Exploring Drug Candidates: All ε-Best Arms Identification in Linear Bandits

Authors:

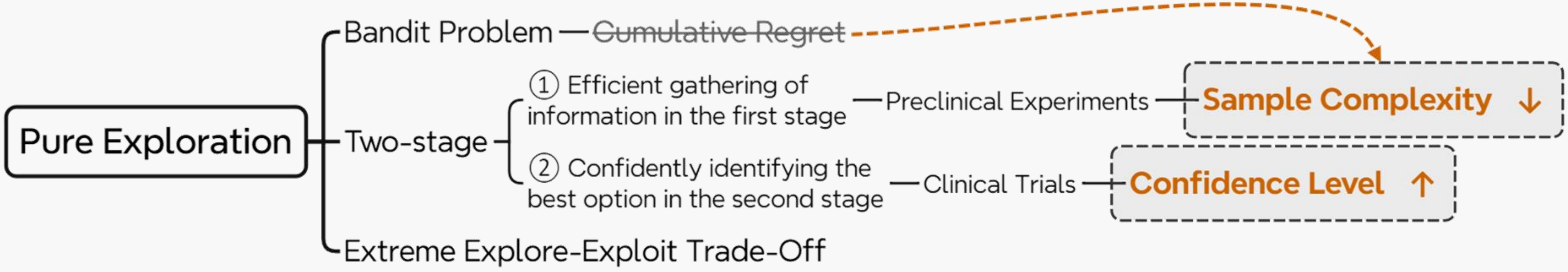
- ***Zhekai Li*** (SEIEE, Shanghai Jiao Tong University)
- **Tianyi Ma** (University of Michigan - Shanghai Jiao Tong University Joint Institute)
- **Cheng Hua** (Antai College of Economics and Management, Shanghai Jiao Tong University)
- **Ruihao Zhu** (Cornell SC Johnson College of Business)

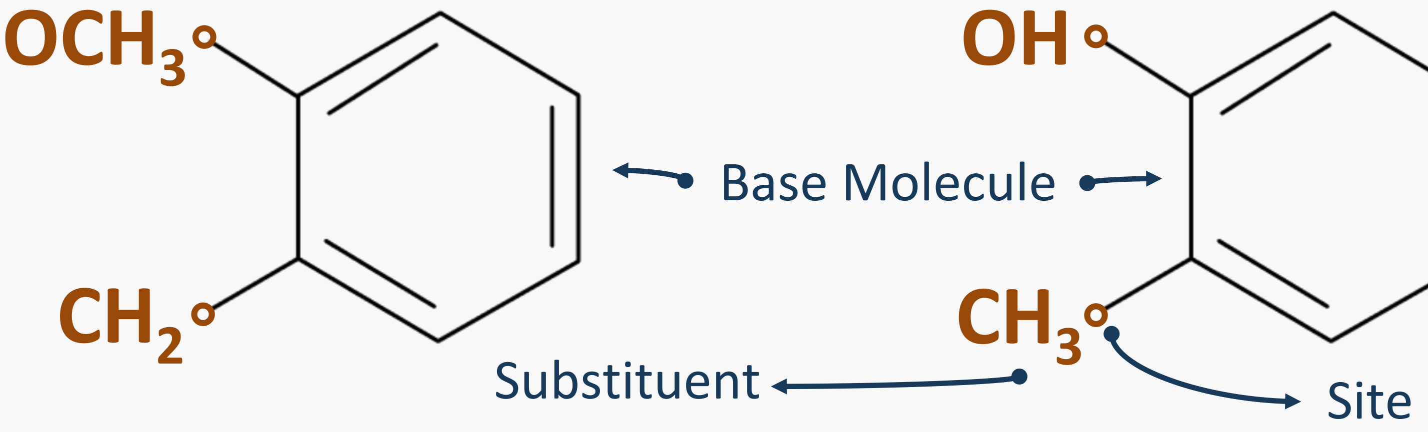
2024 INFORMS Annual Meeting

Motivation – Drug Discovery

An Illustrative Example

- **In Drug discovery:** medical researchers start with a promising molecule for treating a given disease and then **test** potentially millions of variants of this molecule to identify the highly potent candidates for later **clinical trials**





Motivation – Drug Discovery

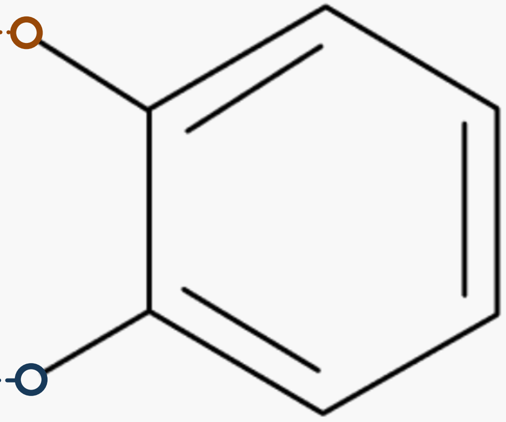
An Illustrative Example

H
OH
$OCOCH_3$
$OCOCH_2CH_3$
OCH_3
NO_2

Six Optional substituents for the first site

CH_3
$CH_2CH_2(C_6H_5)$
$CH_2(C_3H_5)$
$CH_2CHC(CH_3)_2$
CH_2
NO_2

Six Optional substituents for the second site



6×6=36

Combination

($H+CH_3$), ($H+CH_2CH_2(C_6H_5)$), ($H+CH_2(C_3H_5)$), ($H+CH_2$), ($H+NO_2$)

($OH+CH_3$), ($OH+CH_2CH_2(C_6H_5)$), ($OH+CH_2(C_3H_5)$), ($OH+CH_2$), ($OH+NO_2$)

($OCOCH_3+CH_3$), ($OCOCH_3$ $+CH_2CH_2(C_6H_5)$), ($OCOCH_3$ $+CH_2(C_3H_5)$), ($OCOCH_3$ $+CH_2$), ($OCOCH_3$ $+NO_2$)
……

How to handle this increasing scale of variants

Motivation – Why Linear Bandits

Free-Wilson – A Quantitative Structured Relationship (Free and Wilson 1964)

$$\text{Total Effectiveness} = [\text{Effectiveness of Base Molecule}] + \sum_i [\text{Effectiveness of Substituent } i]$$

Combination - grow exponentially

($H+CH_3$), ($H+CH_2CH_2(C_6H_5)$), ($H+CH_2(C_3H_5)$), ($H+CH_2$), ($H+NO_2$)

($OH+CH_3$), ($OH+CH_2CH_2(C_6H_5)$), ($OH+CH_2(C_3H_5)$), ($OH+CH_2$), ($OH+NO_2$)

($OCOCH_3+CH_3$), ($OCOCH_3+CH_2CH_2(C_6H_5)$), ($OCOCH_3+CH_2(C_3H_5)$), ($OCOCH_3+CH_2$), ($OCOCH_3+NO_2$)

……

H
OH
$OCOCH_3$
$OCOCH_2CH_3$
OCH_3
NO_2

CH_3
$CH_2CH_2(C_6H_5)$
$CH_2(C_3H_5)$
$CH_2CHC(CH_3)_2$
CH_2
NO_2

Effectiveness of all the substituents → **Much fewer**

Shared Information
↓
Linear Bandits

Motivation – Why "All ε-Best"

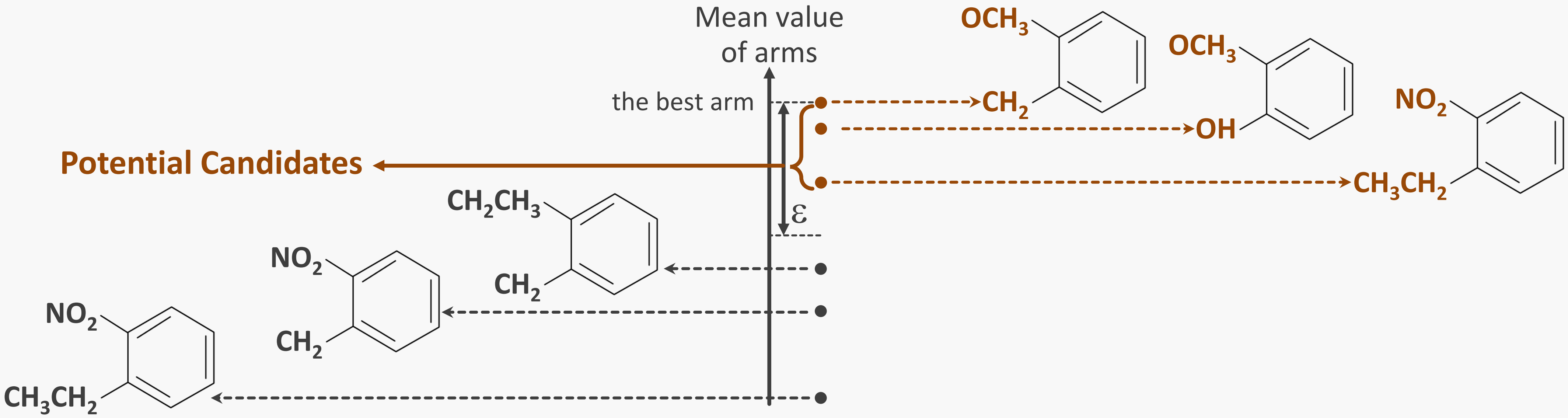
Problem

- Expansion of problem scale → Introducing the linear structure
- Drug Development: preclinical drug discovery → clinical trials

high cost and low efficacy

BAI → Finding All ε-Best Candidates

- Identifying all candidates whose effectiveness is within a range of ε from the best one



