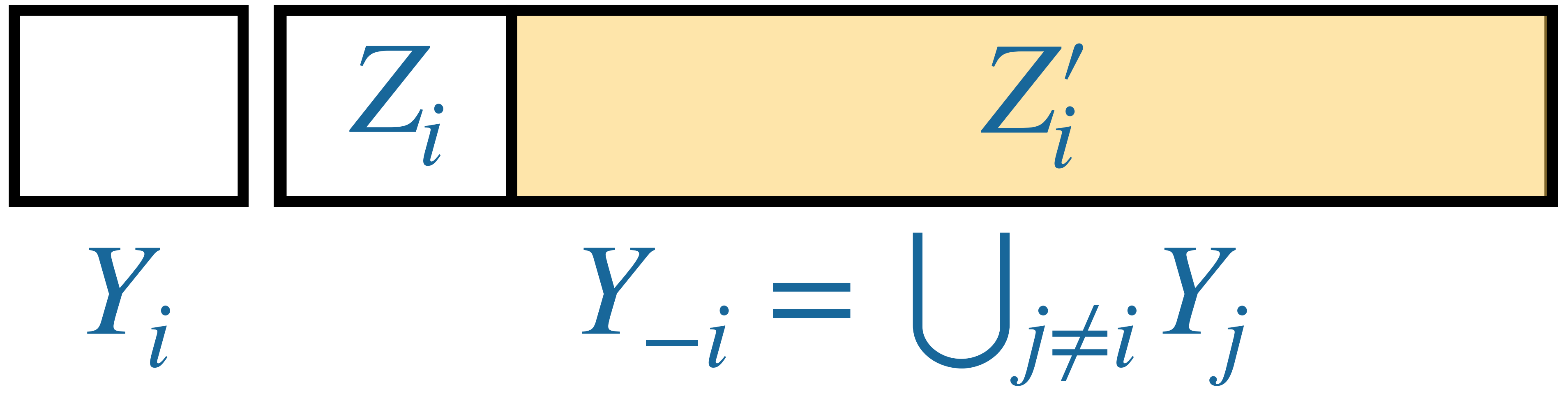
Each agent i will

- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

Mechanism

- For each agent i:
  - $Z_i \leftarrow$ randomly sample $\sigma/\sqrt{cm}$ points from others' submissions Y_{-i}.
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$ # Variance proportional to difference
  - $Z_i' \leftarrow \left\{ z + \epsilon_z, \quad \text{for all } z \in Y_{-i} \backslash Z_i, \quad \text{where } \epsilon_z \sim \mathcal{N}(0, \eta_i^2) \right\}$.
  - Set allocation to each agent, $A_i \leftarrow (Z_i, Z_i', \eta_i^2)$.

Each agent i will



A MECHANISM FOR COLLABORATIVE NORMAL MEAN ESTIMATION

Each agent i will

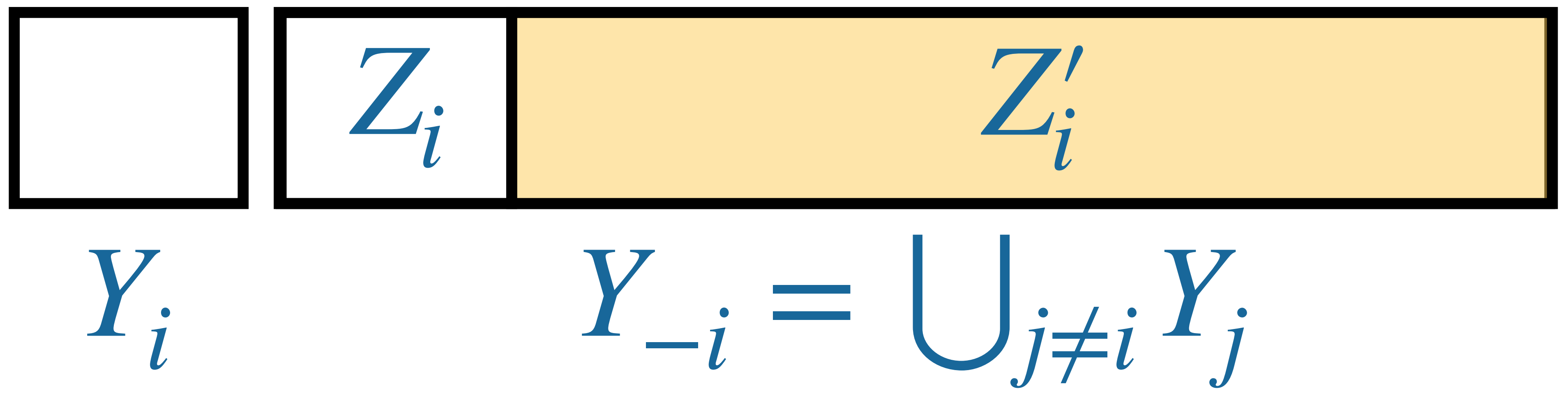
- Choose their strategy $s_i = (n_i, f_i, h_i)$
- Collect n_i points $X_i = \{x_{i,1}, \ldots, x_{i,n_i}\}$ and submit $Y_i = \{y_{i,1}, \ldots, y_{i,n_i'}\} = f_i(X_i)$.

Mechanism

- For each agent i:
  - $Z_i \leftarrow$ randomly sample $\sigma/\sqrt{cm}$ points from others' submissions Y_{-i}.
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$ # Variance proportional to difference
  - $Z_i' \leftarrow \left\{ z + \epsilon_z, \quad \text{for all } z \in Y_{-i} \backslash Z_i, \quad \text{where } \epsilon_z \sim \mathcal{N}(0, \eta_i^2) \right\}$.
  - Set allocation to each agent, $A_i \leftarrow (Z_i, Z_i', \eta_i^2)$.

Each agent i will

- Compute their estimate $h_i(X_i, Y_i, A_i)$



RECOMMENDED STRATEGIES

Mechanisms recommends that agents follow $s_i^\star = (n_i^\star, f_i^\star, h_i^\star)$,

$$n_i^\star = \frac{\sigma}{\sqrt{cm}},$$

$$f_i^\star = \text{identity},$$

$$h_i^\star\Big(X_i, Y_i, \underbrace{(Z_i, Z_i', \eta_i^2)}_{A_i}\Big) = \frac{\sum_{u \in X_i \cup Z_i} u + \frac{1}{1 + \eta_i^2/\sigma^2} \sum_{u \in Z_i'} u}{|X_i \cup Z_i| + \frac{1}{1 + \eta_i^2/\sigma^2} |Z_i'|}$$

Mechanisms recommends that agents follow $s_i^\star = (n_i^\star, f_i^\star, h_i^\star)$,

$$n_i^\star = \frac{\sigma}{\sqrt{cm}},$$

$$f_i^\star = \text{identity},$$

$$h_i^\star\Big(X_i, Y_i, \underbrace{(Z_i, Z_i', \eta_i^2)}_{A_i}\Big) = \frac{\sum_{u \in X_i \cup Z_i} u \;+\; \frac{1}{1+\eta_i^2/\sigma^2} \sum_{u \in Z_i'} u}{|X_i \cup Z_i| \;+\; \frac{1}{1+\eta_i^2/\sigma^2} |Z_i'|}$$

That is collect a sufficient amount of data $n_i^\star$, submit it truthfully $f_i^\star$, and use a weighted average estimator $h_i^\star$.

RECOMMENDED STRATEGIES

Mechanisms recommends that agents follow $s_i^\star = (n_i^\star, f_i^\star, h_i^\star)$,

$$n_i^\star = \frac{\sigma}{\sqrt{cm}},$$

$$f_i^\star = \text{identity},$$

$$h_i^\star\Big(X_i, Y_i, \underbrace{(Z_i, Z_i', \eta_i^2)}_{A_i}\Big) = \frac{\sum_{u \in X_i \cup Z_i} u + \frac{1}{1+\eta_i^2/\sigma^2} \sum_{u \in Z_i'} u}{|X_i \cup Z_i| + \frac{1}{1+\eta_i^2/\sigma^2} |Z_i'|}$$

Z_i' is the corrupted dataset.

That is collect a sufficient amount of data $n_i^\star$, submit it truthfully $f_i^\star$, and use a weighted average estimator $h_i^\star$.

Theorem: The recommended strategy profile $s^\star$ is a Nash equilibrium. Moreover, at $s^\star$, the mechanism is individually rational and approximately efficient with $P(M, s^\star) \leq 2 \cdot \inf_{M,s} P(M, s)$.

Theorem: The recommended strategy profile $s^\star$ is a Nash equilibrium. Moreover, at $s^\star$, the mechanism is individually rational and approximately efficient with $P(M, s^\star) \leq 2 \cdot \inf_{M,s} P(M, s)$.

Theorem (high-dimensional distributions with bounded variance): The recommended strategy profile $s^\star$ is an $\tilde{\mathcal{O}}(1/m)$-approximate Nash equilibrium. Moreover, the mechanism is individually rational and approximately efficient with $P(M, s^\star) \leq \left(2 + \tilde{\mathcal{O}}(1/m)\right) \cdot \inf_{M,s} P(M, s)$.

PROOF OF NASH INCENTIVE-COMPATIBILITY

PROOF OF NASH INCENTIVE-COMPATIBILITY

We need to show that $s^\star = \{(n_i^\star, f_i^\star, h_i^\star)\}_i$ is a Nash equilibrium, i.e

$$p_i(M, (s_i^\star, s_{-i}^\star)) \leq p_i\left(M, (s_i, s_{-i}^\star)\right) \quad \text{for all agents } i \text{ and all deviations } s_i$$

PROOF OF NASH INCENTIVE-COMPATIBILITY

We need to show that $s^\star = \{(n_i^\star, f_i^\star, h_i^\star)\}_i$ is a Nash equilibrium, i.e

$$p_i(M, (s_i^\star, s_{-i}^\star)) \leq p_i\left(M, (s_i, s_{-i}^\star)\right) \quad \text{for all agents } i \text{ and all deviations } s_i$$

Step 1: First, we will show that for any n_i, submitting the data truthfully and using the recommended estimator minimizes the penalty, i.e

$$p_i\left(M, \left((n_i, f_i^\star, h_i^\star), s_{-i}^\star\right)\right) \leq p_i\left(M, \left((n_i, f_i, h_i), s_{-i}^\star\right)\right) \quad \text{for all } (n_i, f_i, h_i) \in \mathbb{N} \times \mathcal{F} \times \mathcal{H}$$

PROOF OF NASH INCENTIVE-COMPATIBILITY

We need to show that $s^\star = \{(n_i^\star, f_i^\star, h_i^\star)\}_i$ is a Nash equilibrium, i.e

$$p_i(M, (s_i^\star, s_{-i}^\star)) \leq p_i\left(M, (s_i, s_{-i}^\star)\right) \quad \text{for all agents } i \text{ and all deviations } s_i$$

