# Novelty Assessment Report

**Paper**: Breaking the Total Variance Barrier: Sharp Sample Complexity for Linear Heteroscedastic Bandits with Fixed Action Set

**PDF URL**: https://openreview.net/pdf?id=IkPJocP3ju

**Venue**: ICLR 2026 Conference Submission

**Year**: 2026

**Report Generated**: 2025-12-29

## Abstract


Recent years have witnessed increasing interests in tackling heteroscedastic noise in bandits and reinforcement learning \citep{zhou2021nearly, zhao2023variance, jia2024does, pacchiano2025second}. In these works, the cumulative variance of the noise $\Lambda = \sum_{t=1}^T \sigma_t^2$, where $\sigma_t^2$ is the variance of the noise at round T, has been used to characterize the statistical complexity of the problem, yielding simple regret bounds of order $\tilde{\mathcal O}(d \sqrt{\Lambda / T^2})$ for linear bandits with heteroscedastic noise \citep{zhou2021nearly, zhao2023variance}. However, with a closer look, $\Lambda$ remains the same order even if the noise is close to zero at half of the rounds, which indicates that the $\Lambda$-dependence is not optimal.

In this paper, we revisit the linear bandit problem with heteroscedastic noise. We consider the setting where the action set is fixed throughout the learning process. We propose a novel variance-adaptive algorithm VAEE (Variance-Aware Exploration with Elimination) for large action set, which actively explores actions that maximizes the information gain among a candidate set of actions that are not eliminated. With the active-exploration strategy, we show that VAEE achieves a simple regret with a nearly harmonic-mean dependent rate, i.e. $\tilde{\mathcal O}\Big(d\Big[\sum_{t = 1}^T \frac{1}{\sigma_t^2} - \sum_{i = 1}^{\tilde{O}(d)} \frac{1}{[\sigma^{(i)}]^2} \Big]^{-\frac{1}{2}}\Big)$ where D is the dimension of the feature space and $\sigma^{(i)}$ is the I-th smallest variance among $\{\sigma_t\}_{t=1}^T$. For finitely many actions, we propose a variance-aware variant of G-optimal design based exploration, which achieves a $\tilde {\mathcal O}$ $\bigg(\sqrt{d \log |\mathcal A| }\Big[ \sum_{t = 1}^T \frac{1}{\sigma_t^2}-\sum_{i = 1}^{\tilde{O}(d)} \frac{1}{[\sigma^{(i)}]^2} \Big]^{-\frac{1}{2}}\bigg)$ simple regret bound. We also establish a nearly matching lower bound for the fixed action set setting indicating that \emph{harmonic-mean} dependent rate is unavoidable. To the best of our knowledge, this is the first work that breaks the $\sqrt{\Lambda}$ barrier for linear bandits with heteroscedastic noise.



**Disclaimer**

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn


## Core Task Landscape

This paper addresses: **Variance-Adaptive Exploration in Linear Bandits with Heteroscedastic Noise**

A total of **26 papers** were analyzed and organized into a taxonomy with **15 categories**.

### Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Variance-Adaptive Exploration Mechanisms**
- **Non-Stationary and Time-Varying Environments**
- **Noise Structure and Dependency Modeling**
- **Specialized Contextual and Structural Settings**
- **Application-Driven Bandit Methods**