Motivation – Why "All ε-Best"

Zhekai Li (SJTU)

Problem

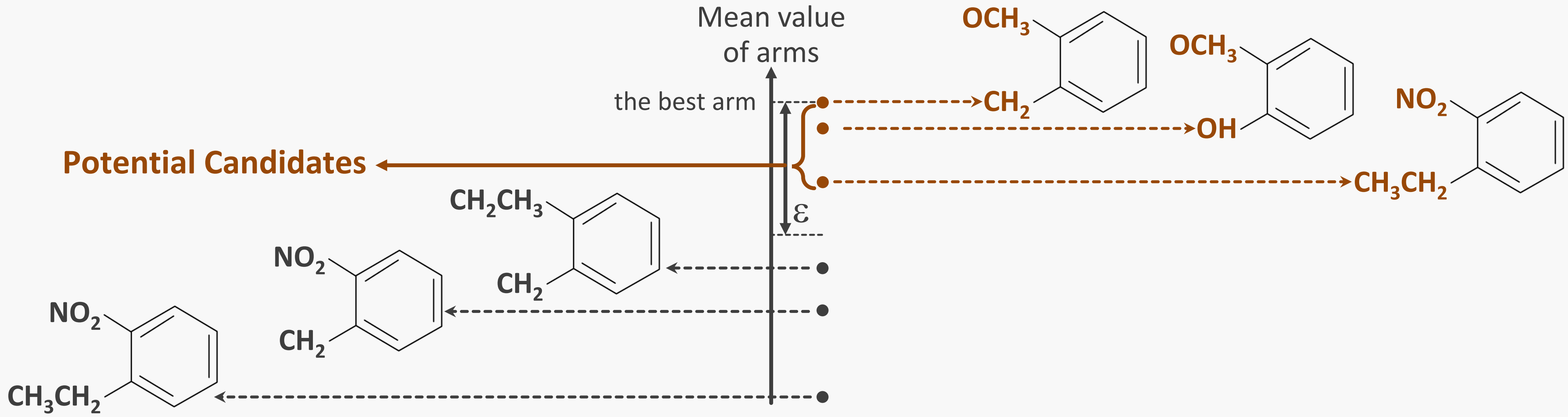
- Expansion of problem scale → Introducing the linear structure
- Drug Development: preclinical drug discovery →the most effective candidates → clinical trials

the likelihood of finding at least one successful, marketable drug

high cost and low efficacy

BAI → Finding All ε-Best Candidates

- Identifying all candidates whose effectiveness is within a range of ε from the best one



Model - All ε-Best Arms Identification in Linear Bandits

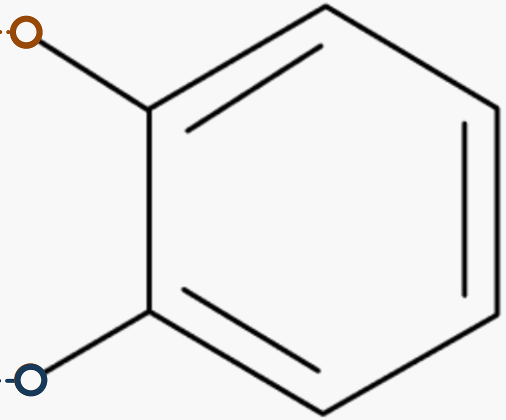
- **A finite set of K arms**, denoted as $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$ and we have $\mu_1 > \mu_2 \geq \ldots \geq \mu_K$
- **The Linear Structure**
  - Denote $\boldsymbol{\theta}$ as the unknown parameter vector
  - Denote $\boldsymbol{A} = \{a_1, a_2, \ldots, a_K\} \subset \mathbf{R}^d$ as the set containing the feature vectors of all arms
  - Observe the reward $X_t = \boldsymbol{a}_{A_t}{}^\top \boldsymbol{\theta} + \eta_t$, where η_t, the noise, is conditionally 1-sub-Gaussian

H
OH
$OCOCH_3$
$OCOCH_2CH_3$
OCH_3
NO_2

Optional substituents for the first site

CH_3
$CH_2CH_2(C_6H_5)$
$CH_2(C_3H_5)$
$CH_2CHC(CH_3)_2$
CH_2
NO_2

Optional substituents for the second site

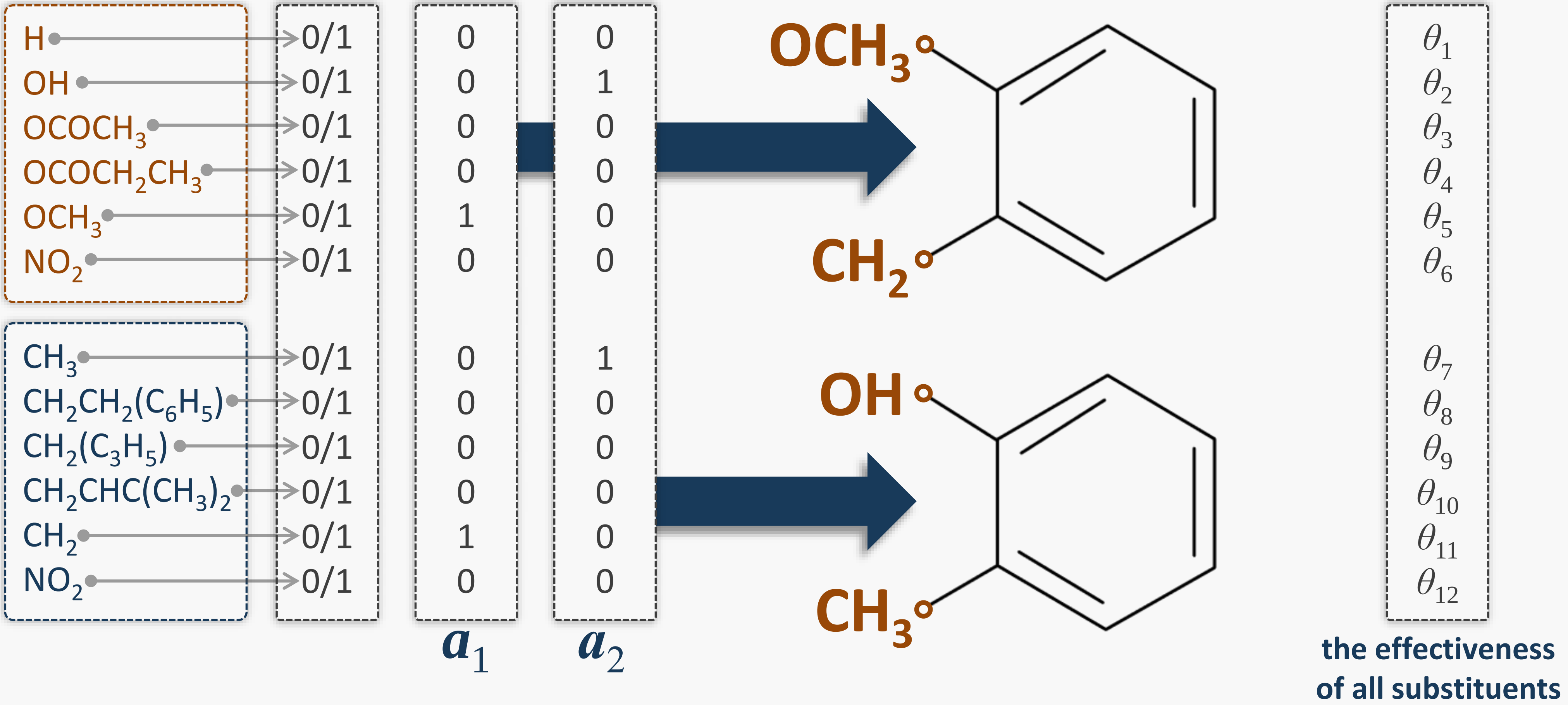


Previous Drug Example

Motivation
Model
Algorithm
Results
Experiment

Model - All ε-Best Arms Identification in Linear Bandits

- **A finite set of K arms**, denoted as $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$ and we have $\mu_1 > \mu_2 \geq \ldots \geq \mu_K$
- **The Linear Structure**
  - Denote $\boldsymbol{\theta}$ as the unknown parameter vector
  - Denote $\boldsymbol{A} = \{a_1, a_2, \ldots, a_K\} \subset \mathbf{R}^d$ as the set containing the feature vectors of all arms
  - Observe the reward $X_t = \boldsymbol{a}_{At}^{\top} \boldsymbol{\theta} + \eta_t$, where η_t, the noise, is conditionally 1-sub-Gaussian



