[1]Cloud Computing Research Institute, China Telecom

[2]Department of Computer and Information Sciences, Temple University

Change Detection Methods for Non-Stationary Stochastic Linear Bandits

Linghang Meng [1]
menglh1@chinatelecom.cn

Jie Wu [1,2]
jiewu@temple.edu

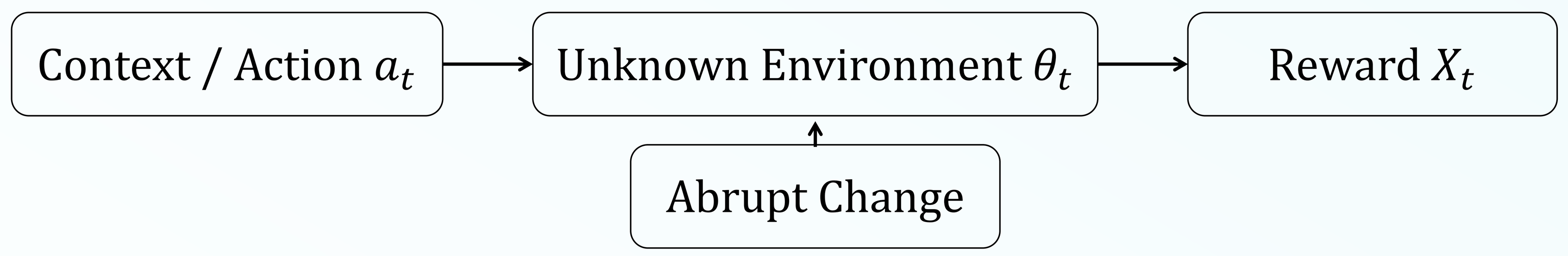
ICASSP2026 Barcelona

Take-home Message

- Abrupt changes break stationarity: outdated observations can mislead standard linear bandit algorithms.
- Linear structure enables efficient detection: a change in the shared parameter shifts rewards across actions.
- LBCD avoids forced exploration, achieves near-optimal regret, and performs strongly in experiments.

Motivation

- In many systems, the environment may remain stable for a while and then change abruptly. Classical linear bandits assume a fixed reward parameter. When the environment changes, outdated observations can mislead the learner and cause large dynamic regret.