### Complete Taxonomy Tree

- Variance-Adaptive Exploration in Linear Bandits with Heteroscedastic Noise Survey Taxonomy
- Variance-Adaptive Exploration Mechanisms
  - Heteroscedastic Noise Adaptation with Fixed Action Sets ★ (4 papers)
  - [0] Breaking the Total Variance Barrier: Sharp Sample Complexity for Linear Heteroscedastic Bandits with Fixed Action Set (Anon et al., 2026) View paper
  - [2] Information directed sampling and bandits with heteroscedastic noise (Kirschner, 2018) View paper
  - [8] Noise-Adaptive Confidence Sets for Linear Bandits and Application to Bayesian Optimization (Jun, 2024) View paper
  - [12] Data Source Adaptive Online Learning under Heteroscedastic Noise (ABH Anand, 2025) View paper
  - Variance-Aware Confidence Set Construction (4 papers)
  - [14] Variance-Dependent Regret Bounds for Linear Bandits and Reinforcement Learning: Adaptivity and Computational Efficiency (Zhao, 2023) View paper
  - [22] Improved Variance-Aware Confidence Sets for Linear Bandits and Linear\n Mixture MDP (Zhang Zi-han, 2021) View paper
  - [25] Improved Regret Analysis for Variance-Adaptive Linear Bandits and Horizon-Free Linear Mixture MDPs (Kim, 2021) View paper
  - Bayesian Variance-Adaptive Approaches (2 papers)
  - [16] Empirical Bound Information-Directed Sampling for Norm-Agnostic Bandits (Suder, 2025) View paper
  - [20] Only Pay for What Is Uncertain: Variance-Adaptive Thompson Sampling (Saha, 2023) View paper
  - Neural Network-Enhanced Variance Adaptation (2 papers)
  - [1] LNUCB-TA: Linear-nonlinear Hybrid Bandit Learning with Temporal Attention (Khosravi, 2025) View paper
  - [7] Variance-Aware Linear UCB with Deep Representation for Neural Contextual Bandits (Bui, 2024) View paper
- Non-Stationary and Time-Varying Environments
  - Variance-Dependent Non-Stationary Regret Bounds (2 papers)
  - [5] Variance-Dependent Regret Bounds for Non-stationary Linear Bandits (Wang Zhiyong, 2024) View paper

- [10] From Theory to Practice with RAVEN-UCB: Addressing Non-Stationarity in Multi-Armed Bandits through Variance Adaptation (Fang, 2025) View paper
- Latent State and Auto-Regressive Dynamics (2 papers)
- [3] Non-Stationary Latent Auto-Regressive Bandits (Trella, 2024) View paper
- [24] Recurrent Neural-Linear Posterior Sampling for Nonstationary Contextual Bandits. (Aditya Ramesh, 2022) View paper
- Time-Varying Gaussian Process Optimization (1 papers)
- [15] No-Regret Gaussian Process Optimization of Time-Varying Functions (Eliabelle Mauduit, 2025) View paper
- High-Dimensional Non-Stationary Recommender Systems (2 papers)
- [4] Non-Stationary Linear Bandits With Dimensionality Reduction for Large-Scale Recommender Systems (Saeed Ghoorchian, 2024) View paper
- [18] Bayesian Non-stationary Linear Bandits for Large-Scale Recommender Systems (Ghoorchian, 2022) View paper

- Noise Structure and Dependency Modeling
  - Non-i.i.d. and Temporally Dependent Noise (1 papers)
  - [6] Linear Bandits with Non-i.i.d. Noise (Clerico, 2025) View paper
  - Best-of-Both-Worlds and Adversarial Corruption Robustness (1 papers)
  - [19] Best-of-Three-Worlds Linear Bandit Algorithm with Variance-Adaptive Regret Bounds (Ito Shinji, 2023) View paper
  - Generalized Linear Models with Unbounded Noise (1 papers)
  - [21] Optimal Online Generalized Linear Regression with Stochastic Noise and Its Application to Heteroscedastic Bandits (Zhao, 2022) View paper
- Specialized Contextual and Structural Settings
  - Federated and Distributed Linear Bandits (1 papers)
  - [11] Federated Linear Bandits with Finite Adversarial Actions (Fan Li, 2023) View paper
  - Network and Peer Influence Learning (1 papers)
  - [13] Learning Peer Influence Probabilities with Linear Contextual Bandits (ShahverdiKondori, 2025) View paper
- Application-Driven Bandit Methods
  - Dynamic Pricing with Contextual Bandits (2 papers)
  - [9] Contextual dynamic pricing with unknown noise: Explore-then-ucb strategy and improved regrets (Y Luo, 2022) View paper
  - [26] Deep Reinforcement Learning for Dynamic Pricing Strategies: Empirical Evidence from E-Commerce Platforms (SM Ameli, n.d.) View paper
  - Wireless Resource Allocation and Scheduling (1 papers)
  - [17] An Attention-Based Reward Framework for Multi-Armed Bandit Scheduling in 5G Uplink (Sanil Fulani, 2025) View paper