Step 1: First, we will show that for any n_i, submitting the data truthfully and using the recommended estimator minimizes the penalty, i.e

$$p_i\left(M,\ ((n_i, f_i^\star, h_i^\star), s_{-i}^\star)\right) \leq p_i\left(M,\ ((n_i, f_i, h_i), s_{-i}^\star)\right) \quad \text{for all } (n_i, f_i, h_i) \in \mathbb{N} \times \mathcal{F} \times \mathcal{H}$$

Step 2: Then, we will show the agent's penalty is minimized when she collects $n_i^\star$ samples under $(f_i^\star, h_i^\star)$, i.e

$$p_i\left(M,\ ((n_i^\star, f_i^\star, h_i^\star), s_{-i}^\star)\right) \leq p_i\left(M,\ ((n_i, f_i^\star, h_i^\star), s_{-i}^\star)\right) \quad \text{for all } n_i \in \mathbb{N}$$

PROOF OF STEP 1

Step 1: First, we will show that for any amount of data collected n_i, submitting it truthfully and using the recommended estimator minimizes the penalty, i.e

$$p_i\left(M, \left((n_i, f_i^\star, h_i^\star), s_{-i}^\star\right)\right) \leq p_i\left(M, \left((n_i, f_i, h_i), s_{-i}^\star\right)\right) \quad \text{for all } (n_i, f_i, h_i) \in \mathbb{N} \times \mathcal{F} \times \mathcal{H}$$

PROOF OF STEP 1

Step 1: First, we will show that for any amount of data collected n_i, submitting it truthfully and using the recommended estimator minimizes the penalty, i.e

$$p_i\Big(M,\ \big((n_i, f_i^\star, h_i^\star), s_{-i}^\star\big)\Big) \le p_i\Big(M,\ \big((n_i, f_i, h_i), s_{-i}^\star\big)\Big) \quad \text{for all } (n_i, f_i, h_i) \in \mathbb{N} \times \mathscr{F} \times \mathscr{H}$$

We need to show, for all $(n_i, f_i, h_i) \in \mathbb{N} \times \mathscr{F} \times \mathscr{H}$,

$$\sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star\left(X_i, f_i^\star(X_i), A_i\right) - \mu\right)^2 \right] + cn_i \;\le\; \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \right] + cn_i$$

PROOF OF STEP 1

Step 1: First, we will show that for any amount of data collected n_i, submitting it truthfully and using the recommended estimator minimizes the penalty, i.e

$$p_i\left(M,\ ((n_i, f_i^\star, h_i^\star), s_{-i}^\star)\right) \leq p_i\left(M,\ ((n_i, f_i, h_i), s_{-i}^\star)\right) \quad \text{for all } (n_i, f_i, h_i) \in \mathbb{N} \times \mathscr{F} \times \mathscr{H}$$

We need to show, for all $(n_i, f_i, h_i) \in \mathbb{N} \times \mathscr{F} \times \mathscr{H}$,

$$\sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star\left(X_i, f_i^\star(X_i), A_i\right) - \mu\right)^2\right] + \not{c}n_i \;\leq\; \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] + \not{c}n_i$$

Step 1: First, we will show that for any amount of data collected n_i, submitting it truthfully and using the recommended estimator minimizes the penalty, i.e

$$p_i\left(M,\ ((n_i, f_i^\star, h_i^\star), s_{-i}^\star)\right) \leq p_i\left(M,\ ((n_i, f_i, h_i), s_{-i}^\star)\right) \quad \text{for all } (n_i, f_i, h_i) \in \mathbb{N} \times \mathscr{F} \times \mathscr{H}$$

We need to show, for all $(n_i, f_i, h_i) \in \mathbb{N} \times \mathscr{F} \times \mathscr{H}$,

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_\mu\left[\left(h_i^\star\left(X_i, f_i^\star(X_i), A_i\right) - \mu\right)^2\right] + \cancel{c}n_i \ \leq\ \sup_{\mu\in\mathbb{R}} \mathbb{E}_\mu\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] + \cancel{c}n_i$$

Or equivalently,

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_\mu\left[\left(h_i^\star\left(X_i, f_i^\star(X_i), A_i\right) - \mu\right)^2\right] \ =\ \inf_{f_i, h_i}\ \sup_{\mu\in\mathbb{R}}\ \mathbb{E}_\mu\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right]$$

DIGRESSION: MINIMAX NORMAL MEAN ESTIMATION

We are given $X_1^n = \{X_1, \ldots, X_n\}$, drawn i.i.d from $\mathcal{N}(\mu, \sigma^2)$ where σ^2 is known. Let $h(X_1^n)$ be an estimator for μ. We wish to show

DIGRESSION: MINIMAX NORMAL MEAN ESTIMATION

We are given $X_1^n = \{X_1, \ldots, X_n\}$, drawn i.i.d from $\mathcal{N}(\mu, \sigma^2)$ where σ^2 is known. Let $h(X_1^n)$ be an estimator for μ. We wish to show

$$\text{minimax risk} = \inf_{\hat{\mu}} \sup_{\mu \in \mathbb{R}} \mathbb{E}_{X_1^n}\left[\left(\mu - h(X_1^n)\right)^2\right] = \frac{\sigma^2}{n}$$

DIGRESSION: MINIMAX NORMAL MEAN ESTIMATION

We are given $X_1^n = \{X_1, \ldots, X_n\}$, drawn i.i.d from $\mathcal{N}(\mu, \sigma^2)$ where σ^2 is known. Let $h(X_1^n)$ be an estimator for μ. We wish to show

$$\text{minimax risk} = \inf_{\hat{\mu}} \sup_{\mu \in \mathbb{R}} \mathbb{E}_{X_1^n}\left[\left(\mu - h(X_1^n)\right)^2\right] = \frac{\sigma^2}{n}$$

Upper bound via an estimator: We can use the sample mean $h_{\text{sm}}(X) = (X_1 + \ldots + X_n)/n$.

We are given $X_1^n = \{X_1, \ldots, X_n\}$, drawn i.i.d from $\mathcal{N}(\mu, \sigma^2)$ where σ^2 is known. Let $h(X_1^n)$ be an estimator for μ. We wish to show

$$\text{minimax risk} = \inf_{\hat{\mu}} \sup_{\mu \in \mathbb{R}} \mathbb{E}_{X_1^n}\left[\left(\mu - h(X_1^n)\right)^2\right] = \frac{\sigma^2}{n}$$

Upper bound via an estimator: We can use the sample mean $h_{\text{sm}}(X) = (X_1 + \ldots + X_n)/n$.

$$\text{minimax risk} \leq \sup_{\mu \in \mathbb{R}} \mathbb{E}_{X_1^n}\left[\left(\mu - h_{\text{sm}}(X_1^n)\right)^2\right] = \frac{\sigma^2}{n}$$

DIGRESSION: MINIMAX RISK FOR NORMAL MEAN ESTIMATION

Lower bound via Bayes' risk: Choose a prior Λ for μ. Then lower bound via the Bayes' risk under Λ.

Lower bound via Bayes' risk: Choose a prior Λ for μ. Then lower bound via the Bayes' risk under Λ.

We will use $\Lambda = \mathcal{N}(0, \tau^2)$. Then, for any estimator h,

Lower bound via Bayes' risk: Choose a prior Λ for μ. Then lower bound via the Bayes' risk under Λ.

We will use $\Lambda = \mathcal{N}(0, \tau^2)$. Then, for any estimator h,

$$\sup_{\mu \in \mathbb{R}} \mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2\right] \geq \mathbb{E}_{\mu \sim \Lambda}\left[\mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2 \mid \mu\right]\right]$$

← sup ≥ avg

DIGRESSION: MINIMAX RISK FOR NORMAL MEAN ESTIMATION

Lower bound via Bayes' risk: Choose a prior Λ for μ. Then lower bound via the Bayes' risk under Λ.

We will use $\Lambda = \mathcal{N}(0,\tau^2)$. Then, for any estimator h,

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2 \mid \mu\right]\right]$$

$$= \mathbb{E}_{X_1^n}\left[\mathbb{E}_{\mu\sim\Lambda}\left[(\mu - h(X_1^n))^2 \mid X_1^n\right]\right]$$

← sup $\geq$ avg

← swap order of expectation

DIGRESSION: MINIMAX RISK FOR NORMAL MEAN ESTIMATION

Lower bound via Bayes' risk: Choose a prior Λ for μ. Then lower bound via the Bayes' risk under Λ.

We will use $\Lambda = \mathcal{N}(0,\tau^2)$. Then, for any estimator h,

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2 \,|\, \mu\right]\right]$$

← sup ≥ avg

$$= \mathbb{E}_{X_1^n}\left[\mathbb{E}_{\mu\sim\Lambda}\left[(\mu - h(X_1^n))^2 \,|\, X_1^n\right]\right]$$

← swap order of expectation

Now, minimize inner expectation w.r.t h.

(i) As μ, X_1^n is jointly Gaussian, $\mu \,|\, X_1^n$ is also Gaussian.

(ii) Then choose $h =$ posterior mean.

Lower bound via Bayes' risk: Choose a prior Λ for μ. Then lower bound via the Bayes' risk under Λ.

We will use $\Lambda = \mathcal{N}(0,\tau^2)$. Then, for any estimator h,

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2 \,|\, \mu\right]\right]$$ ← sup ≥ avg

$$= \mathbb{E}_{X_1^n}\left[\mathbb{E}_{\mu\sim\Lambda}\left[(\mu - h(X_1^n))^2 \,|\, X_1^n\right]\right]$$ ← swap order of expectation

Now, minimize inner expectation w.r.t h.

(i) As μ, X_1^n is jointly Gaussian, $\mu \,|\, X_1^n$ is also Gaussian.

(ii) Then choose $h =$ posterior mean.

$$\geq \mathbb{E}_{X_1^n}\left[\frac{\sigma^2}{n + \sigma^2/\tau^2}\right]$$

Lower bound via Bayes' risk: Choose a prior Λ for μ. Then lower bound via the Bayes' risk under Λ.

We will use $\Lambda = \mathcal{N}(0, \tau^2)$. Then, for any estimator h,

$$\sup_{\mu \in \mathbb{R}} \mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{X_1^n}\left[(\mu - h(X_1^n))^2 \,|\, \mu\right]\right]$$

← sup ≥ avg

$$= \mathbb{E}_{X_1^n}\left[\mathbb{E}_{\mu\sim\Lambda}\left[(\mu - h(X_1^n))^2 \,|\, X_1^n\right]\right]$$

← swap order of expectation

Now, minimize inner expectation w.r.t h.

(i) As μ, X_1^n is jointly Gaussian, $\mu \,|\, X_1^n$ is also Gaussian.