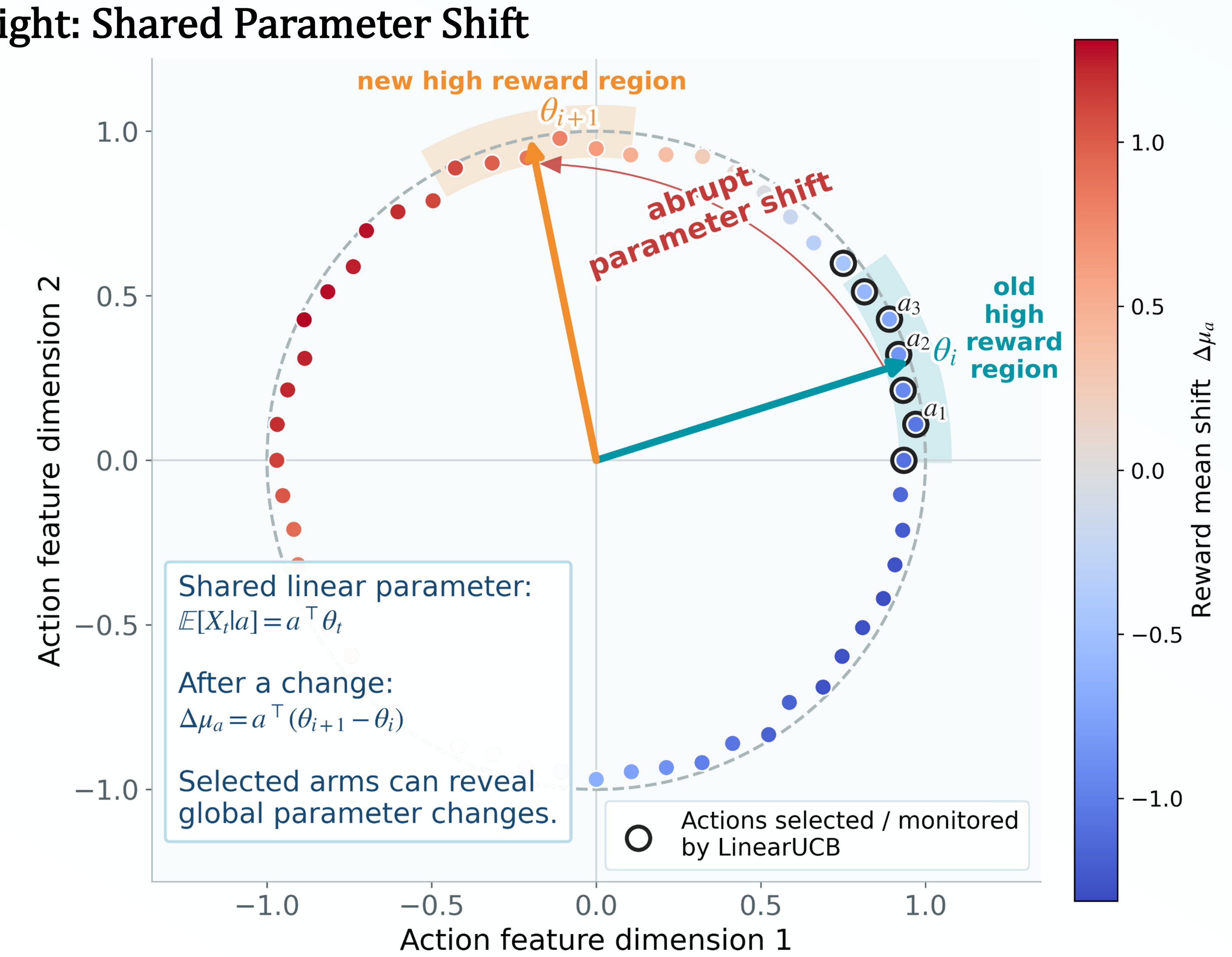
Problem Formulation

Piecewise Stationary Linear Bandits

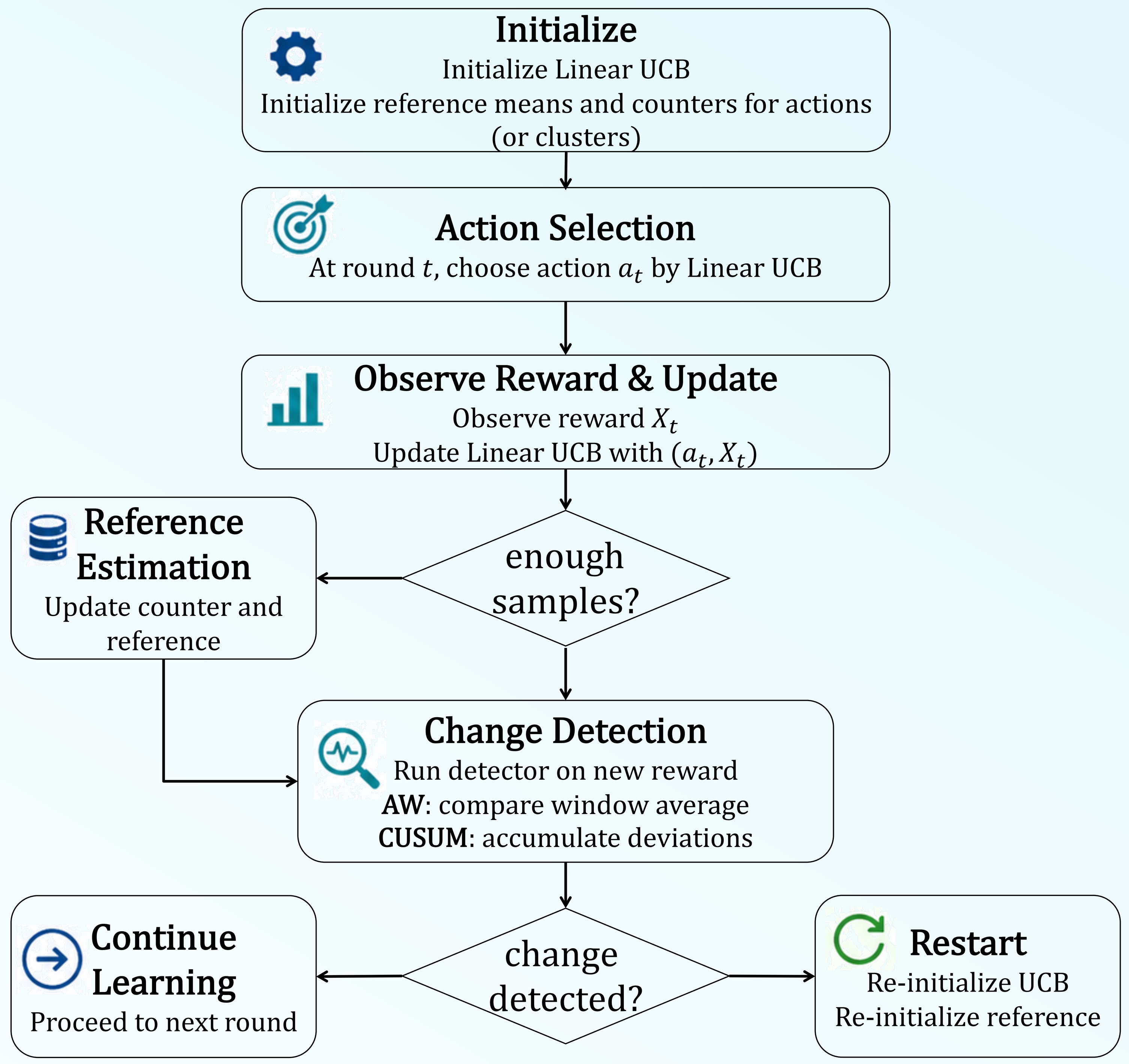
- At each round t, the leaner chooses an action $a_t \in \mathcal{A} \subset \mathbb{R}^d$ and observes $X_t = \langle A_t, \theta_t \rangle + \eta_t$.
- Environment is piecewise stationary $\theta_t = \theta^{(i)}$. Changepoints are unknown.
- The leaner aims to minimize dynamic regret