### Narrative

Core task: variance-adaptive exploration in linear bandits with heteroscedastic noise. The field addresses sequential decision-making under uncertainty when reward noise varies across actions or contexts, requiring algorithms that adapt their exploration strategies to heterogeneous variance structures. The taxonomy organizes research into several main branches: Variance-Adaptive Exploration Mechanisms develop core algorithmic principles for handling action-dependent noise, often through confidence-bound adjustments or information-directed sampling; Non-Stationary and Time-Varying Environments extend these ideas to settings where reward distributions drift over time; Noise Structure and Dependency Modeling examines correlated or non-iid noise patterns; Specialized Contextual and Structural Settings incorporate domain-specific constraints such as graph structures or latent representations; and Application-Driven Bandit Methods translate theoretical insights into practical domains like dynamic pricing and recommendation systems. Representative works include Heteroscedastic IDS[2], which pioneered information-theoretic approaches to variance adaptation, and LNUCB-TA[1], which combines variance-aware confidence bounds with time-adaptive mechanisms.

A particularly active line of work focuses on designing tight, instance-dependent confidence intervals that scale with local noise levels, as seen in Noise Adaptive Confidence[8] and Adaptive Heteroscedastic Learning[12], which refine exploration bonuses to avoid over-exploration in low-variance regions. Meanwhile, several studies tackle the interplay between nonstationarity and heteroscedasticity, such as Variance Dependent Nonstationary[5] and Nonstationary Latent Bandits[3], exploring how variance adaptation must account for temporal drift. The original paper, Heteroscedastic Bandits[0], sits squarely within the Variance-Adaptive Exploration Mechanisms branch, specifically addressing heteroscedastic noise adaptation with fixed action sets. Its emphasis on principled variance estimation and regret bounds aligns closely with Noise Adaptive Confidence[8] and Adaptive Heteroscedastic Learning[12], though it appears to focus more directly on theoretical guarantees for fixed-arm scenarios rather than the time-varying or contextual extensions explored by neighbors like Nonstationary Latent Bandits[3].

## Related Works in Same Category

The following **3 sibling papers** share the same taxonomy leaf node with the original paper:

### 1. Information directed sampling and bandits with heteroscedastic noise

**Authors**: Kirschner, Johannes, Krause, Andreas, Johannes Kirschner, et al. (6 authors total) | **Year/Venue**: 2018 | **URL**: View paper

#### Abstract

In the stochastic bandit problem, the goal is to maximize an unknown function via a sequence of noisy function evaluations. Typically, the observation noise is assumed to be independent of the evaluation point and satisfies a tail bound taken uniformly on the domain. In this work, we consider the setting of heteroscedastic noise, that is, we explicitly allow the noise distribution to depend on the evaluation point. We show that this leads to new trade-offs for information and regret, which are n...

#### Relationship Analysis

Both papers belong to the category of heteroscedastic noise adaptation with fixed action sets, addressing variance-dependent exploration in linear bandits. The original paper focuses on breaking the $\sqrt{\Lambda}$ barrier by achieving harmonic-mean dependent regret bounds through variance-aware elimination and G-optimal design, while the candidate paper (Kirschner & Krause, 2018) introduces Information Directed Sampling (IDS) as a frequentist framework that balances regret and information gain through the regret-information ratio. The key difference is that the original paper achieves sharper instance-dependent bounds by exploiting the harmonic mean of variances with elimination strategies, whereas the candidate paper proposes a more general IDS framework with weighted least squares estimation but does not establish the same refined harmonic-mean dependence.

---

### 2. Noise-Adaptive Confidence Sets for Linear Bandits and Application to Bayesian Optimization

**Authors**: Jun, Kwang-Sung, Kim Jungtaek, Kwang-Sung Jun, Jungtaek Kim | **Year/Venue**: 2024 | **URL**: View paper

### Abstract

Adapting to a priori unknown noise level is a very important but challenging problem in sequential decision-making as efficient exploration typically requires knowledge of the noise level, which is often loosely specified. We report significant progress in addressing this issue for linear bandits in two respects. First, we propose a novel confidence set that is `semi-adaptive' to the unknown sub-Gaussian parameter $\sigma_*^2$ in the sense that the (normalized) confidence width scales with $\sqr...