(ii) Then choose $h =$ posterior mean.

$$\geq \mathbb{E}_{X_1^n}\left[\frac{\sigma^2}{n + \sigma^2/\tau^2}\right]$$

$$= \frac{\sigma^2}{n + \sigma^2/\tau^2} \quad \to \frac{\sigma^2}{n} \qquad \text{as } \tau^2 \to \infty$$

PROOF OF STEP 1

We will apply the same recipe to prove step 1,

$$\inf_{f_i, h_i} \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i \left(X_i, f_i(X_i), A_i \right) - \mu \right)^2 \right] = \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star \left(X_i, f_i^\star(X_i), A_i \right) - \mu \right)^2 \right]$$

PROOF OF STEP 1

We will apply the same recipe to prove step 1,

$$\inf_{f_i, h_i} \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i \left(X_i, f_i(X_i), A_i \right) - \mu \right)^2 \right] = \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star \left(X_i, f_i^\star(X_i), A_i \right) - \mu \right)^2 \right]$$

Two challenges:

1. Not just the estimator h_i but also the submission function f_i.

PROOF OF STEP 1

We will apply the same recipe to prove step 1,

$$\inf_{f_i, h_i} \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i \left(X_i, f_i(X_i), A_i \right) - \mu \right)^2 \right] = \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star \left(X_i, f_i^\star(X_i), A_i \right) - \mu \right)^2 \right]$$

Two challenges:

1. Not just the estimator h_i but also the submission function f_i.
2. The data available to the agent is not i.i.d!
   - The corruption is data-dependent.

PROOF OF STEP 1

We will apply the same recipe to prove step 1,

$$\inf_{f_i,h_i}\sup_{\mu\in\mathbb{R}} \mathbb{E}_\mu\left[\left(h_i\left(X_i, f_i(X_i), A_i\right)-\mu\right)^2\right] = \sup_{\mu\in\mathbb{R}}\mathbb{E}_\mu\left[\left(h_i^\star\left(X_i, f_i^\star(X_i), A_i\right)-\mu\right)^2\right]$$

Two challenges:

1. Not just the estimator h_i but also the submission function f_i.
2. The data available to the agent is not i.i.d!
   - The corruption is data-dependent.

- Set noise variance $\eta_i^2 = \alpha^2\left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$ # V
- $Z_i' \leftarrow \left\{z + \epsilon_z, \quad \text{for all } z \in Y_{-i}\backslash Z_i, \quad \text{where } \epsilon_z \sim \mathcal{N}(0, \eta_i^2)\right\}$.

PROOF OF STEP 1

We will apply the same recipe to prove step 1,

$$\inf_{f_i, h_i} \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i \left(X_i, f_i(X_i), A_i \right) - \mu \right)^2 \right] = \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star \left(X_i, f_i^\star(X_i), A_i \right) - \mu \right)^2 \right]$$

Two challenges:

1. Not just the estimator h_i but also the submission function f_i.
2. The data available to the agent is not i.i.d!
   - The corruption is data-dependent.
   - In fact, X_i, Z_i, Z_i' is not even jointly Gaussian.

- Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i) \right)^2$ # V
- $Z_i' \leftarrow \left\{ z + \epsilon_z, \quad \text{for all } z \in Y_{-i} \backslash Z_i, \quad \text{where } \epsilon_z \sim \mathcal{N}(0, \eta_i^2) \right\}.$

PROOF OF STEP 1: UPPER BOUND

We show

$$\inf_{f_i, h_i} \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i \left(X_i, f_i(X_i), A_i \right) - \mu \right)^2 \right] \leq \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star \left(X_i, f_i^\star(X_i), A_i \right) - \mu \right)^2 \right]$$

PROOF OF STEP 1: UPPER BOUND

We show

$$\inf_{f_i, h_i} \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i \left(X_i, f_i(X_i), A_i \right) - \mu \right)^2 \right] \leq \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star \left(X_i, f_i^\star(X_i), A_i \right) - \mu \right)^2 \right]$$

$$= \mathbb{E}_{Z \sim \mathcal{N}(0,1)} \left[\left(\frac{(m-2) n_i^\star}{\left(\sigma^2 + \alpha^2 \left(\sigma^2 / n_i + \sigma^2 / n_i^\star \right) Z^2 \right)} + \frac{n_i + n_i^\star}{\sigma^{-2}} \right)^{-1} \right]$$

PROOF OF STEP 1: UPPER BOUND

We show

$$\inf_{f_i, h_i} \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i \left(X_i, f_i(X_i), A_i \right) - \mu \right)^2 \right] \leq \sup_{\mu \in \mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star \left(X_i, f_i^\star(X_i), A_i \right) - \mu \right)^2 \right]$$

$$= \mathbb{E}_{Z \sim \mathcal{N}(0,1)} \left[\left(\frac{(m-2) n_i^\star}{\left(\sigma^2 + \alpha^2 \left(\sigma^2 / n_i + \sigma^2 / n_i^\star \right) Z^2 \right)} + \frac{n_i + n_i^\star}{\sigma^{-2}} \right)^{-1} \right] =: R_\infty(n_i) \quad \text{(say)}$$

PROOF OF STEP 1: UPPER BOUND

We show

$$\inf_{f_i,h_i} \sup_{\mu\in\mathbb{R}} \mathbb{E}_\mu \left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu \right)^2 \right] \le \sup_{\mu\in\mathbb{R}} \mathbb{E}_\mu \left[\left(h_i^\star\left(X_i, f_i^\star(X_i), A_i\right) - \mu \right)^2 \right]$$

$$= \mathbb{E}_{Z\sim\mathcal{N}(0,1)} \left[\left(\frac{(m-2)n_i^\star}{\left(\sigma^2 + \alpha^2\left(\sigma^2/n_i + \sigma^2/n_i^\star\right) Z^2\right)} + \frac{n_i + n_i^\star}{\sigma^{-2}} \right)^{-1} \right] =: R_\infty(n_i) \quad \text{(say)}$$

Proof idea:

- When $f_i^\star =$ identity, first condition on X_i, Z_i, then $Z_i' \sim \mathcal{N}(0, \sigma^2 + \eta^2)$.

PROOF OF STEP 1: LOWER BOUND

Choose prior $\Lambda = \mathcal{N}(0, \tau^2)$ for μ. Then for any f_i, h_i, we have

PROOF OF STEP 1: LOWER BOUND

Choose prior $\Lambda = \mathcal{N}(0, \tau^2)$ for μ. Then for any f_i, h_i, we have

$$\sup_{\mu \in \mathbb{R}} \mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \mu\right]\right]$$

← sup ≥ avg

PROOF OF STEP 1: LOWER BOUND

Choose prior $\Lambda = \mathcal{N}(0, \tau^2)$ for μ. Then for any f_i, h_i, we have

$$\sup_{\mu \in \mathbb{R}} \mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \mu\right]\right]$$

← sup ≥ avg

$$= \mathbb{E}_{\text{data}\sim\mu}\left[\mathbb{E}_{\mu\sim\Lambda}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \text{data}\right]\right]$$

← Swap order of expectation

PROOF OF STEP 1: LOWER BOUND

Choose prior $\Lambda = \mathcal{N}(0,\tau^2)$ for μ. Then for any f_i, h_i, we have

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \mu\right]\right]$$

← sup ≥ avg

$$= \mathbb{E}_{\text{data}\sim\mu}\left[\mathbb{E}_{\mu\sim\Lambda}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \text{data}\right]\right]$$

← Swap order of expectation

Choose h_i = posterior mean to minimize w.r.t h_i.

PROOF OF STEP 1: LOWER BOUND

Choose prior $\Lambda = \mathcal{N}(0,\tau^2)$ for μ. Then for any f_i, h_i, we have

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \mu\right]\right]$$

← sup ≥ avg

$$= \mathbb{E}_{\text{data}\sim\mu}\left[\mathbb{E}_{\mu\sim\Lambda}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \text{data}\right]\right]$$

← Swap order of expectation

Choose h_i = posterior mean to minimize w.r.t h_i.

- μ, X_i, Z_i, Z_i' is not jointly Gaussian, but $\mu \mid X_i, Z_i, Z_i'$ is Gaussian.

PROOF OF STEP 1: LOWER BOUND

Choose prior $\Lambda = \mathcal{N}(0,\tau^2)$ for μ. Then for any f_i, h_i, we have

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \mu\right]\right]$$

← sup ≥ avg

$$= \mathbb{E}_{\text{data}\sim\mu}\left[\mathbb{E}_{\mu\sim\Lambda}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \text{data}\right]\right]$$

← Swap order of expectation

Choose h_i = posterior mean to minimize w.r.t h_i.

- μ, X_i, Z_i, Z_i' is not jointly Gaussian, but $\mu \mid X_i, Z_i, Z_i'$ is Gaussian.

$$\geq \mathbb{E}_{\text{data}}\left[\left(|Z_i'|\left(\sigma^2+\alpha^2\left(\frac{1}{|f_i(X_i)|}\sum_{y\in f_i(X_i)} y - \frac{1}{|Z_i|}\sum_{z\in Z_i} z\right)^2\right)^{-1} + \frac{|X_i|+|Z_i|}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}\right]$$

PROOF OF STEP 1: LOWER BOUND

Choose prior $\Lambda = \mathcal{N}(0,\tau^2)$ for μ. Then for any f_i, h_i, we have

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] \;\geq\; \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \mu\right]\right]$$

← sup ≥ avg

$$= \mathbb{E}_{\text{data}\sim\mu}\left[\mathbb{E}_{\mu\sim\Lambda}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \text{data}\right]\right]$$

← Swap order of expectation

Choose h_i = posterior mean to minimize w.r.t h_i.

- μ, X_i, Z_i, Z_i' is not jointly Gaussian, but $\mu \,|\, X_i, Z_i, Z_i'$ is Gaussian.

$$\geq \mathbb{E}_{\text{data}}\left[\left(|Z_i'|\left(\sigma^2 + \alpha^2\left(\frac{1}{|f_i(X_i)|}\sum_{y\in f_i(X_i)} y - \frac{1}{|Z_i|}\sum_{z\in Z_i} z\right)^2\right)^{-1} + \frac{|X_i| + |Z_i|}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}\right]$$

$= \;\ldots\; = R_\tau(n_i)$ (say) ← To minimize w.r.t f_i, choose $f_i(X_i) = \left\{\left(1 + \sigma^2/(|X|\tau^2)\right)^{-1} x, \forall x \in X_i\right\}$ and apply Hardy-Littlewood inequality.

PROOF OF STEP 1: LOWER BOUND

Choose prior $\Lambda = \mathcal{N}(0,\tau^2)$ for μ. Then for any f_i, h_i, we have

$$\sup_{\mu\in\mathbb{R}} \mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2\right] \geq \mathbb{E}_{\mu\sim\Lambda}\left[\mathbb{E}_{\text{data}\sim\mu}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \mu\right]\right]$$

← sup ≥ avg

$$= \mathbb{E}_{\text{data}\sim\mu}\left[\mathbb{E}_{\mu\sim\Lambda}\left[\left(h_i\left(X_i, f_i(X_i), A_i\right) - \mu\right)^2 \middle| \text{data}\right]\right]$$

← Swap order of expectation

Choose $h_i =$ posterior mean to minimize w.r.t h_i.