Model - All ε-Best Arms Identification in Linear Bandits

- **A finite set of K arms**, denoted as $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$ and we have $\mu_1 > \mu_2 \geq \ldots \geq \mu_K$
- **The Linear Structure**
  - Denote $\boldsymbol{\theta}$ as the unknown parameter vector
  - Denote $\boldsymbol{A} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, a_K\} \subset \mathbf{R}^d$ as the set containing the feature vectors of all arms
  - Observe the reward $X_t = \boldsymbol{a}_{A_t}{}^\top \boldsymbol{\theta} + \eta_t$, where η_t, the noise, is conditionally 1-sub-Gaussian
- **Task: Denote the set of all ε-best arms with mean vector $\boldsymbol{\mu}$ as $G_\varepsilon(\boldsymbol{\mu}) := \{\boldsymbol{i}: \boldsymbol{\mu_i} \geq \boldsymbol{\mu_1} - \varepsilon\}$**
  - Additive ε-Best Arm: given $\varepsilon > 0$, an arm i is deemed ε-best if $\mu_i \geq \mu_1 - \varepsilon$
- **Performance Metric: the sample complexity**
  - Confidence level δ is fixed → Fixed-Confidence Setting

$$
\begin{aligned}
\min \quad & \mathbf{E}_{\boldsymbol{\mu}}[\tau_\delta] \\
\text{s.t.} \quad & \mathbf{P}_{\boldsymbol{\mu}}(\tau_\delta < \infty, \textit{recommended set equals } G_\varepsilon(\mu)) \geq 1 - \delta
\end{aligned}
$$

Related Literature

- Pure Exploration with different tasks. **Mannor and Tsitsiklis (2004), Even-Dar et al. (2006), Russo (2020), Komiyama et al. (2023), Kalyanakrishnan and Stone (2010), Kalyanakrishnan et al. (2012), Locatelli et al. (2016), Abernethy et al. (2016), Garivier and Kaufmann (2016),**
- Linear Bandits in Pure Exploration. **Abbasi-Yadkori et al. (2011), Gabillon et al. (2012), Hoffman et al. (2014), Soare et al. (2014), Fiez et al. (2019), Ŕeda et al. (2021), Yang and Tan (2021), Azizi et al. (2023)**
- Model Misspecification. **Ghosh et al. (2017), Lattimore et al. (2020), Reda et al. (2021), Ahn et al. (2024)**
- All ε-Best Arms Identification. **Mason et al. (2020), Al Marjani et al. (2022)**

Current Status of Research

- All ε-best arms identification in stochastic bandits **Mason et al. (2020) – lower bound and two algorithms**
- **Limited to the stochastic setting and is hard to be applied in problems with a large number of choices**

Research Question

- **How to solve the All ε-Best Arms Identification in Linear Bandits?**
  - The description of the problem complexity
  - The algorithm and the upper bound
  - Extensions and other insights (Misspecification and GLM)

LinFACTE: Linear Fast Arm Classification with Threshold Estimation

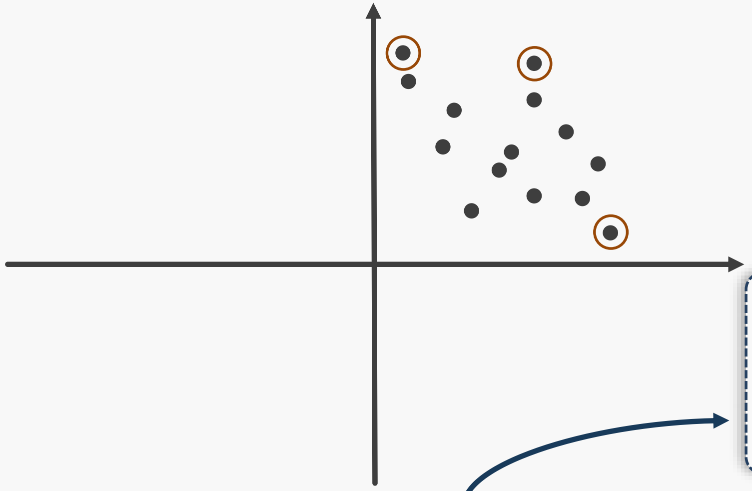
General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design** { **G-optimal (arm's confidence region)** / **$\mathcal{XY}$-optimal (gap's confidence region)**

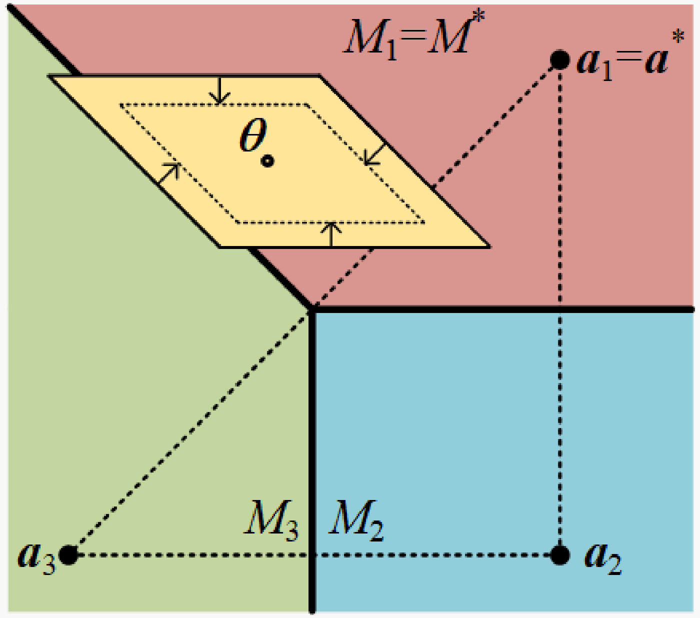


2-Dimensional Arm Space

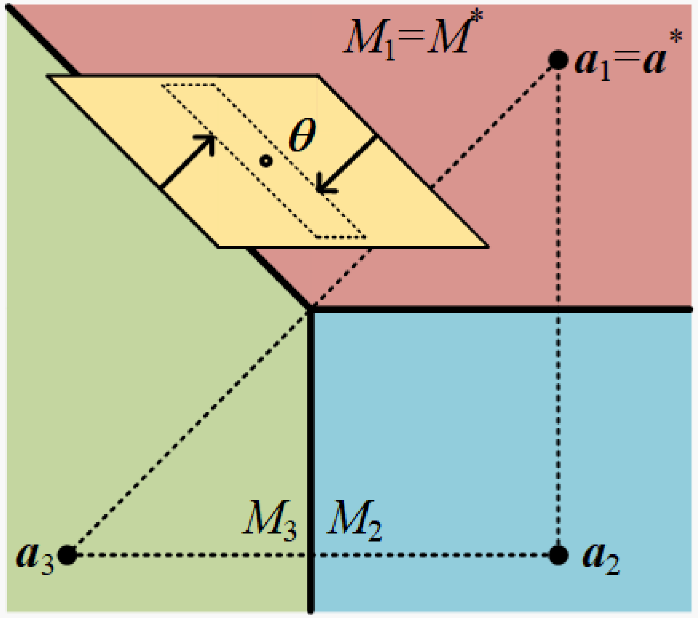
- **Given confidence level → Minimized budget**
- **Only a small number of arms need to be sampled**
- **Dramatically reducing the problem's difficulty**

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design** { **G-optimal (arm's confidence region)** / **$\mathcal{XY}$-optimal (gap's confidence region)**



Uniform Contraction Based on G-Optimal Design



More Purposeful Contraction of XY-Optimal Design

LinFACTE: Linear Fast Arm Classification with Threshold Estimation

General Structure

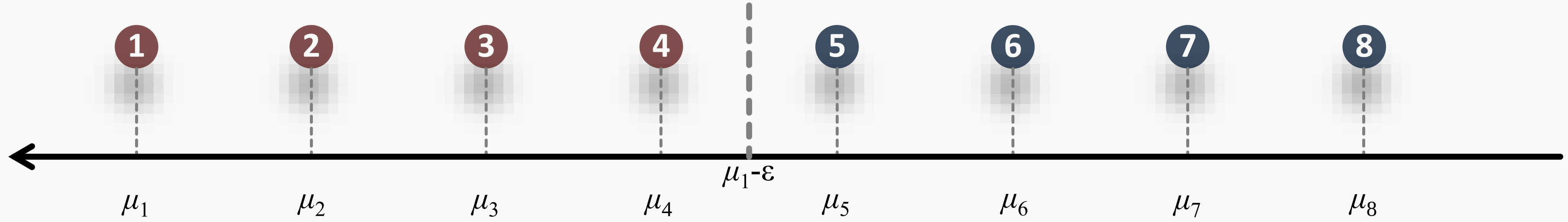
- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms



LinFACTE: Linear Fast Arm Classification with Threshold Estimation

General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

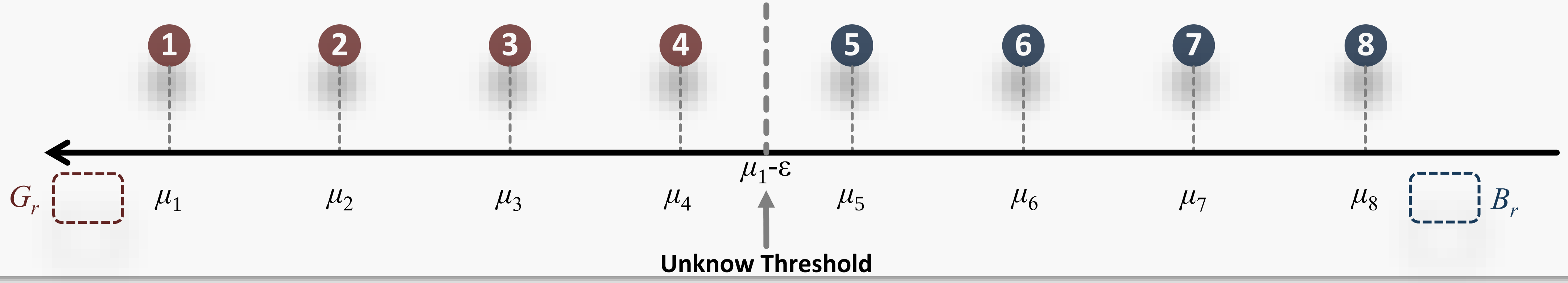
Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms

- Update two sets of arms, that is, G_r and B_r, classifying arms that are empirically ε-best and not ε-best



LinFACTE: Linear Fast Arm Classification with Threshold Estimation

General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

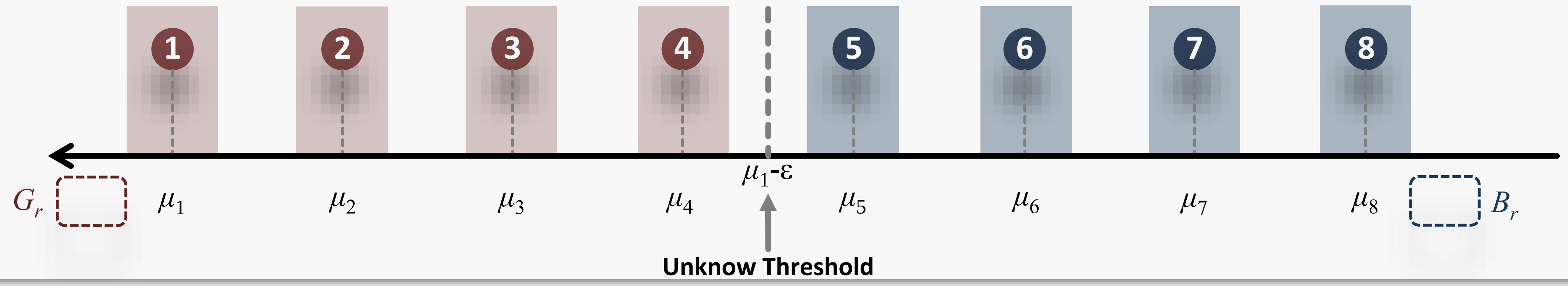
Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms

- Update two sets of arms, that is, G_r and B_r, classifying arms that are empirically ε-best and not ε-best
- Maintain confidence radius of each arm and confidence bound of the unknown threshold $\mu_1 - \varepsilon$



LinFACTE: Linear Fast Arm Classification with Threshold Estimation

General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

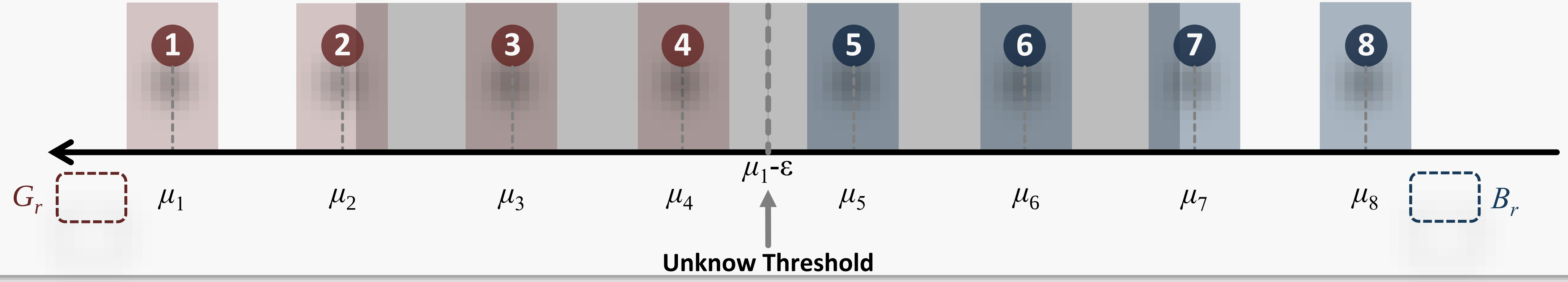
Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms

- Update two sets of arms, that is, G_r and B_r, classifying arms that are empirically ε-best and not ε-best
- Maintain confidence radius of each arm and confidence bound of the unknown threshold $\mu_1 - \varepsilon$



LinFACTE: Linear Fast Arm Classification with Threshold Estimation

General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

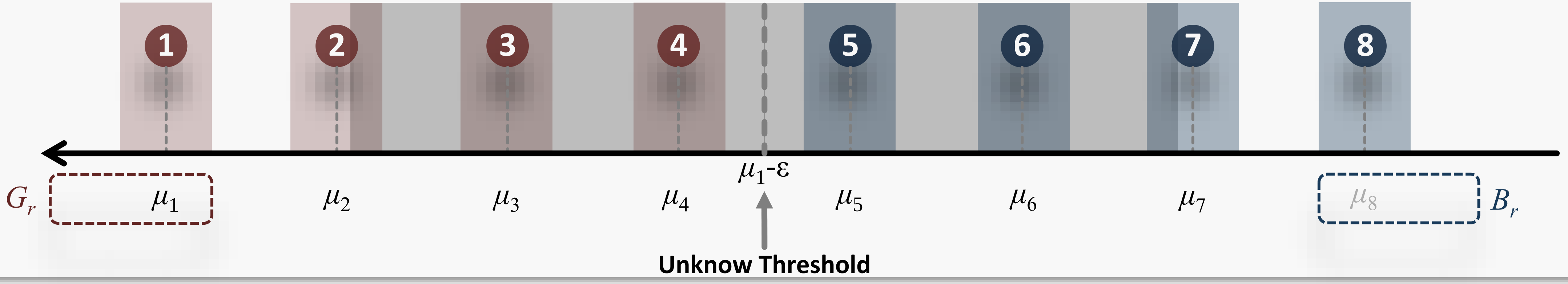
Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms

- Update two sets of arms, that is, G_r and B_r, classifying arms that are empirically ε-best and not ε-best
- Maintain confidence radius of each arm and confidence bound of the unknown threshold $\mu_1 - \varepsilon$



LinFACTE: Linear Fast Arm Classification with Threshold Estimation

General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

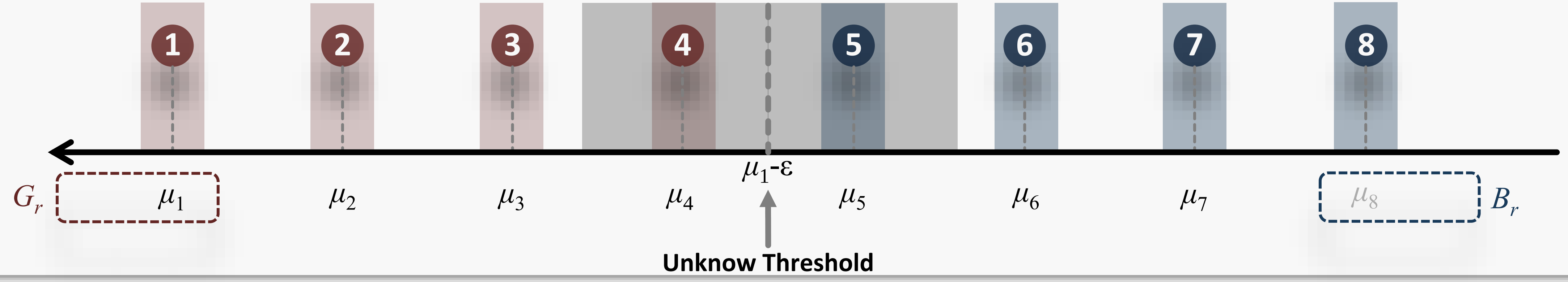
Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms

- Update two sets of arms, that is, G_r and B_r, classifying arms that are empirically ε-best and not ε-best
- Maintain confidence radius of each arm and confidence bound of the unknown threshold $\mu_1 - \varepsilon$



LinFACTE: Linear Fast Arm Classification with Threshold Estimation

General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

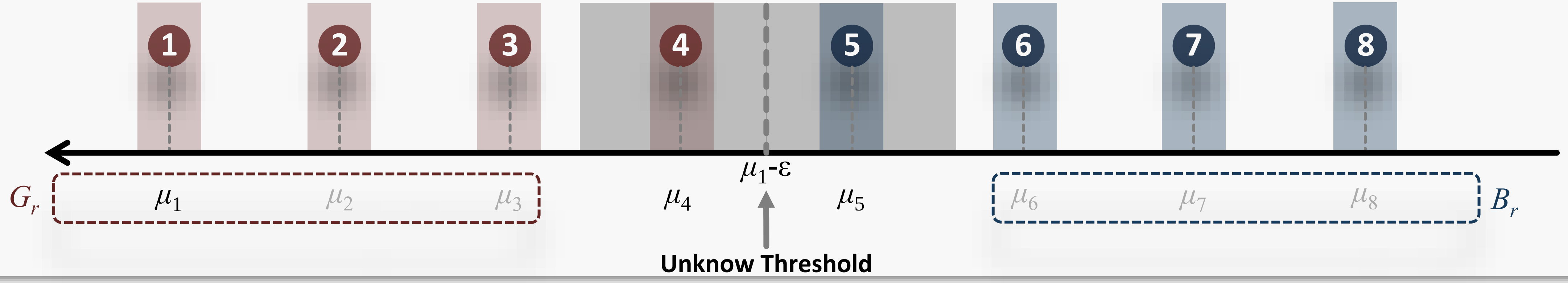
Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms

- Update two sets of arms, that is, G_r and B_r, classifying arms that are empirically ε-best and not ε-best
- Maintain confidence radius of each arm and confidence bound of the unknown threshold $\mu_1 - \varepsilon$



LinFACTE: Linear Fast Arm Classification with Threshold Estimation

General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

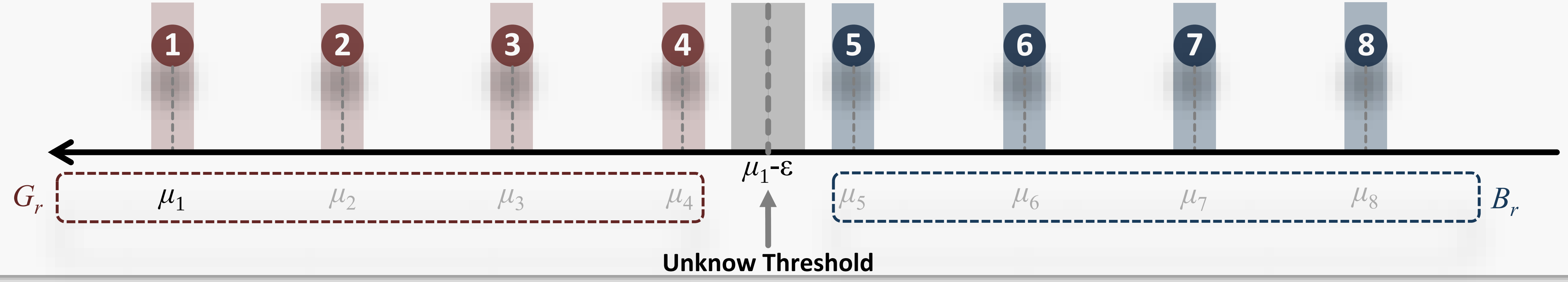
Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms

- Update two sets of arms, that is, G_r and B_r, classifying arms that are empirically ε-best and not ε-best
- Maintain confidence radius of each arm and confidence bound of the unknown threshold $\mu_1 - \varepsilon$



LinFACTE: Linear Fast Arm Classification with Threshold Estimation

Zhekai Li (SJTU)

General Structure

- LinFACTE is a **phase-based elimination and classification algorithm with five general components**
- Initialization → sampling → estimation → classification → stopping and decision

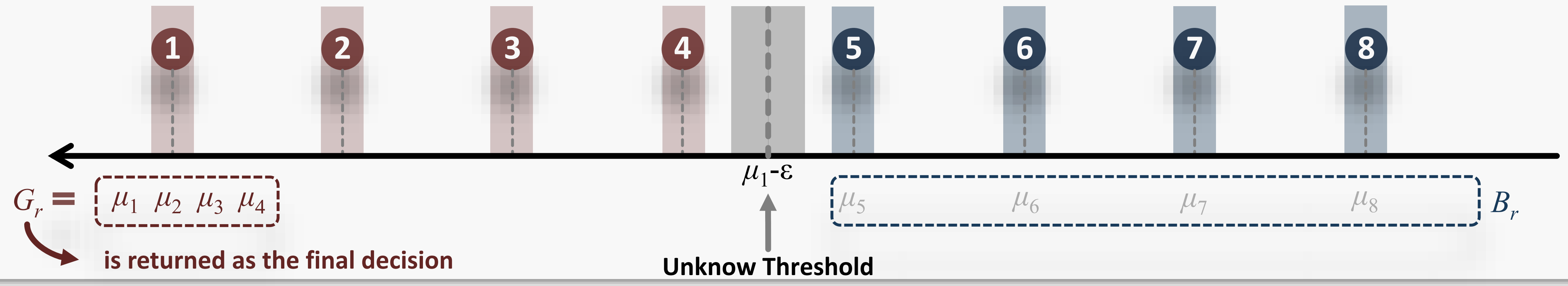
Phase Iteration

Sampling and Estimation

- A probabilistic guarantee that the true mean value is within a range of the estimated mean value for each arm
- Challenge: Extremely large arm space → solved by **the optimal design - G-optimal (arm's confidence region)**

Toy Example of Eight Arms

- Update two sets of arms, that is, G_r and B_r, classifying arms that are empirically ε-best and not ε-best
- Maintain confidence radius of each arm and confidence bound of the unknown threshold $\mu_1 - \varepsilon$



Theoretical Guarantees

Lower Bound

- Consider a set of arms where arm i follows a normal distribution. Any δ-PAC algorithm must satisfy

$$\frac{\mathbb{E}_{\boldsymbol{\mu}}\left[\tau_\delta\right]}{\log\left(1/2.4\delta\right)} \geq (\Gamma^*)^{-1} = \min_{\boldsymbol{p}\in S_K} \max_{(i,j,m)\in\mathcal{X}} \max\left\{\frac{2\|\boldsymbol{a}_i-\boldsymbol{a}_j\|^2_{\boldsymbol{V}_{\boldsymbol{p}}^{-1}}}{\left(\boldsymbol{a}_i^\top\boldsymbol{\theta}-\boldsymbol{a}_j^\top\boldsymbol{\theta}+\varepsilon\right)^2}, \frac{2\|\boldsymbol{a}_1-\boldsymbol{a}_m\|^2_{\boldsymbol{V}_{\boldsymbol{p}}^{-1}}}{\left(\boldsymbol{a}_1^\top\boldsymbol{\theta}-\boldsymbol{a}_m^\top\boldsymbol{\theta}-\varepsilon\right)^2}\right\}$$

Upper Bound

Define $\Delta = \min(\alpha_\varepsilon, \beta_\varepsilon)/8$. Based on the **G-optimal** design, with a probability of at least $1-\delta$, the expected sampling budget of LinFACTE has the following upper bound

$$\mathbb{E}\left[T_G \mid \mathcal{E}\right] = O\left(d\Delta^{-2}\log\left(\frac{K}{\delta}\log_2\left(\Delta^{-2}\right)\right) + d^2\log\left(\Delta^{-1}\right)\right)$$

Define $\Delta = \min(\alpha_\varepsilon, \beta_\varepsilon)/8$. Based on the **$\mathcal{XY}$-optimal** design, with a probability of at least $1-\delta$, the expected sampling budget of LinFACTE has the following upper bound

$$\mathbb{E}\left[T_{\mathcal{XY}} \mid \mathcal{E}\right] = O\left(d\log\left(\Delta^{-1}\right)\log\left(\frac{K}{\delta}\log_2\left(\Delta^{-2}\right)\right)(\Gamma^*)^{-1} + r(\epsilon)\log\left(\Delta^{-1}\right)\right)$$

- Algorithm with $\mathcal{XY}$-optimal design → **near optimal up to some logarithmic factors**
- Model Extension → more general results → applicability of LinFACTE
  - Extendable to **misspecified linear bandits**
  - Extendable to **generalized linear model (GLM)**

Motivation

Model

Algorithm

Results

Experiment

Experiment – Setting

Baselines

- **Bayesian optimization** with a knowledge gradient acquisition function **(Negoescu et al. 2011)**
- **BayesGap**: a gap-based algorithm for the best arm identification (BAI) **(Hoffman et al. 2014)**
- **m-LinGapE** and **LinGIFA**: two gap-based algorithms for the top m identification **(Ŕeda et al. 2021)**
- **Lazy Track-Threshold-and-Stop**: track and stop algorithm for the threshold bandit **(Tewari et al. 2024)**