### Relationship Analysis

Both papers belong to the same taxonomy category of heteroscedastic noise adaptation with fixed action sets, addressing variance-dependent exploration in linear bandits. The original paper focuses on breaking the √Λ barrier by achieving harmonic-mean dependent regret bounds through variance-aware exploration and elimination strategies (VAEE and VAGD algorithms), while the candidate paper (Jun & Kim) proposes noise-adaptive confidence sets that semi-adapt to unknown sub-Gaussian parameters and fully adapt to bounded noise variances, achieving improved regret bounds through novel confidence set constructions (LOSAN and LOFAV algorithms). The key difference is that the original paper explicitly targets harmonic-mean dependence on variance sequences with active elimination, whereas the candidate paper focuses on constructing tighter confidence sets that adapt to noise levels without requiring prior knowledge of exact variance values.

---

## 3. Data Source Adaptive Online Learning under Heteroscedastic Noise

**Authors**: ABH Anand, A Saha, TK Buening, H Luo | **Year/Venue**: 2025 | **URL**: View paper


### Abstract

â¦ We highlight two natural baselines for the heterogeneous multisource bandit problem. The â¦ on adaptive exploration across multiple sources characterized by differing noise variances â¦


### Relationship Analysis

Both papers belong to the category of heteroscedastic noise adaptation with fixed action sets, addressing variance-dependent exploration in bandit settings. The original paper focuses on linear bandits with continuous action spaces and achieves harmonic-mean dependent regret bounds that break the √Λ barrier through variance-aware exploration and elimination strategies. The candidate paper addresses multi-armed bandits with multiple heterogeneous data sources (each with distinct noise variances), proposing a source-selection mechanism that adaptively filters high-variance sources while maintaining UCB-based arm selection, achieving regret bounds dependent on the minimum source variance rather than the action-dependent noise structure studied in the original work.

# Contributions Analysis

**Overall novelty summary.** The paper proposes VAEE, a variance-adaptive algorithm for linear bandits with heteroscedastic noise in fixed action sets. It resides in the 'Heteroscedastic Noise Adaptation with Fixed Action Sets' leaf, which contains four papers total including this work. This leaf sits within the broader 'Variance-Adaptive Exploration Mechanisms' branch, indicating a moderately populated research direction focused on action-dependent noise variance. The paper's core claim is that existing cumulative variance bounds (Λ-dependence) are suboptimal and proposes a harmonic-mean-based characterization for tighter regret guarantees.

The taxonomy reveals neighboring leaves addressing variance-aware confidence sets, Bayesian approaches, and neural network enhancements, all within the same parent branch. Adjacent branches tackle non-stationary environments and noise dependency modeling, suggesting the field has diversified into temporal dynamics and correlation structures. The paper's focus on fixed action sets distinguishes it from time-varying environment methods in sibling branches, while its theoretical emphasis on regret bounds connects it to confidence-set construction techniques explored in parallel leaves. The scope notes clarify that methods without explicit variance incorporation belong elsewhere, positioning this work firmly in the variance-adaptive core.

Among ten candidates examined across three contributions, no refutable prior work was identified. Contribution A (VAEE algorithm) examined four candidates with zero refutations; Contribution B (G-optimal design exploration) examined six candidates, also with zero refutations; Contribution C (lower bound) was not evaluated against prior work. This limited search scope—covering roughly one-third of the taxonomy's twenty-six papers—suggests the analysis captures immediate neighbors but may not reflect the full landscape. The absence of refutations among examined candidates indicates the specific algorithmic mechanisms and theoretical characterizations appear distinct within this sample.

Based on the top-ten semantic matches examined, the work appears to occupy a recognizable but not densely crowded position within variance-adaptive linear bandits. The harmonic-mean variance characterization and elimination-based exploration strategy differentiate it from the four sibling papers in its leaf, though the limited search scope prevents definitive claims about novelty relative to the broader field. The taxonomy structure suggests this is an active area with multiple complementary approaches, and the paper's theoretical contributions address a specific gap in variance-dependent regret analysis.