- μ, X_i, Z_i, Z_i' is not jointly Gaussian, but $\mu \,|\, X_i, Z_i, Z_i'$ is Gaussian.

$$\geq \mathbb{E}_{\text{data}}\left[\left(\left(|Z_i'|\left(\sigma^2 + \alpha^2\left(\frac{1}{|f_i(X_i)|}\sum_{y\in f_i(X_i)} y - \frac{1}{|Z_i|}\sum_{z\in Z_i} z\right)^2\right)\right)^{-1} + \frac{|X_i| + |Z_i|}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}\right]$$

$= \ldots = R_\tau(n_i)$ (say) ← To minimize w.r.t f_i, choose $f_i(X_i) = \left\{\left(1 + \sigma^2/(|X|\tau^2)\right)^{-1} x, \forall x \in X_i\right\}$ and apply Hardy-Littlewood inequality.

$\to R_\infty(n_i)$ as $\tau \to \infty$

PROOF OF STEP 2

Step 2: Then, we will show the agent's penalty is minimized when she collects n_i samples under $(f_i^\star, h_i^\star)$, i.e

$$p_i\left(M,\ \left((n_i^\star, f_i^\star, h_i^\star), s_{-i}^\star\right)\right) \leq p_i\left(M,\ \left((n_i, f_i^\star, h_i^\star), s_{-i}^\star\right)\right) \quad \text{for all } n_i \in \mathbb{N}$$

Step 2: Then, we will show the agent's penalty is minimized when she collects n_i samples under $(f_i^\star, h_i^\star)$, i.e

$$p_i\left(M,\ ((n_i^\star, f_i^\star, h_i^\star), s_{-i}^\star)\right) \le p_i\left(M,\ ((n_i, f_i^\star, h_i^\star), s_{-i}^\star)\right) \quad \text{for all } n_i \in \mathbb{N}$$

From Step 1 we have,

$$\text{RHS} = p_i\left(M,\ ((n_i, f_i^\star, h_i^\star), s_{-i}^\star)\right) = \mathbb{E}_{Z\sim\mathcal{N}(0,1)}\left[\left(\frac{(m-2)n_i^\star}{\left(\sigma^2 + \alpha^2\left(\sigma^2/n_i + \sigma^2/n_i^\star\right) Z^2\right)} + \frac{n_i + n_i^\star}{\sigma^{-2}}\right)^{-1}\right] + cn_i$$

PROOF OF STEP 2

Step 2: Then, we will show the agent's penalty is minimized when she collects n_i samples under $(f_i^\star, h_i^\star)$, i.e

$$p_i\left(M, \left((n_i^\star, f_i^\star, h_i^\star), s_{-i}^\star\right)\right) \leq p_i\left(M, \left((n_i, f_i^\star, h_i^\star), s_{-i}^\star\right)\right) \quad \text{for all } n_i \in \mathbb{N}$$

From Step 1 we have,

$$\text{RHS} = p_i\left(M, \left((n_i, f_i^\star, h_i^\star), s_{-i}^\star\right)\right) = \mathbb{E}_{Z\sim\mathcal{N}(0,1)}\left[\left(\frac{(m-2)n_i^\star}{\left(\sigma^2 + \alpha^2\left(\sigma^2/n_i + \sigma^2/n_i^\star\right)Z^2\right)} + \frac{n_i + n_i^\star}{\sigma^{-2}}\right)^{-1}\right] + cn_i$$

- The term inside $\mathbb{E}$ is convex in n_i. Hence so is $p_i\left(M, \left((n_i, f_i^\star, h_i^\star), s_{-i}^\star\right)\right)$.
- Minimized at $n_i = n_i^\star$ (by our choice of α).

CHOICE OF α AND BOUNDING THE SOCIAL PENALTY

- For each agent i:
  - $Z_i \leftarrow$ sample $n^\star = \sigma/\sqrt{cm}$ points from others' subm
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$

CHOICE OF α AND BOUNDING THE SOCIAL PENALTY

- For each agent i:
  - $Z_i \leftarrow$ sample $n^\star = \sigma/\sqrt{cm}$ points from others' subm
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$

We set α to be the smallest number larger than $\sqrt{n_i^\star}$ such that $G(\alpha) = 0$, where,

CHOICE OF α AND BOUNDING THE SOCIAL PENALTY

- For each agent i:
  - $Z_i \leftarrow$ sample $n^\star = \sigma/\sqrt{cm}$ points from others' subm
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$

We set α to be the smallest number larger than $\sqrt{n_i^\star}$ such that $G(\alpha) = 0$, where,

$$G(\alpha) := \left(\frac{m-4}{m-2}\frac{4\alpha^2}{\sigma/\sqrt{cm}} - 1\right)\frac{4\alpha}{\sqrt{\sigma}(m/c)^{1/4}} - \left(4(m+1)\frac{\alpha^2}{\sigma\sqrt{m/c}} - 1\right)\sqrt{2\pi}\exp\left(\frac{\sigma\sqrt{m/c}}{8\alpha^2}\right)\mathrm{Erfc}\left(\frac{\sqrt{\sigma}(m/c)^{1/4}}{2\sqrt{2}\alpha}\right)$$

CHOICE OF α AND BOUNDING THE SOCIAL PENALTY

- For each agent i:
  - $Z_i \leftarrow$ sample $n^\star = \sigma/\sqrt{cm}$ points from others' subm
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$

We set α to be the smallest number larger than $\sqrt{n_i^\star}$ such that $G(\alpha) = 0$, where,

$$G(\alpha) := \left(\frac{m-4}{m-2}\frac{4\alpha^2}{\sigma/\sqrt{cm}} - 1\right)\frac{4\alpha}{\sqrt{\sigma}(m/c)^{1/4}} - \left(4(m+1)\frac{\alpha^2}{\sigma\sqrt{m/c}} - 1\right)\sqrt{2\pi}\exp\left(\frac{\sigma\sqrt{m/c}}{8\alpha^2}\right)\mathrm{Erfc}\left(\frac{\sqrt{\sigma}(m/c)^{1/4}}{2\sqrt{2}\alpha}\right)$$

- $G(\alpha) = 0$: step 2 of NIC (collect a sufficient amount of data).

CHOICE OF α AND BOUNDING THE SOCIAL PENALTY

- For each agent i:
  - $Z_i \leftarrow$ sample $n^\star = \sigma/\sqrt{cm}$ points from others' subm
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$

We set α to be the smallest number larger than $\sqrt{n_i^\star}$ such that $G(\alpha) = 0$, where,

$$G(\alpha) := \left(\frac{m-4}{m-2}\frac{4\alpha^2}{\sigma/\sqrt{cm}} - 1\right)\frac{4\alpha}{\sqrt{\sigma}(m/c)^{1/4}} - \left(4(m+1)\frac{\alpha^2}{\sigma\sqrt{m/c}} - 1\right)\sqrt{2\pi}\exp\left(\frac{\sigma\sqrt{m/c}}{8\alpha^2}\right)\mathrm{Erfc}\left(\frac{\sqrt{\sigma}(m/c)^{1/4}}{2\sqrt{2}\alpha}\right)$$

- $G(\alpha) = 0$: step 2 of NIC (collect a sufficient amount of data).
- $\alpha^2 \geq n_i^\star$: step 1 of NIC (sufficiently penalize untruthful agents).

CHOICE OF α AND BOUNDING THE SOCIAL PENALTY

- For each agent i:
  - $Z_i \leftarrow$ sample $n^\star = \sigma/\sqrt{cm}$ points from others' subm
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$

We set α to be the smallest number larger than $\sqrt{n_i^\star}$ such that $G(\alpha) = 0$, where,

$$G(\alpha) := \left(\frac{m-4}{m-2}\frac{4\alpha^2}{\sigma/\sqrt{cm}} - 1\right)\frac{4\alpha}{\sqrt{\sigma}(m/c)^{1/4}} - \left(4(m+1)\frac{\alpha^2}{\sigma\sqrt{m/c}} - 1\right)\sqrt{2\pi}\exp\left(\frac{\sigma\sqrt{m/c}}{8\alpha^2}\right)\operatorname{Erfc}\left(\frac{\sqrt{\sigma}(m/c)^{1/4}}{2\sqrt{2}\alpha}\right)$$

- $G(\alpha) = 0$: step 2 of NIC (collect a sufficient amount of data).
- $\alpha^2 \geq n_i^\star$: step 1 of NIC (sufficiently penalize untruthful agents).
- "smallest number larger than": for efficiency (don't over-penalize truthful agents).

CHOICE OF α AND BOUNDING THE SOCIAL PENALTY

- For each agent i:
  - $Z_i \leftarrow$ sample $n^\star = \sigma/\sqrt{cm}$ points from others' subm
  - Set noise variance $\eta_i^2 = \alpha^2 \left(\text{mean}(Y_i) - \text{mean}(Z_i)\right)^2$

We set α to be the smallest number larger than $\sqrt{n_i^\star}$ such that $G(\alpha) = 0$, where,

$$G(\alpha) := \left(\frac{m-4}{m-2}\frac{4\alpha^2}{\sigma/\sqrt{cm}} - 1\right)\frac{4\alpha}{\sqrt{\sigma}(m/c)^{1/4}} - \left(4(m+1)\frac{\alpha^2}{\sigma\sqrt{m/c}} - 1\right)\sqrt{2\pi}\exp\left(\frac{\sigma\sqrt{m/c}}{8\alpha^2}\right)\mathrm{Erfc}\left(\frac{\sqrt{\sigma}(m/c)^{1/4}}{2\sqrt{2}\alpha}\right)$$

- $G(\alpha) = 0$: step 2 of NIC (collect a sufficient amount of data).
- $\alpha^2 \geq n_i^\star$: step 1 of NIC (sufficiently penalize untruthful agents).
- “smallest number larger than”: for efficiency (don't over-penalize truthful agents).