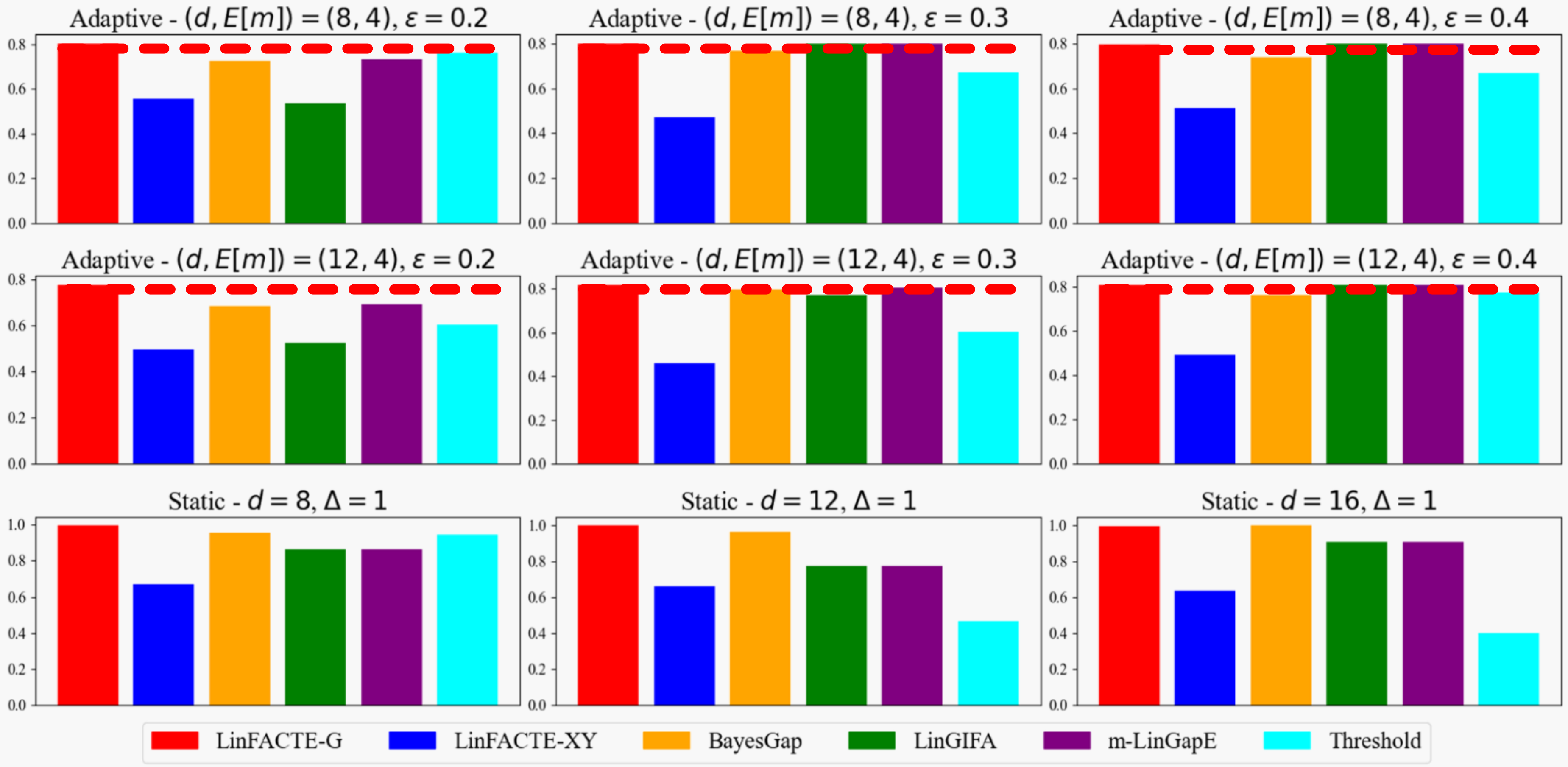
Dataset

- Synthetic Data → LinFACTE's superiority in various edge cases
  - Adaptive Setting
  - Static Setting
- Real Data From Drug Discovery → LinFACTE's applicability in real-world applications