  - $R_T = \sum_{t=1}^{T} [\max_{a \in \mathcal{A}} \langle a, \theta_t \rangle - \langle A_t, \theta_t \rangle]$

$R(T) = \Omega(d\sqrt{MT})$

Key Insight: Shared Parameter Shift



Linear Bandits with Change Detection



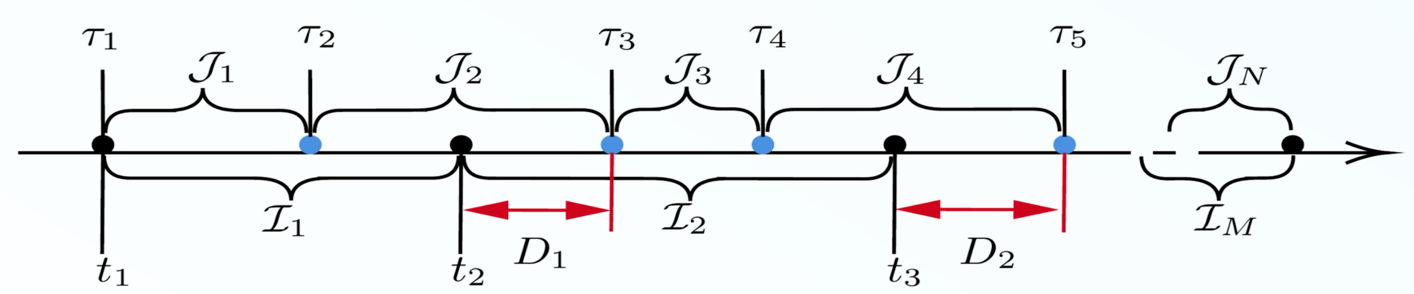
Detector	Statistic	Alarm Rule
LBCD-AW	Window Average	$\bar{X}_W - \hat{\mu}$
LBCD-CUSUM	Cumulative Deviation	$g^+ > h$ or $g^- > h$