+ Wolfram Mathematica

OUTLINE

1. Mechanism design for collaborative normal mean estimation

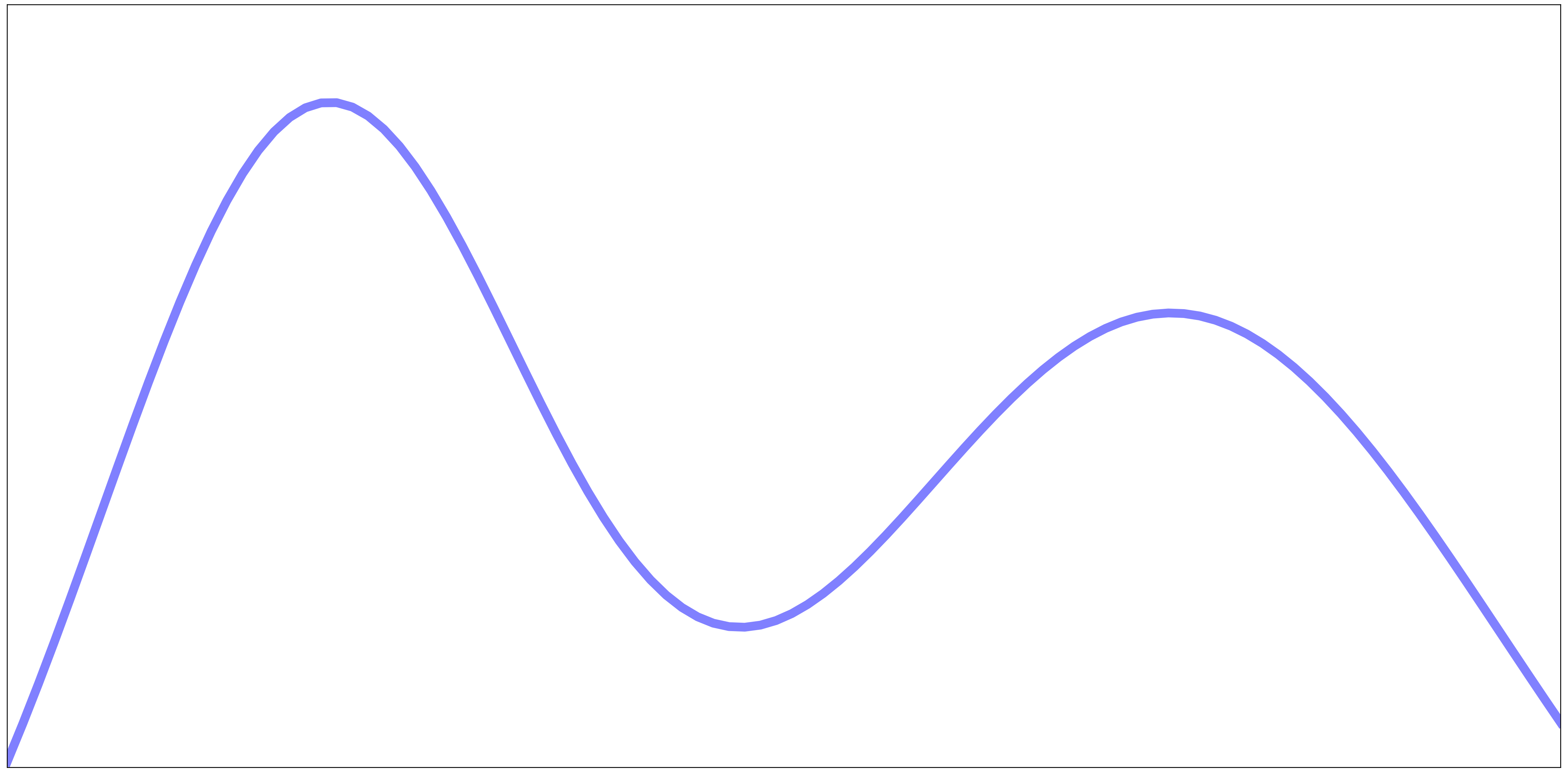
(Chen, Zhu, Kandasamy, *NeurIPS 2023*)

   - Intuitions, overview of results
   - Problem formalism
   - Mechanism and theoretical analysis

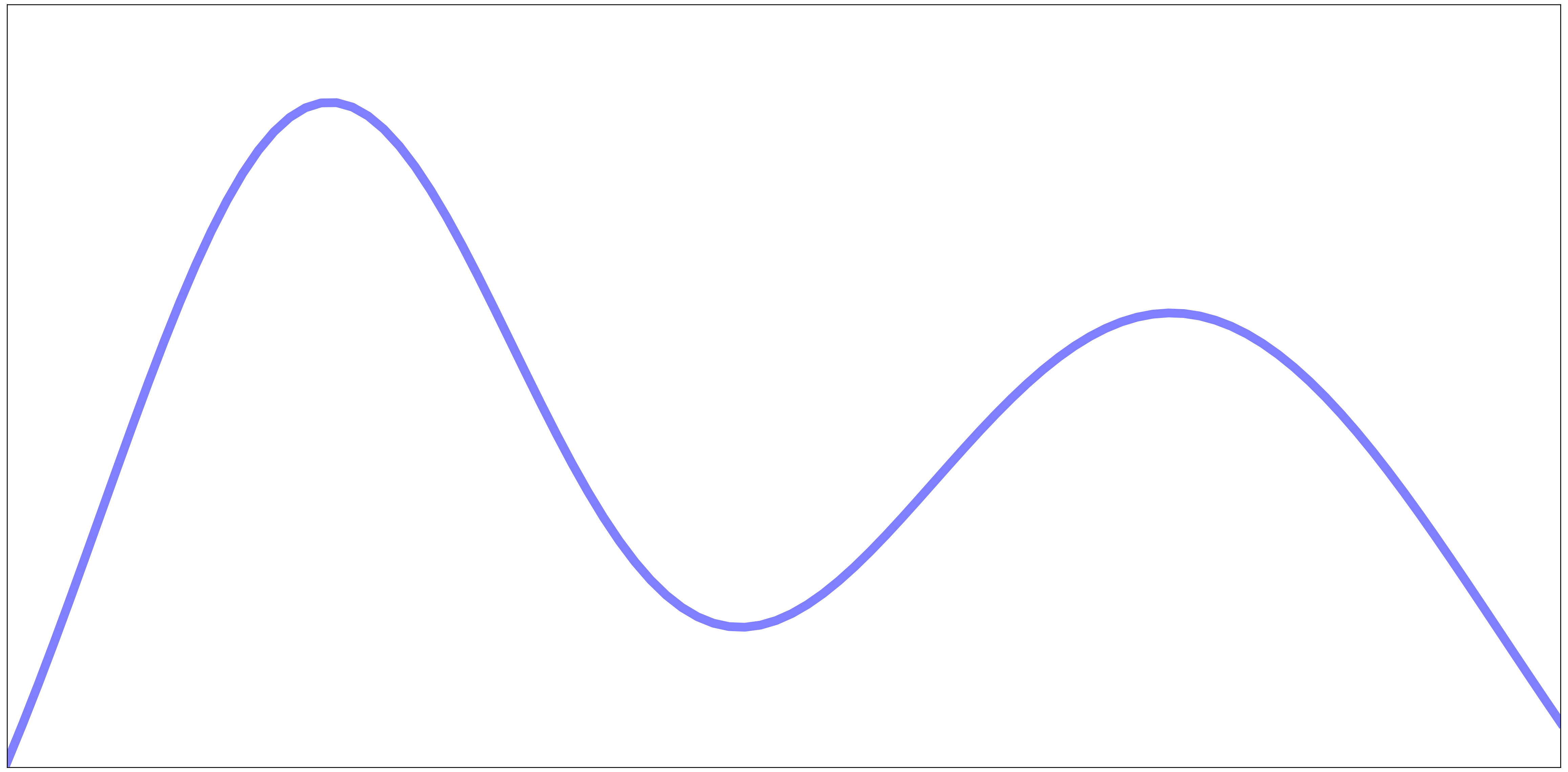
2. **Extensions** **(Clinton, Chen, Zhu, Kandasamy, *Ongoing work*)**

   - **Multiple distributions with asymmetric data collection capabilities**
   - **Collaborative supervised learning and experiment design**