Experiment – Simulation Results

Motivation
Model
Algorithm
Results
Experiment

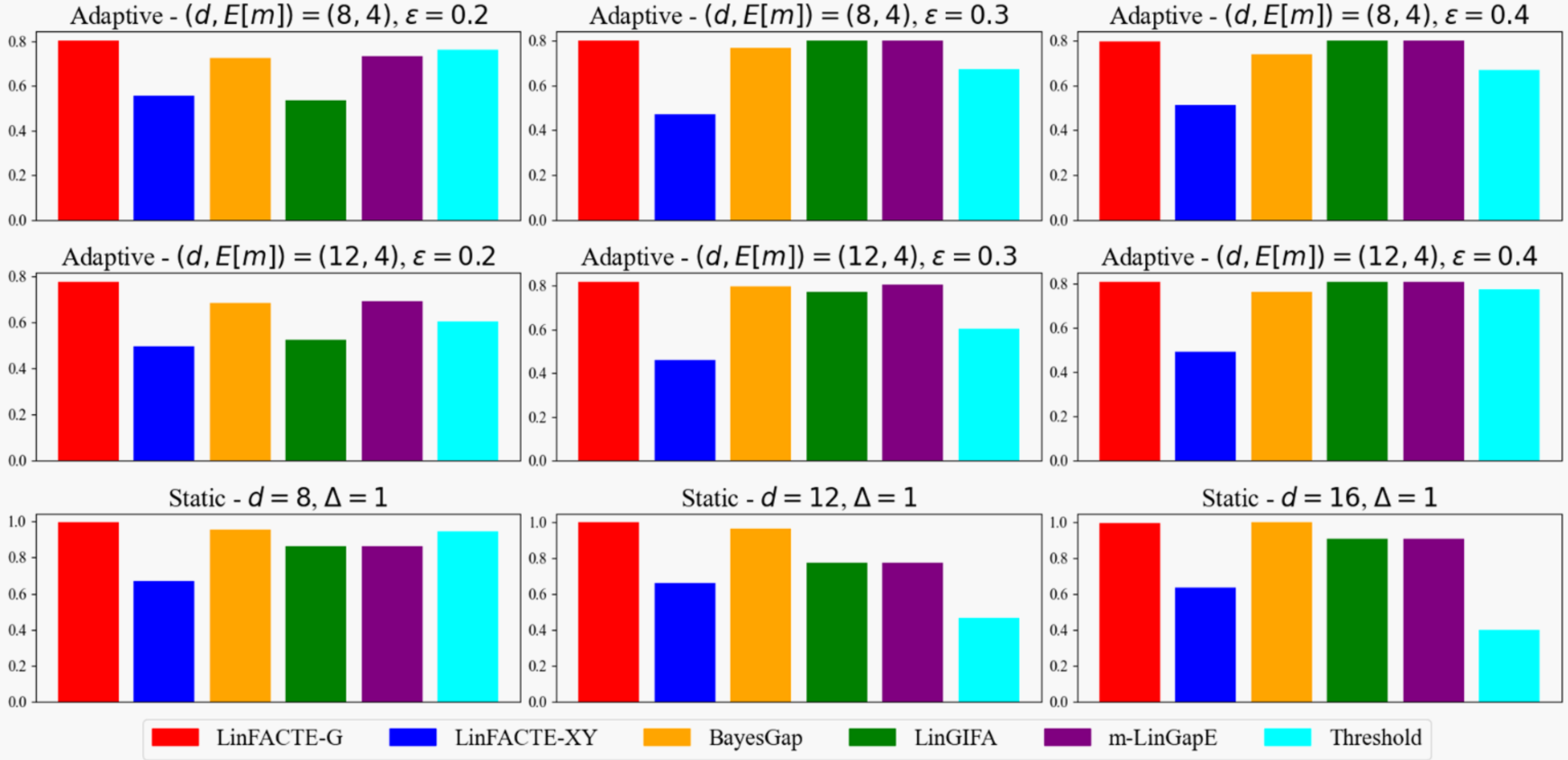
Synthetic Data – F1 Score



Motivation
Model
Algorithm
Results
Experiment

Experiment – Simulation Results

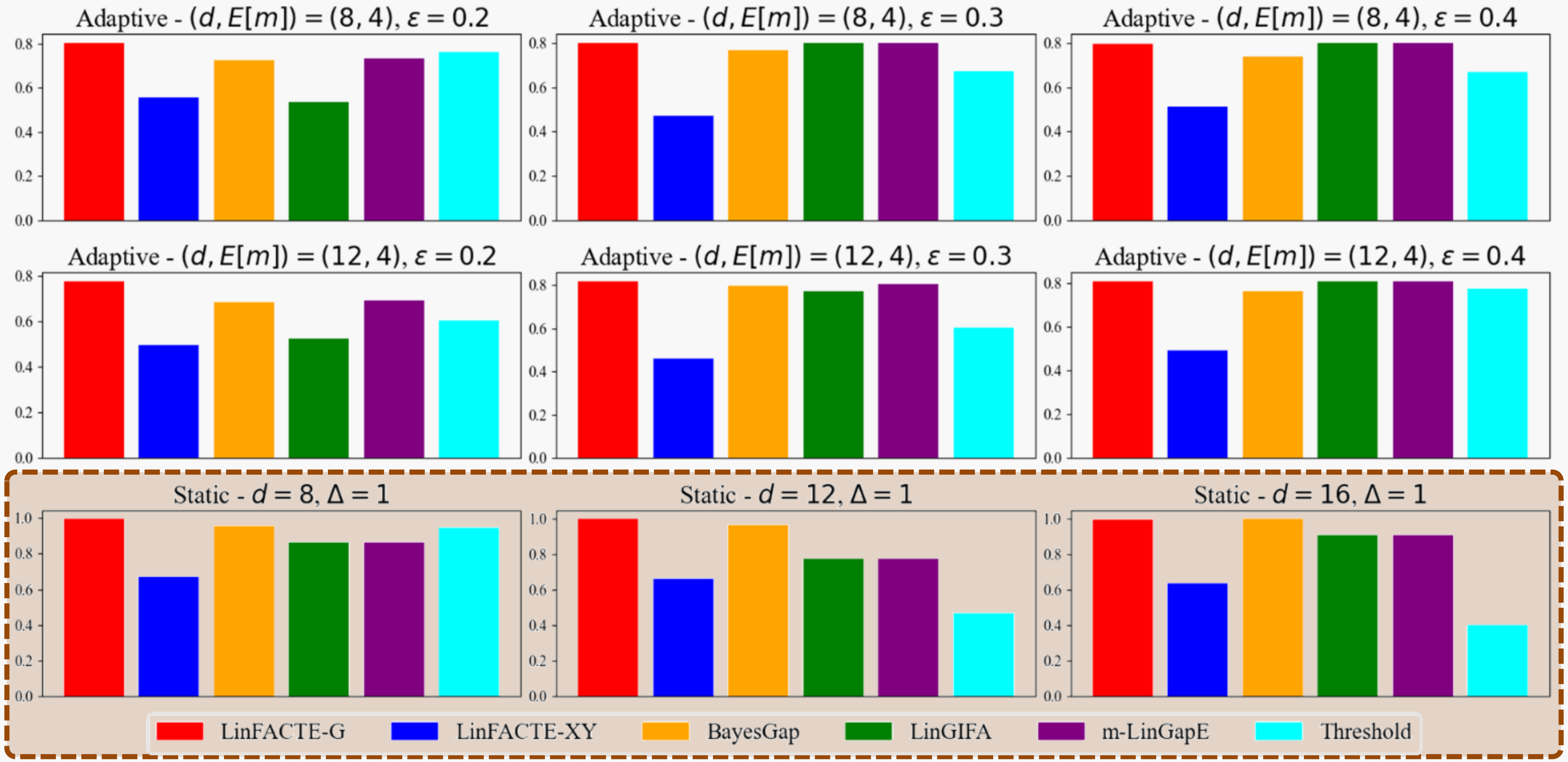
Synthetic Data – F1 Score



Superiority and Adaptivity in Complex Edge Cases

Experiment – Simulation Results

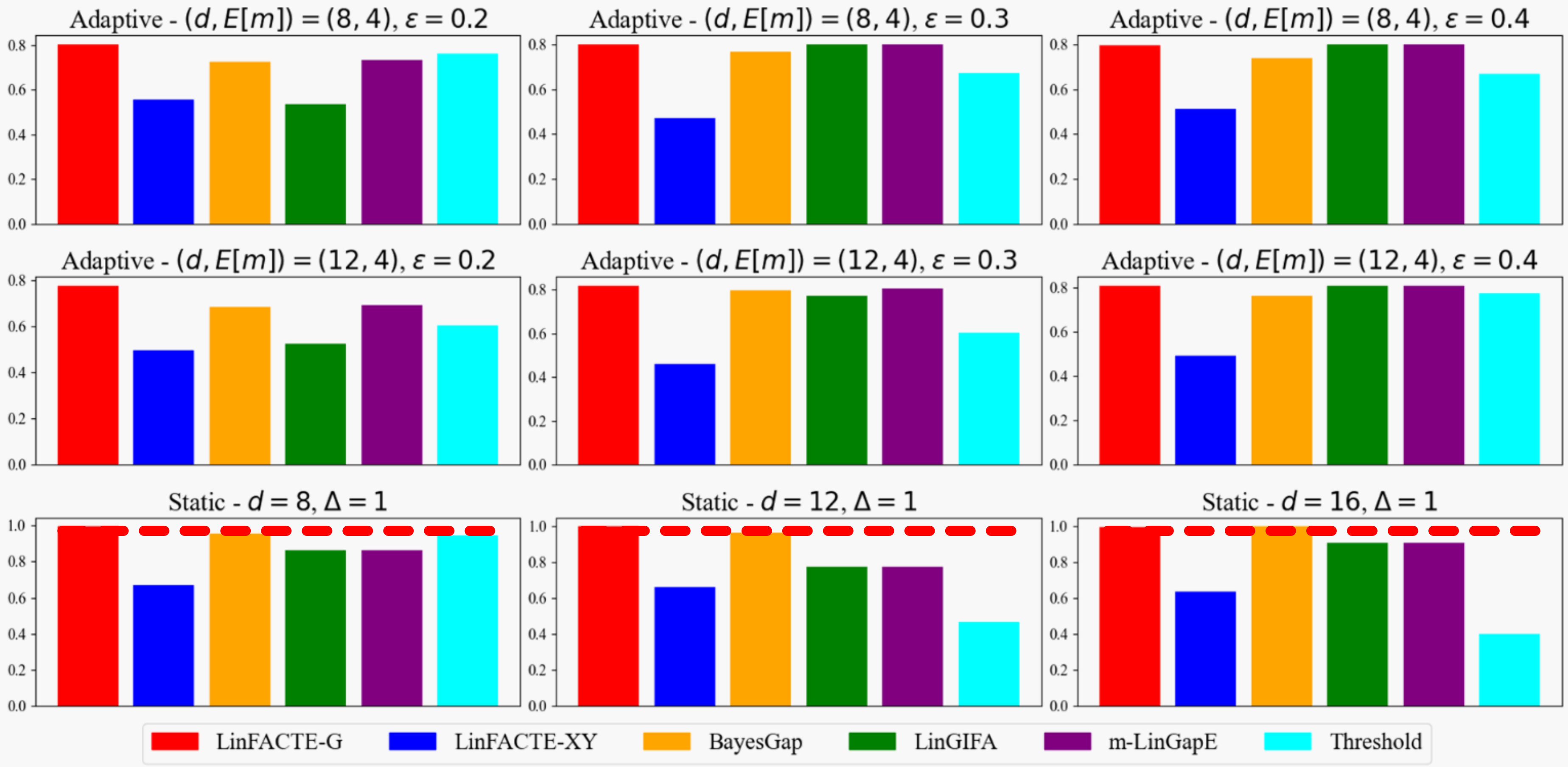
Synthetic Data – F1 Score



Experiment – Simulation Results

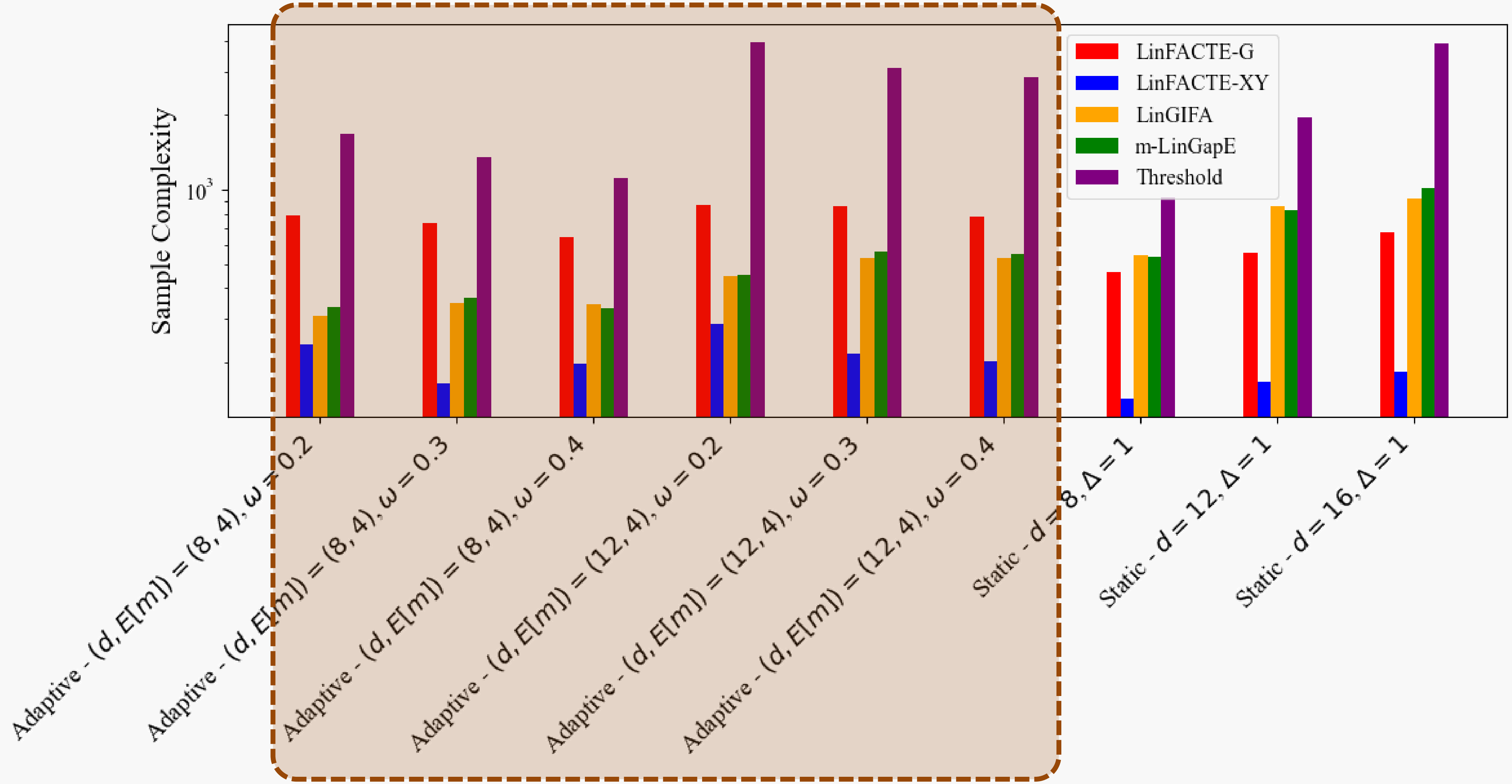
Motivation
Model
Algorithm
Results
Experiment