Theoretical Guarantees

Simplified high-probability bound:

$$\tilde{O}\left(d\sqrt{(M+F)T}\right) + \tilde{O}\left(\sqrt{d(M+D+F)^3}\right)$$

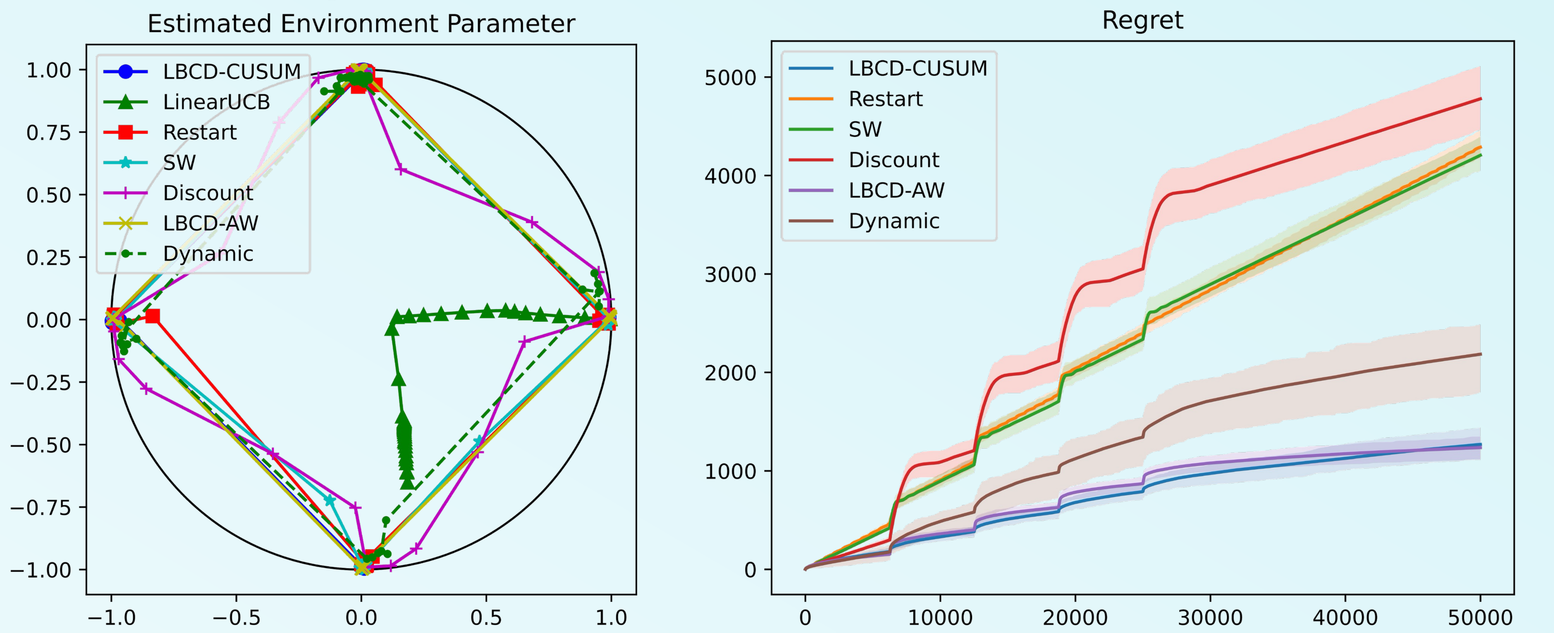
- M: Number of segments;
- F: number of false alarms;
- D: total detection delay.