COLLABORATIVE SUPERVISED LEARNING AND EXPERIMENT DESIGN

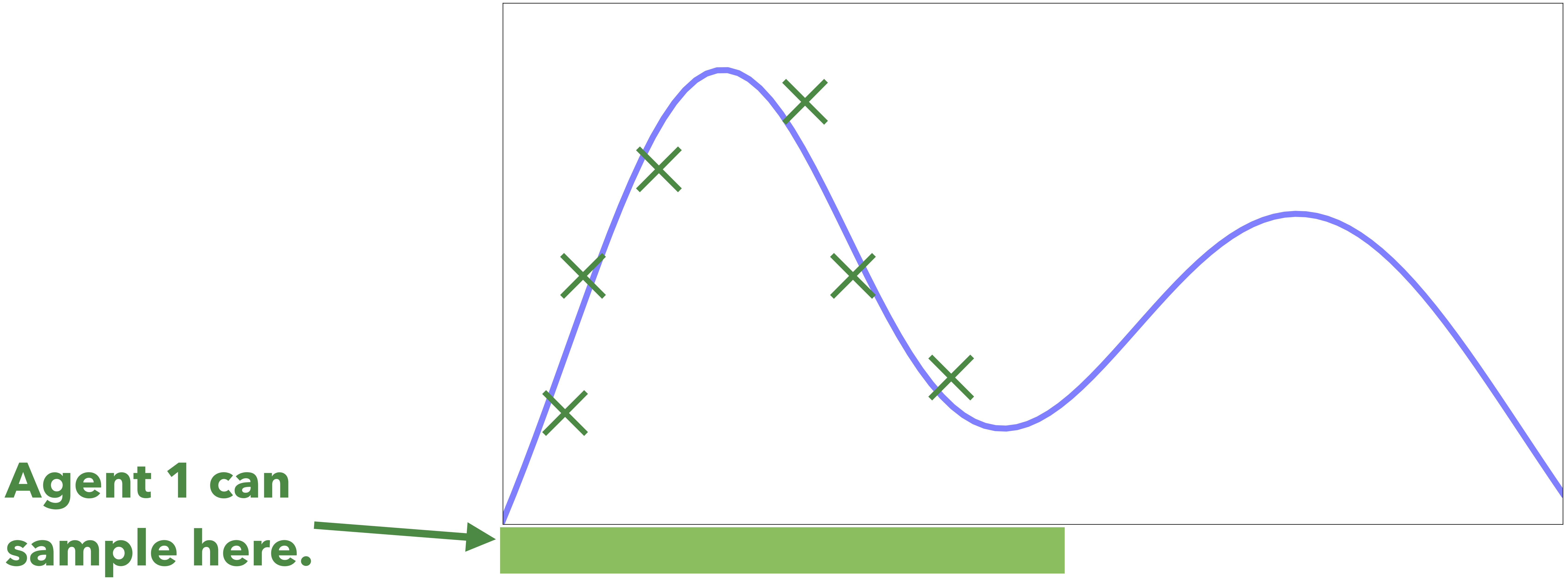


COLLABORATIVE SUPERVISED LEARNING AND EXPERIMENT DESIGN



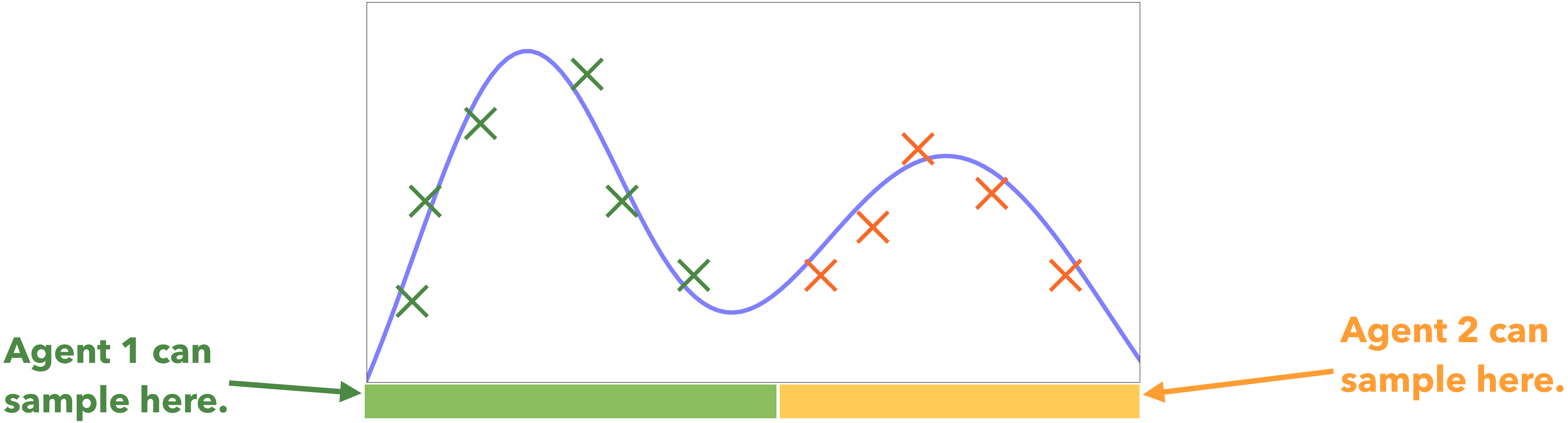
Data sharing when there is asymmetric data collection capabilities.

COLLABORATIVE SUPERVISED LEARNING AND EXPERIMENT DESIGN

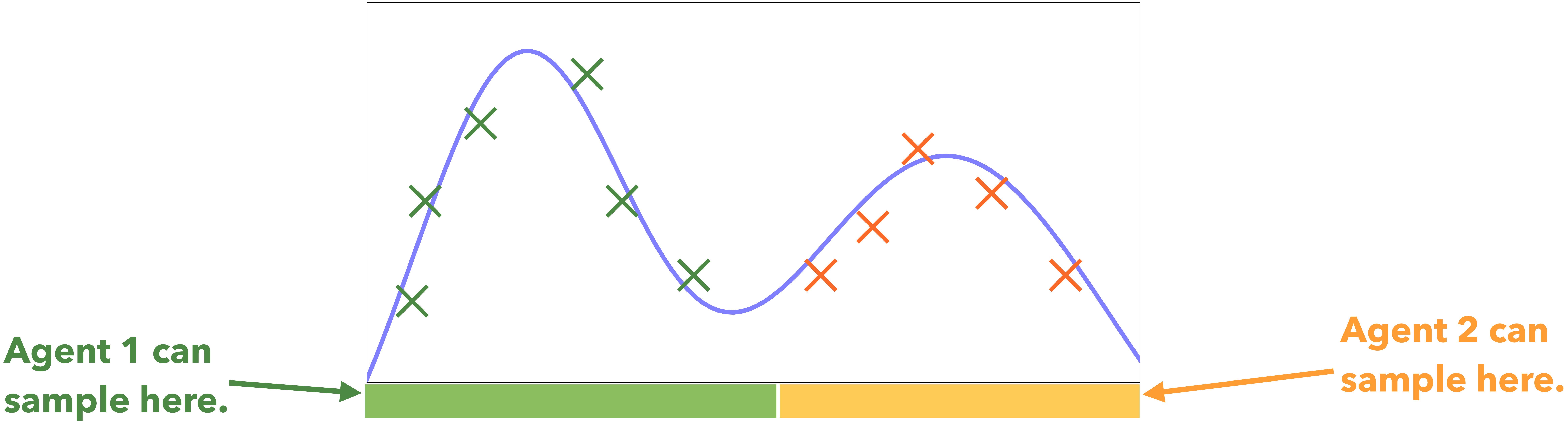


Data sharing when there is asymmetric data collection capabilities.

COLLABORATIVE SUPERVISED LEARNING AND EXPERIMENT DESIGN



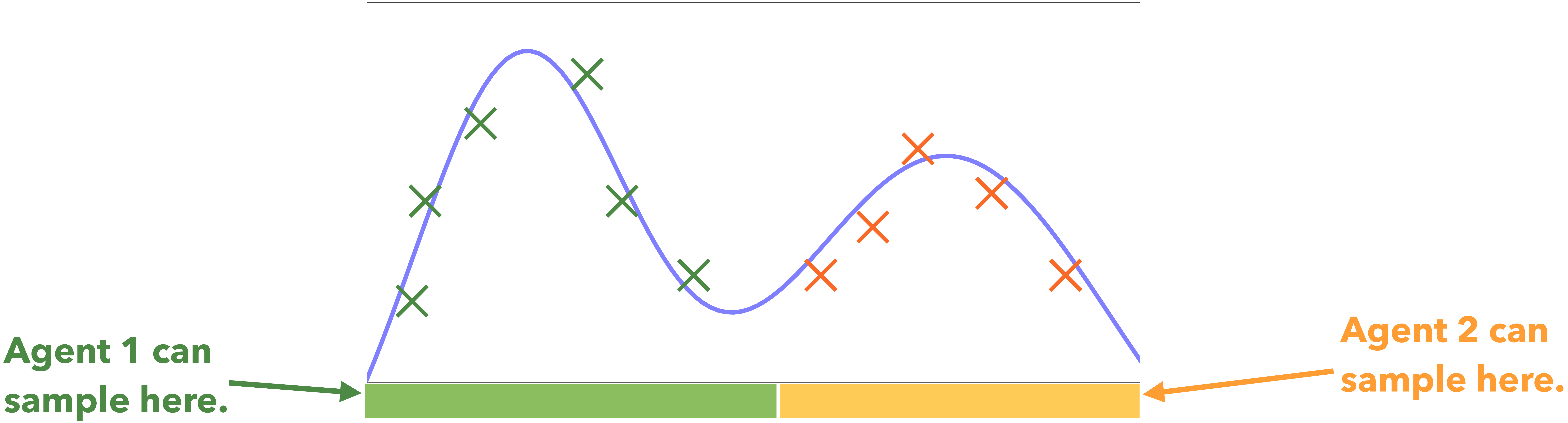
Data sharing when there is asymmetric data collection capabilities.



Data sharing when there is asymmetric data collection capabilities.

E.g: hospitals in different locations, researchers with different experimental equipment etc.

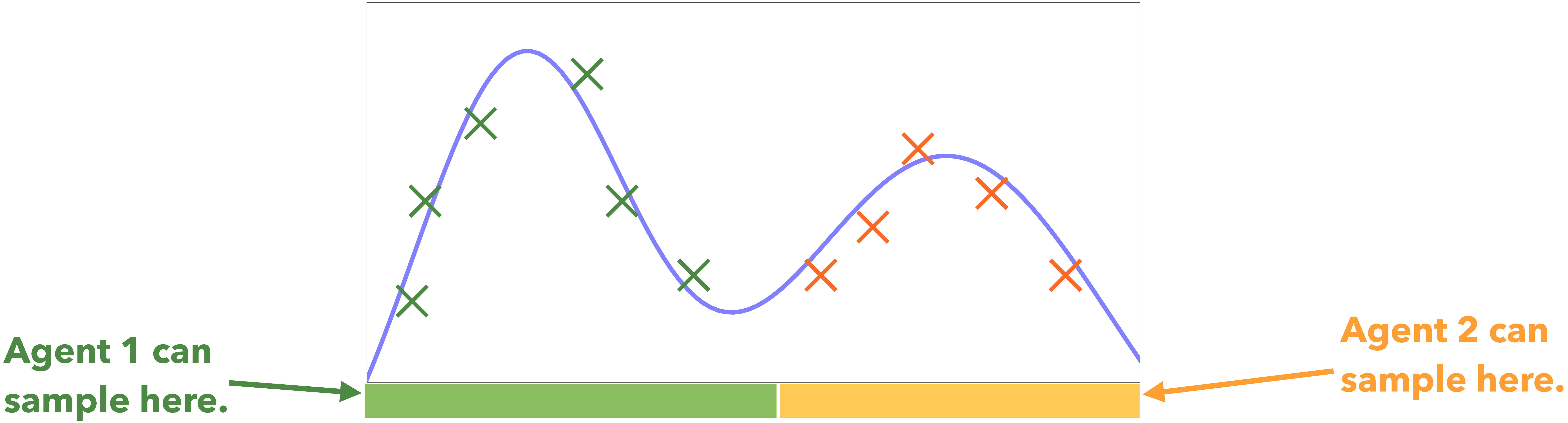
COLLABORATIVE SUPERVISED LEARNING AND EXPERIMENT DESIGN



Data sharing when there is asymmetric data collection capabilities.

E.g: hospitals in different locations, researchers with different experimental equipment etc.

+ Agents will be more willing to collaborate due to complementarity of data.