Synthetic Data – F1 Score



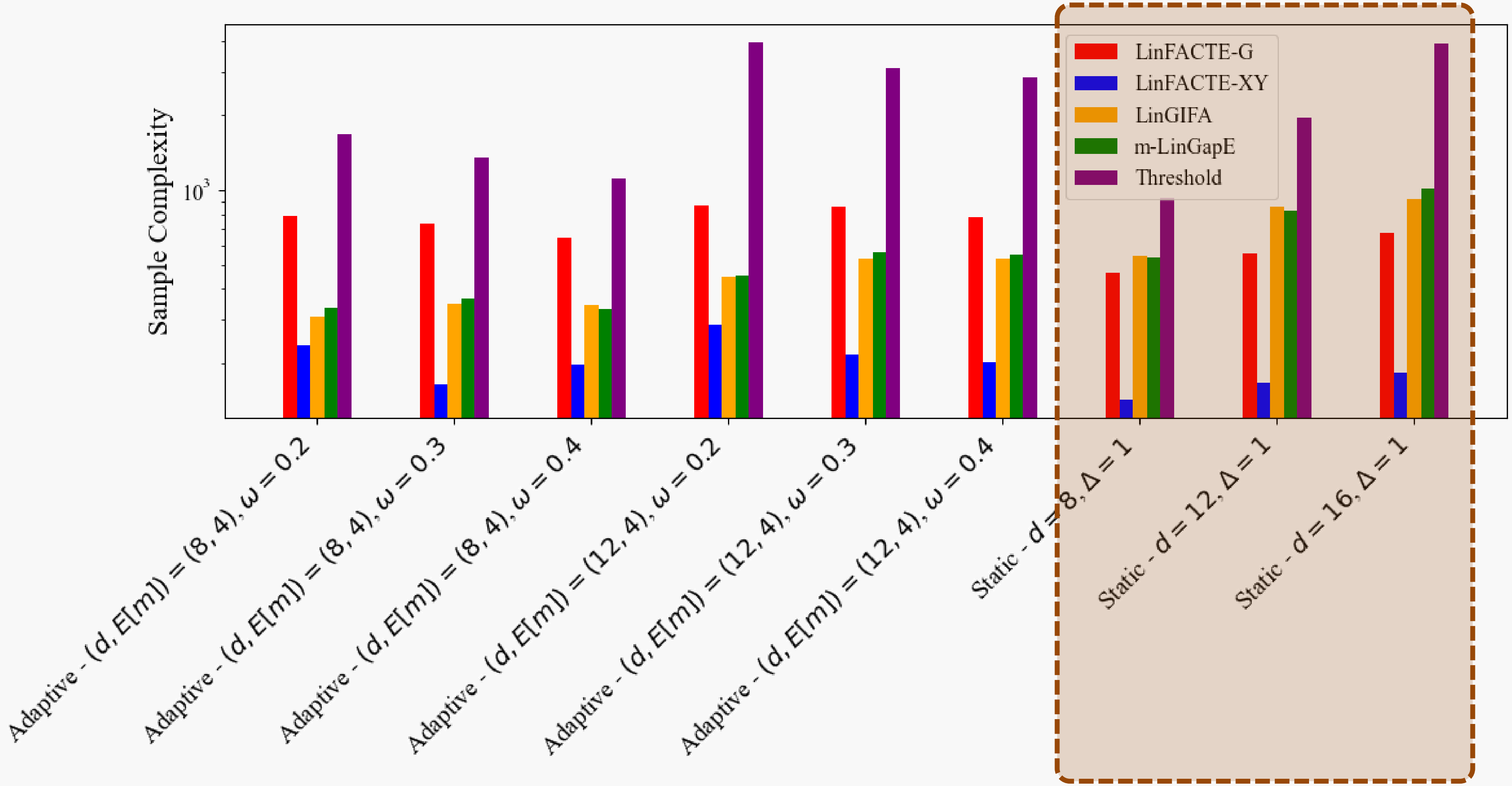
Experiment – Simulation Results

Synthetic Data – Sample Complexity



Experiment – Simulation Results

Synthetic Data – Sample Complexity

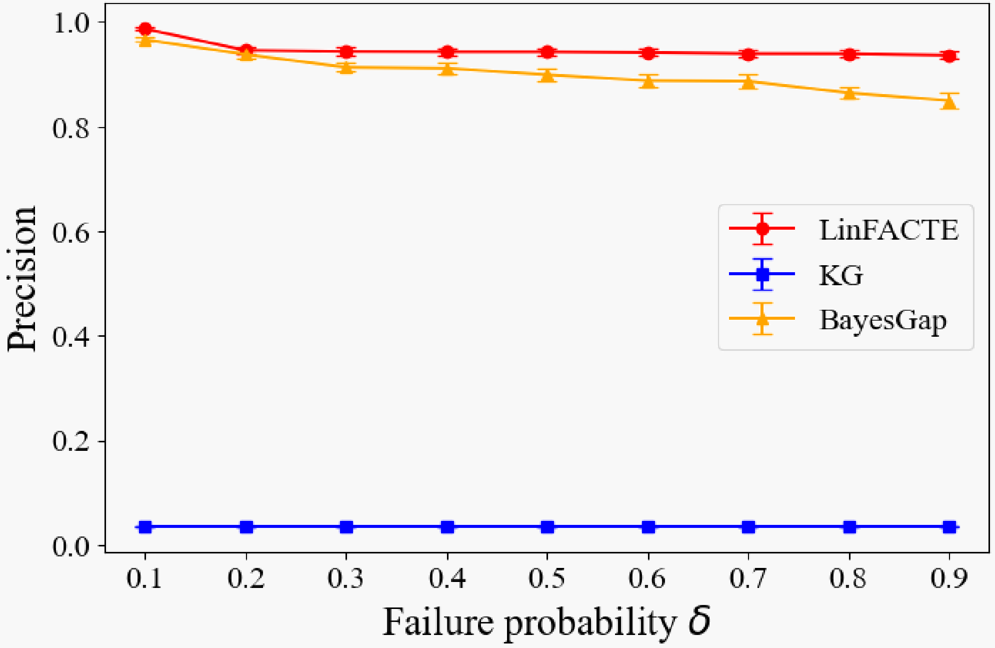


Experiment – Simulation Results

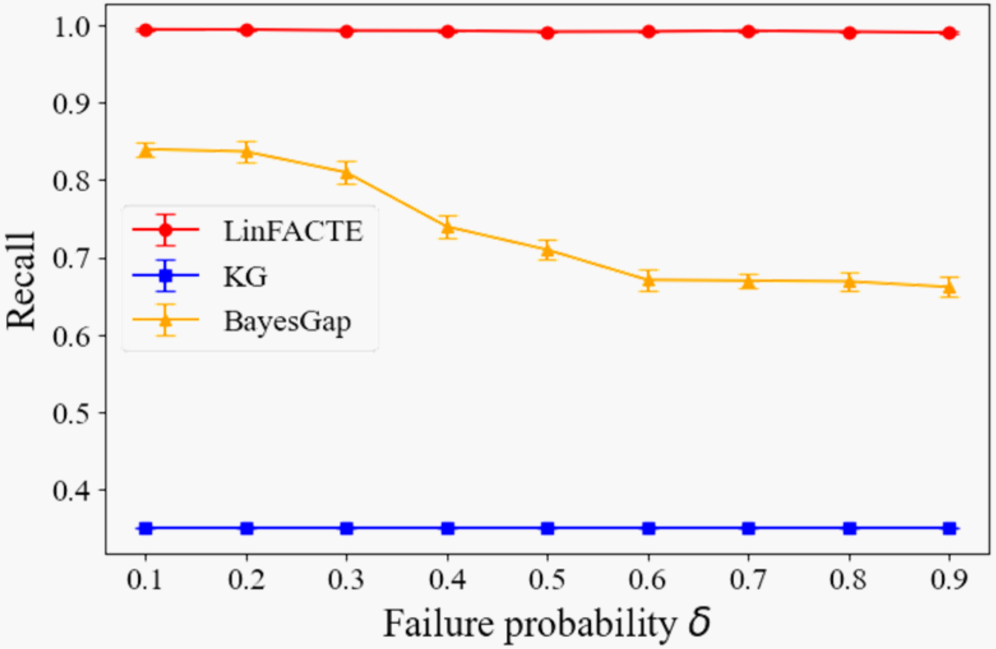
Drug Discovery with Free-Wilson Model (Negoescu et al. 2011)

arm space is extremely large

measures the accuracy of positive predictions



measures the ability to find all actual positive cases



- LinFACTE shows outstanding advantages in computational complexity → **1min < 4mins << 2 hours**
- In fact, LinFACTE is more suitable for real experiment
  - All other algorithms propose one drug and do one experiment
  - LinFACTE can propose different drugs and do multiple experiments in a batch

Conclusion

- New **Setting**: All ε-Best Arms Identification + Linear Bandits
- First Information-Theoretic **Lower Bound**
- Matching **Upper Bound**
- **Model Extensions** to Misspecified Linear Bandits and GLM
- **Numerical Simulations** with Synthetic Data and Real Data

For any further questions, please contact: ***zhekai_li.work@sjtu.edu.cn***

Thank you for your attention!

For any further questions, please contact: ***zhekai_li.work@sjtu.edu.cn***