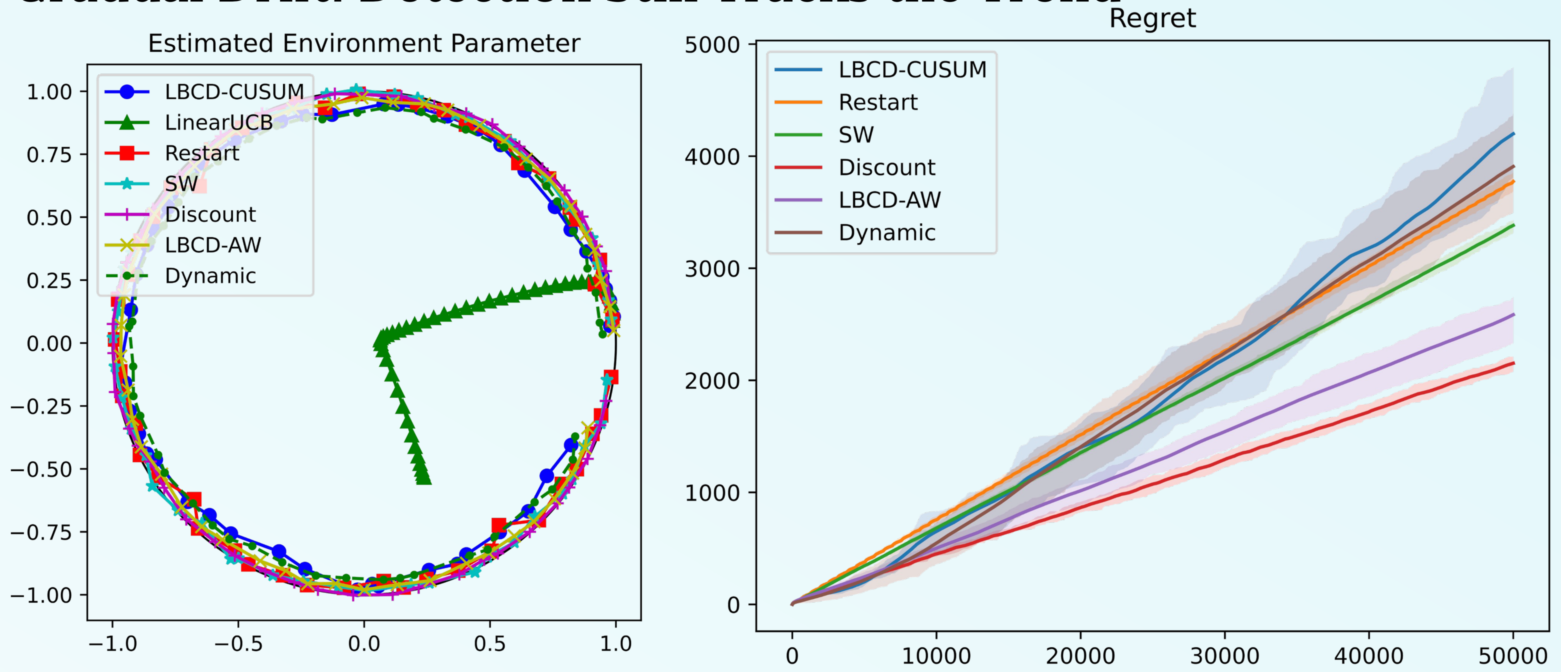
- LBCD-AW achieves nearly optimal regret in T without forced exploration, matching the lower-bound rate.