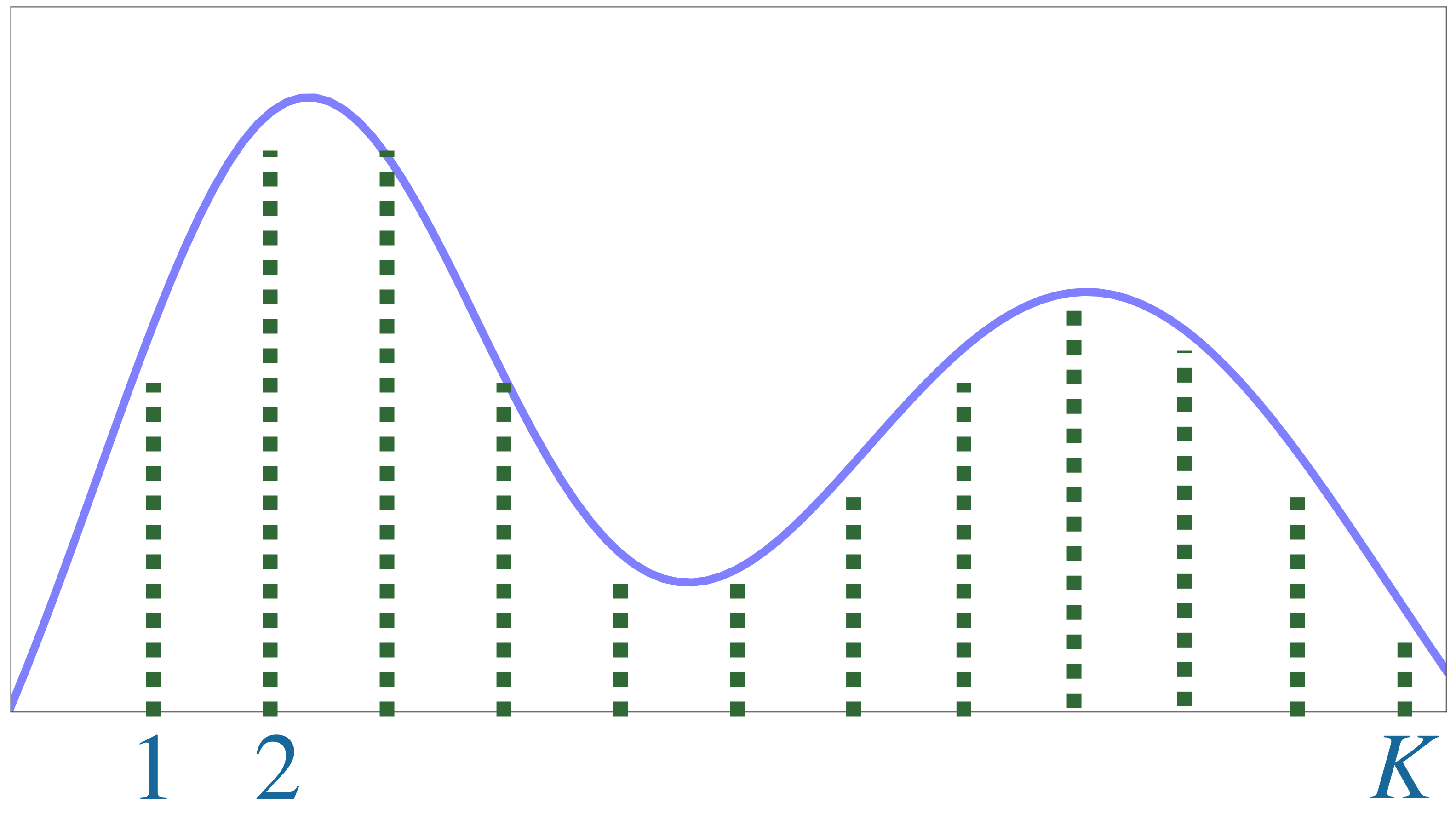
Data sharing when there is asymmetric data collection capabilities.

E.g: hospitals in different locations, researchers with different experimental equipment etc.

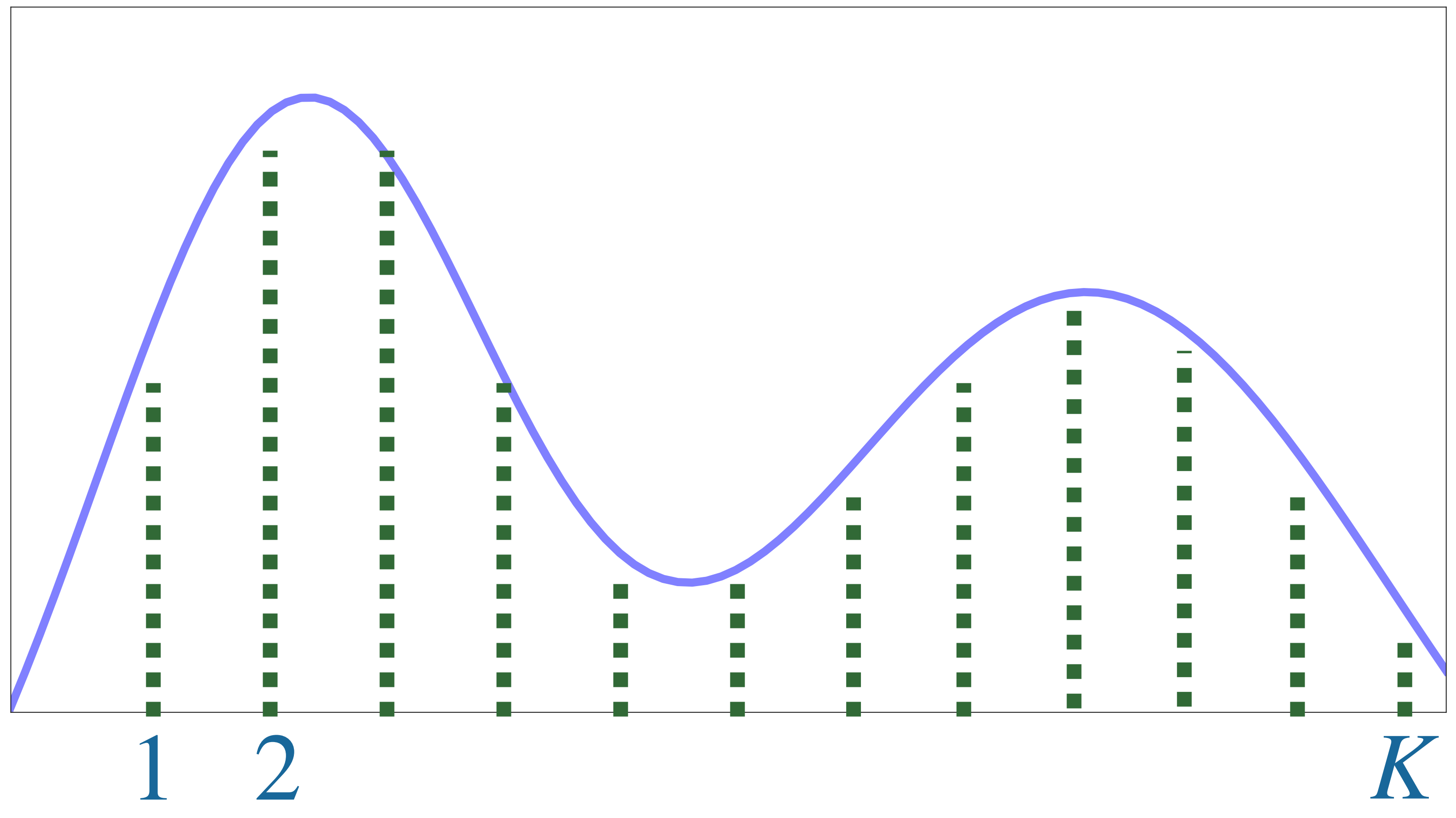
+ Agents will be more willing to collaborate due to complementarity of data.

– No way to validate an agent's data with other similar data.

Consider estimating K distributions (e.g discretizing the domain)



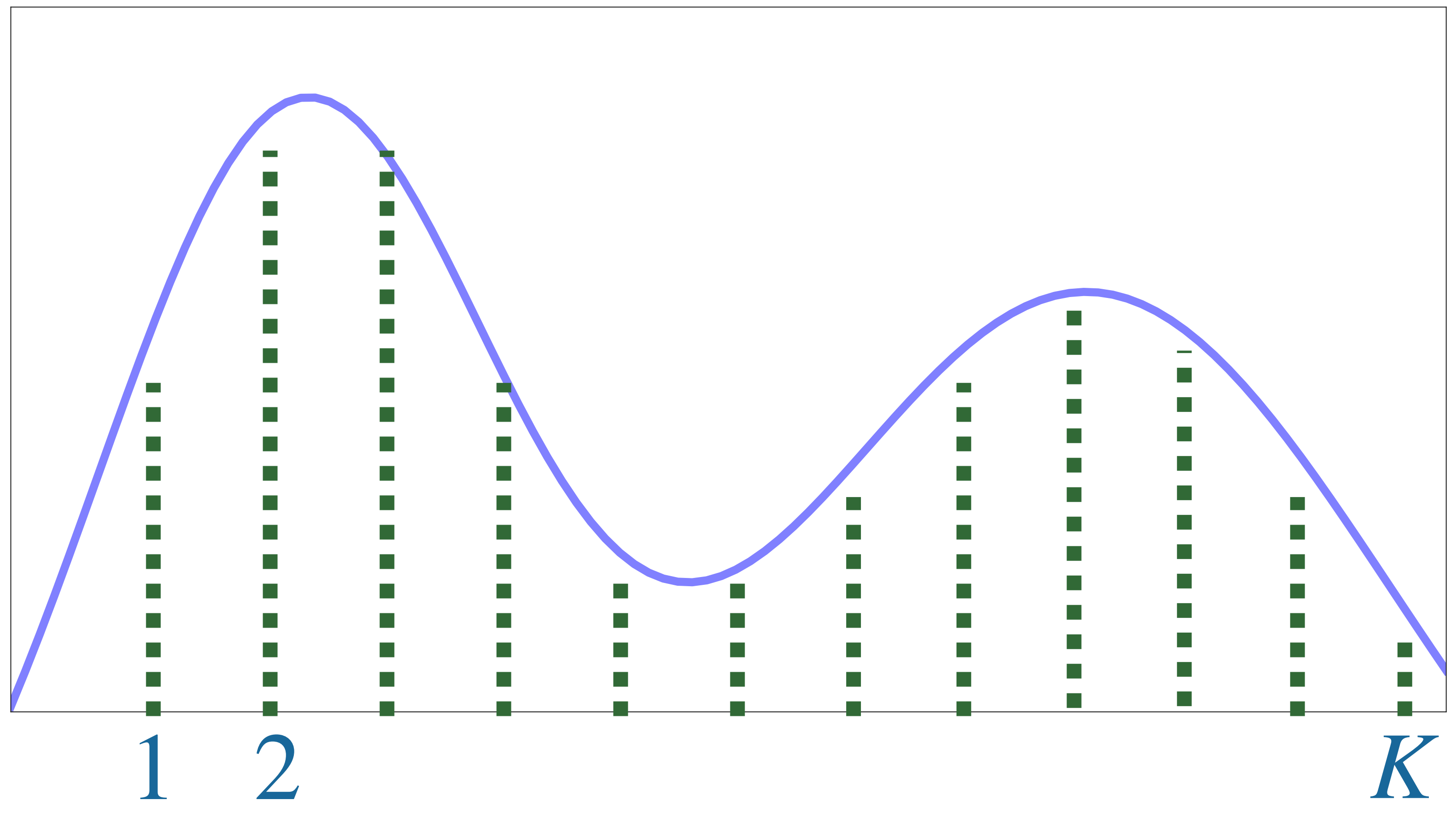
Consider estimating K distributions (e.g discretizing the domain)



Agent i can sample from distribution k at cost $c_{i,k}$.