---

This paper presents **3 main contributions**, each analyzed against relevant prior work:

## Contribution 1: Variance-Aware Exploration with Elimination (VAEE) algorithm for large action sets

**Description**: The authors introduce VAEE, a variance-adaptive algorithm designed for linear bandits with large or infinite action sets. The algorithm maintains a candidate set of promising actions and actively explores those that maximize information gain subject to elimination rules, achieving a simple regret bound with nearly harmonic-mean dependent rate on the noise variances.

This contribution was assessed against **4 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

### 1. Federated Linear Bandits with Finite Adversarial Actions

**URL**: View paper

**Brief Assessment**

Federated Adversarial[11] focuses on federated linear bandits with finite adversarial action sets in a distributed setting, not variance-adaptive exploration with elimination for large/infinite action sets in a centralized setting.

---

### 2. Risk-Aware Linear Bandits: Theory and Applications in Smart Order Routing

**URL**: View paper

**Brief Assessment**

Risk-Aware Smart Routing[29] focuses on risk-aware linear bandits with mean-variance objectives for smart order routing, not variance-adaptive exploration with elimination for heteroscedastic noise in general linear bandits.

---

### 3. Multi-Metric Adaptive Experimental Design under Fixed Budget with Validation

**URL**: View paper

**Brief Assessment**

Multi-Metric Adaptive[28] focuses on multi-metric adaptive experimental design with exploration and validation phases for treatment selection, not variance-adaptive linear bandits with large action sets. The candidate addresses a different problem domain (A/B testing with multiple metrics) rather than heteroscedastic linear bandits.

### 4. Variance-aware decision making with linear function approximation under heavy-tailed rewards

**URL**: View paper

**Brief Assessment**

Heavy-Tailed Variance[27] focuses on linear bandits with heavy-tailed rewards (finite variances only), not heteroscedastic noise with varying sub-gaussian variances. The candidate uses adaptive Huber regression for robustness, while VAEE uses variance-weighted least squares with active exploration and elimination for heteroscedastic settings.

## Contribution 2: Variance-aware G-optimal design based exploration for finite action sets

**Description**: The authors develop a variance-adaptive variant of G-optimal design based exploration specifically for finite action sets. This approach achieves improved simple regret bounds with better dependence on the dimension d compared to the infinite action set case, replacing the $\sqrt{d}$ factor with $\sqrt{d \log |A|}$.

This contribution was assessed against **6 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

### 1. Adaptive Data Collection for Policy Evaluation, Multi-task Learning and Llm Alignment

**URL**: View paper

**Brief Assessment**

Adaptive Policy Evaluation[32] focuses on policy evaluation and LLM alignment applications, not on linear bandits with heteroscedastic noise. The candidate's context mentions G-optimal design and variance-aware techniques only in passing without detailed algorithmic development for finite action linear bandits.

### 2. Declaration of Committee

**URL**: View paper

**Brief Assessment**

Declaration Committee[31] focuses on constrained MDPs with linear function approximation and primal-dual frameworks, not on variance-aware G-optimal design exploration for finite action linear bandits. The technical settings and problems are fundamentally different.

### 3. Federated Linear Bandits with Finite Adversarial Actions

**URL**: View paper

**Brief Assessment**

Federated Adversarial[11] addresses federated learning with finite adversarial actions but does not propose variance-aware G-optimal design exploration. The candidate's focus is on federated communication protocols, not variance-adaptive exploration strategies.

### 4. No-Regret Linear Bandits under Gap-Adjusted Misspecification

**URL**: View paper

**Brief Assessment**

Gap-Adjusted Misspecification[30] focuses on misspecified linear bandits with gap-adjusted approximation errors, not variance-aware exploration for heteroscedastic noise in finite action sets.

### 5. Risk-Aware Linear Bandits: Theory and Applications in Smart Order Routing

**URL**: View paper

**Brief Assessment**

Risk-Aware Smart Routing[29] uses G-optimal design for mean-variance optimization in financial applications, not for variance-adaptive exploration in heteroscedastic linear bandits with improved dimension dependence.