Experiments

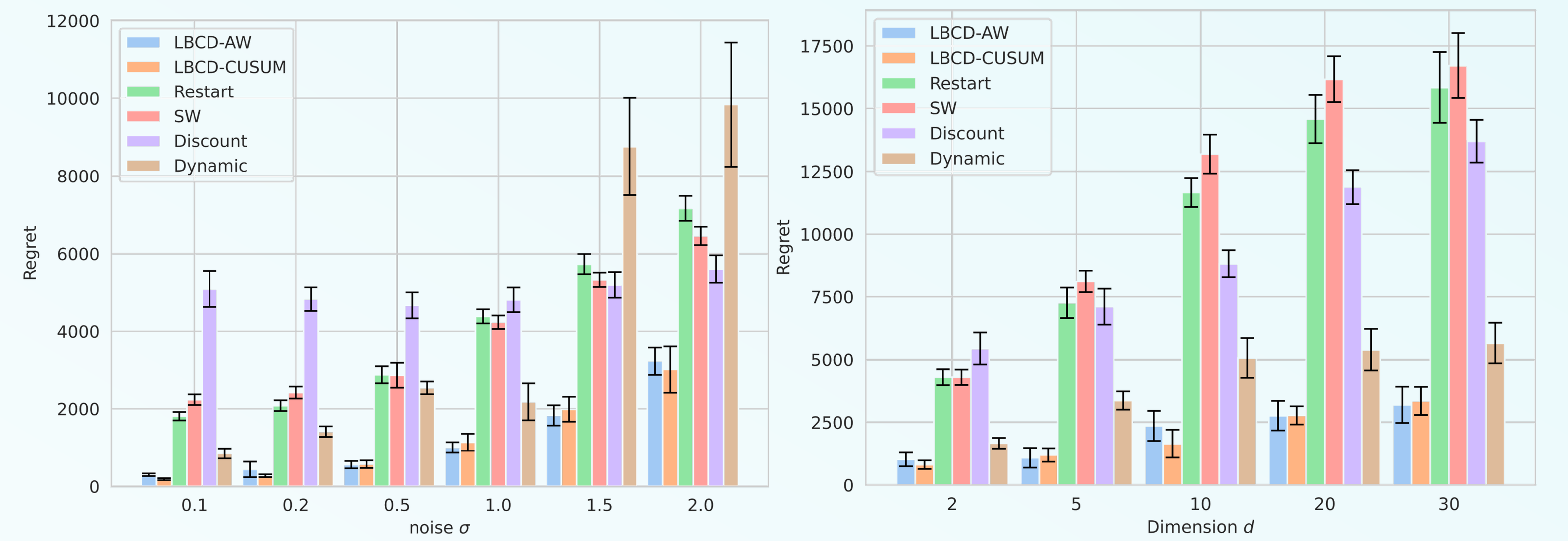
Abrupt Changes: Fast Detection and Restart



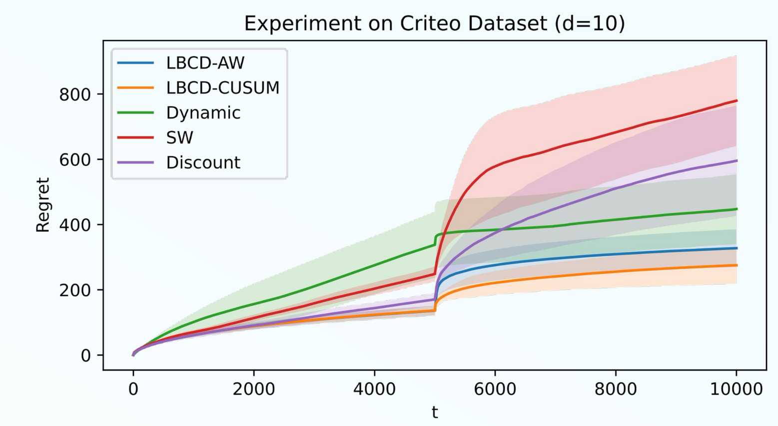
Gradual Drift: Detection Still Tracks the Trend



Robustness to Noise and Dimension



Real Dataset: Criteo Live Traffic



For large candidate action sets, the actions are clustered by contextual features, and change detection is performed at the cluster level.