Penalty, $p_i = \sum_{k=1}^{K} \text{est-err}_k + \sum_{k=1}^{K} c_{i,k} n_{i,k}$

Consider estimating K distributions (e.g discretizing the domain)



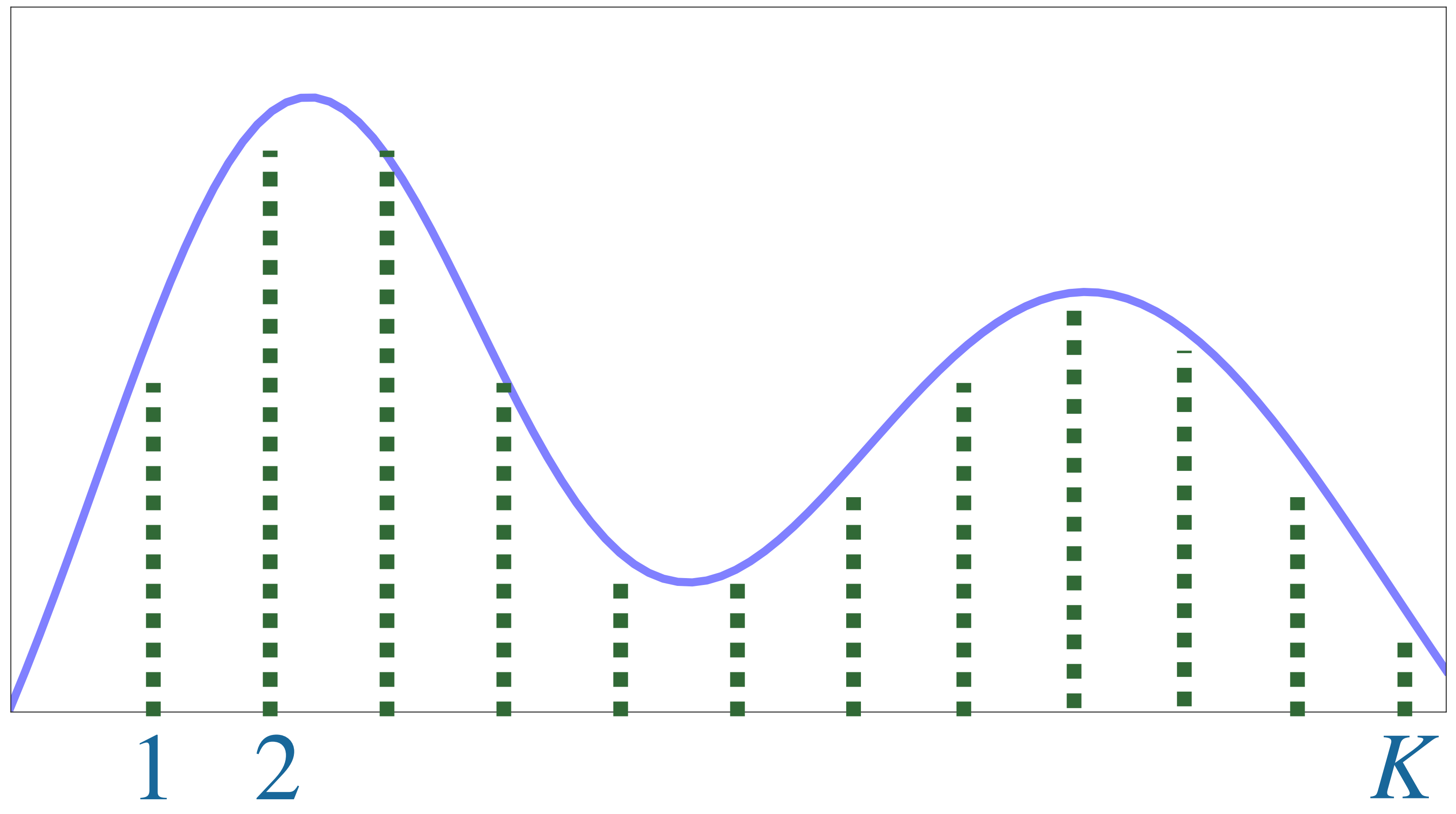
Agent i can sample from distribution k at cost $c_{i,k}$.

Penalty, $p_i = \sum_{k=1}^{K} \text{est-err}_k + \sum_{k=1}^{K} c_{i,k} n_{i,k}$

Overview of our solution:

- Uses axiomatic bargaining to define idealized *collaboration targets* assuming agents will always report truthfully.

Consider estimating K distributions (e.g discretizing the domain)



Agent i can sample from distribution k at cost $c_{i,k}$.

Penalty, $p_i = \sum_{k=1}^{K} \text{est-err}_k + \sum_{k=1}^{K} c_{i,k} n_{i,k}$

Overview of our solution:

- Uses axiomatic bargaining to define idealized *collaboration targets* assuming agents will always report truthfully.
- Enforces truthful behaviour, via corruption and other techniques.

Theorem: There exists a NIC and IR mechanism for which,

$$P(M, s^\star) \leq 8\sqrt{m} \cdot \inf_{M,s} P(M, s)$$

m: number of agents

Theorem: There exists a NIC and IR mechanism for which,

$$P(M, s^\star) \leq 8\sqrt{m} \cdot \inf_{M,s} P(M, s)$$

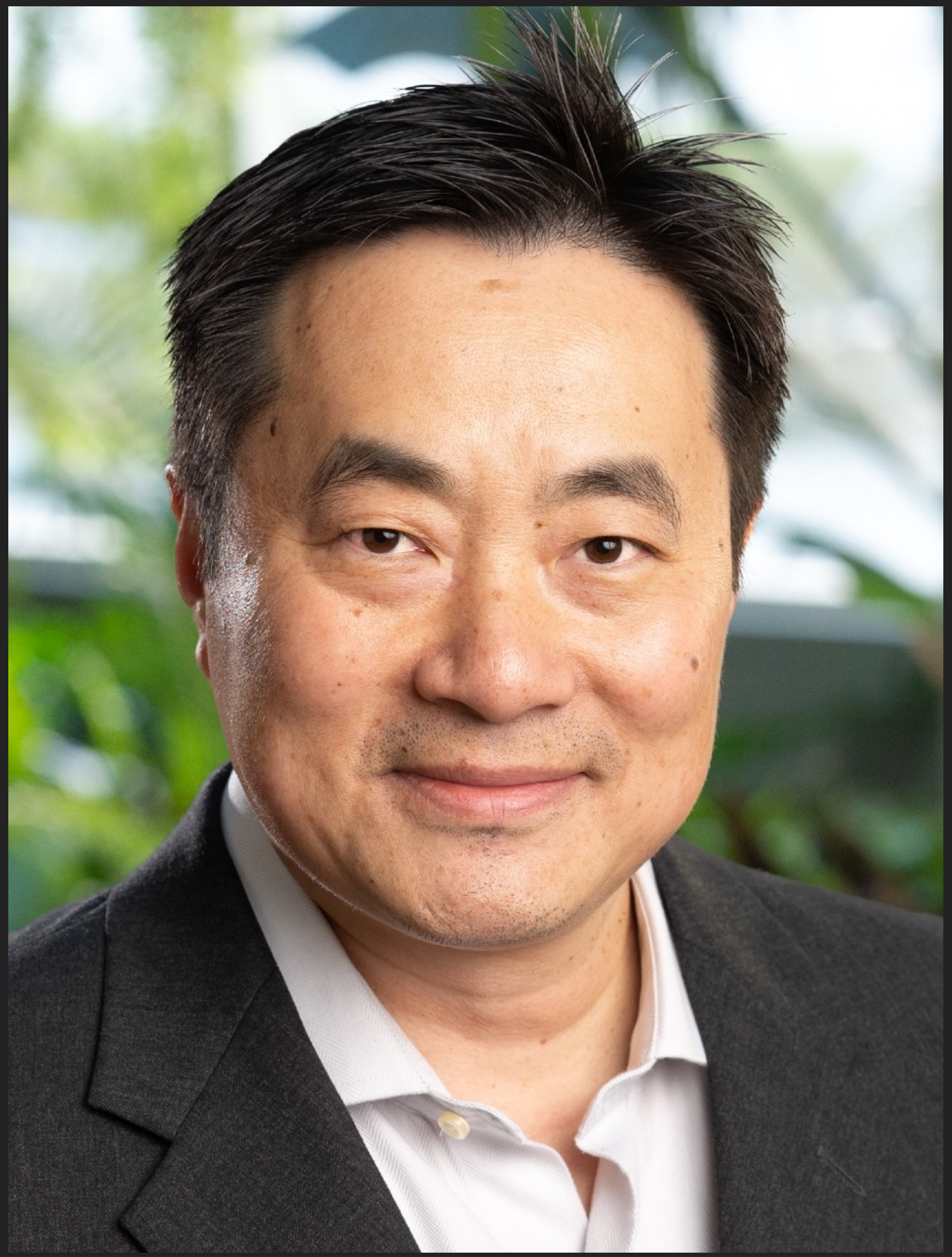
m: number of agents

Theorem (hardness): There exists a set of costs $\{c_{i,k}\}_{i,k}$ such that for any mechanism M and any Nash equilibrium $s^\star$ of this mechanism, we have

$$P(M, s^\star) \geq \mathcal{O}\left(\sqrt{m}\right) \cdot \inf_{M,s} P(M, s)$$

Yiding Chen

Alex Clinton



Jerry Zhu

THANK YOU!

kandasamy@cs.wisc.edu

SUMMARY

- Data sharing has many benefits
  - Maximize the value created by data.
  - Democratize data.
- But strategic agents can free-ride in naive mechanisms, either by not contributing data, or contributing fabricated datasets.
- For mean estimation, our mechanism is IR and NIC while achieving a factor 2 of the global minimum social penalty.

MORE MECHANISMS AND RESULTS

When the mechanism deploys an estimate for agents in a downstream application ($\mathcal{S} = \mathbb{N} \times \mathcal{F}$):

When the mechanism deploys an estimate for agents in a downstream application ($\mathcal{S} = \mathbb{N} \times \mathcal{F}$):

Theorem: For all $\epsilon > 0$, there exists a NIC and IR mechanism M_ϵ such that $P(M_\epsilon, s^\star) \leq (1 + \epsilon) \cdot \inf_{M,s} P(M, s)$.

When the mechanism deploys an estimate for agents in a downstream application ($\mathcal{S} = \mathbb{N} \times \mathcal{F}$):

Theorem: For all $\epsilon > 0$, there exists a NIC and IR mechanism M_ϵ such that $P(M_\epsilon, s^\star) \leq (1 + \epsilon) \cdot \inf_{M,s} P(M, s)$.

When agents have to report truthfully ($\mathcal{S} = \mathbb{N} \times \mathcal{H}$):

MORE MECHANISMS AND RESULTS

When the mechanism deploys an estimate for agents in a downstream application ($\mathcal{S} = \mathbb{N} \times \mathcal{F}$):

Theorem: For all $\epsilon > 0$, there exists a NIC and IR mechanism M_ϵ such that $P(M_\epsilon, s^\star) \leq (1 + \epsilon) \cdot \inf_{M,s} P(M, s)$.

When agents have to report truthfully ($\mathcal{S} = \mathbb{N} \times \mathcal{H}$):

Theorem: The *"pool and share, but only if you contribute enough data"* mechanism is NIC and IR and achieves the global minimum penalty $\inf_{M,s} P(M, s)$.