### 6. Risk-Aware Linear Bandits with Application in Smart Order Routing

**URL**: View paper

**Brief Assessment**

Risk-Aware Smart Routing[33] focuses on risk-aware linear bandits with mean-variance objectives for financial applications, while the original paper addresses heteroscedastic noise in best-arm identification. The candidate uses G-optimal design for variance minimization in a risk-aversion context, not for variance-adaptive exploration under heteroscedastic noise.

## Contribution 3: Nearly matching lower bound demonstrating harmonic-mean dependence is unavoidable

**Description**: The authors prove a nearly matching instance-dependent lower bound for the fixed action set setting, showing that the harmonic-mean dependence on noise variances is fundamental to the problem. This establishes that their algorithms are essentially optimal in their variance dependence.

This contribution was assessed against **0 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

# Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

# References


- [0] Breaking the Total Variance Barrier: Sharp Sample Complexity for Linear Heteroscedastic Bandits with Fixed Action Set View paper
- [1] LNUCB-TA: Linear-nonlinear Hybrid Bandit Learning with Temporal Attention View paper
- [2] Information directed sampling and bandits with heteroscedastic noise View paper
- [3] Non-Stationary Latent Auto-Regressive Bandits View paper
- [4] Non-Stationary Linear Bandits With Dimensionality Reduction for Large-Scale Recommender Systems View paper
- [5] Variance-Dependent Regret Bounds for Non-stationary Linear Bandits View paper

- [6] Linear Bandits with Non-i.i.d. Noise View paper
- [7] Variance-Aware Linear UCB with Deep Representation for Neural Contextual Bandits View paper
- [8] Noise-Adaptive Confidence Sets for Linear Bandits and Application to Bayesian Optimization View paper
- [9] Contextual dynamic pricing with unknown noise: Explore-then-ucb strategy and improved regrets View paper
- [10] From Theory to Practice with RAVEN-UCB: Addressing Non-Stationarity in Multi-Armed Bandits through Variance Adaptation View paper
- [11] Federated Linear Bandits with Finite Adversarial Actions View paper
- [12] Data Source Adaptive Online Learning under Heteroscedastic Noise View paper
- [13] Learning Peer Influence Probabilities with Linear Contextual Bandits View paper
- [14] Variance-Dependent Regret Bounds for Linear Bandits and Reinforcement Learning: Adaptivity and Computational Efficiency View paper
- [15] No-Regret Gaussian Process Optimization of Time-Varying Functions View paper
- [16] Empirical Bound Information-Directed Sampling for Norm-Agnostic Bandits View paper
- [17] An Attention-Based Reward Framework for Multi-Armed Bandit Scheduling in 5G Uplink View paper
- [18] Bayesian Non-stationary Linear Bandits for Large-Scale Recommender Systems View paper
- [19] Best-of-Three-Worlds Linear Bandit Algorithm with Variance-Adaptive Regret Bounds View paper
- [20] Only Pay for What Is Uncertain: Variance-Adaptive Thompson Sampling View paper
- [21] Optimal Online Generalized Linear Regression with Stochastic Noise and Its Application to Heteroscedastic Bandits View paper
- [22] Improved Variance-Aware Confidence Sets for Linear Bandits and Linear\n Mixture MDP View paper
- [23] Improved Variance-Aware Confidence Sets for Linear Bandits and Linear Mixture MDP View paper
- [24] Recurrent Neural-Linear Posterior Sampling for Nonstationary Contextual Bandits. View paper
- [25] Improved Regret Analysis for Variance-Adaptive Linear Bandits and Horizon-Free Linear Mixture MDPs View paper
- [26] Deep Reinforcement Learning for Dynamic Pricing Strategies: Empirical Evidence from E-Commerce Platforms View paper
- [27] Variance-aware decision making with linear function approximation under heavy-tailed rewards View paper
- [28] Multi-Metric Adaptive Experimental Design under Fixed Budget with Validation View paper
- [29] Risk-Aware Linear Bandits: Theory and Applications in Smart Order Routing View paper
- [30] No-Regret Linear Bandits under Gap-Adjusted Misspecification View paper
- [31] Declaration of Committee View paper
- [32] Adaptive Data Collection for Policy Evaluation, Multi-task Learning and Llm Alignment View paper
- [33] Risk-Aware Linear Bandits with Application in Smart Order Routing View paper