POLITECNICO
MILANO 1863

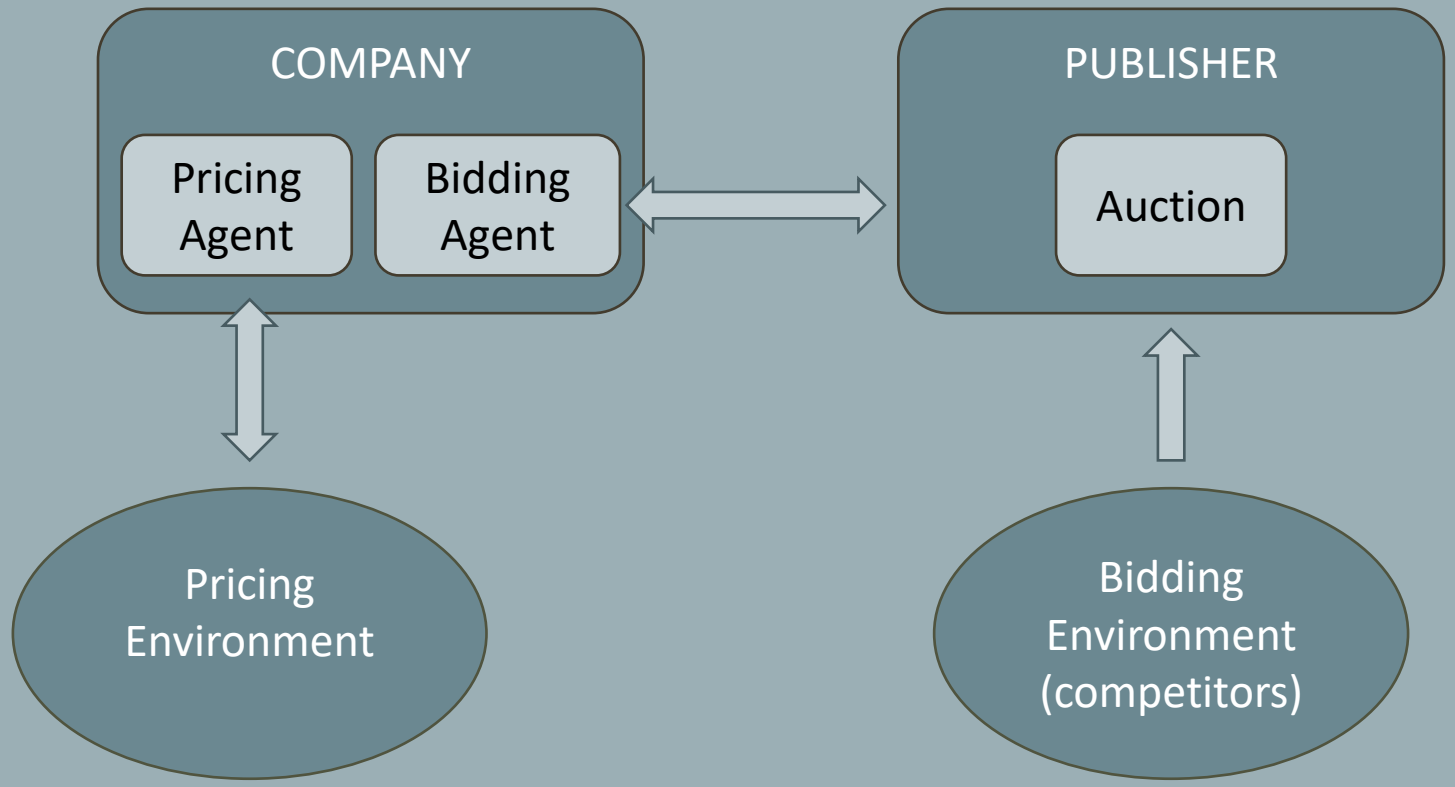
ONLINE LEARNING APPLICATIONS PROJECT

Stefano Arcaro
Gabriele Farace
Antonio Napolitano
Davide Remondina
Sofia Yang

REQUIREMENT 1

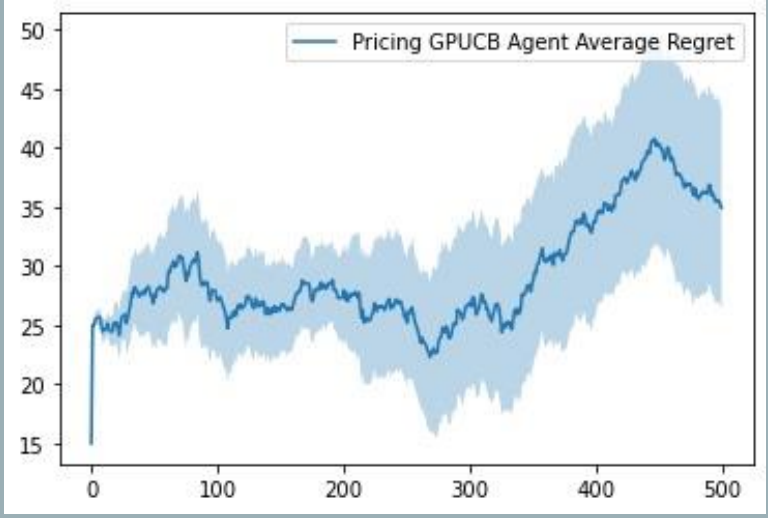
BIDDING ENVIRONMENT (COMPETITORS)	Stochastic with uniformely distributed bids
PRICING ENVIRONMENT	Stochastic with: • demands sampled from a binomial distribution with success probability given by the conversion rate $(1 - p)$ • number of trials = number of visits obtained in the bidding interaction
AUCTION TYPE	Second Price (Truthful)
BIDDING AGENTS	• UCB-like • Primal-Dual (Multiplicative Pacing)
PRICING AGENT	Gaussian Process (continuous action set)

INTERACTION



PRICING ONLY

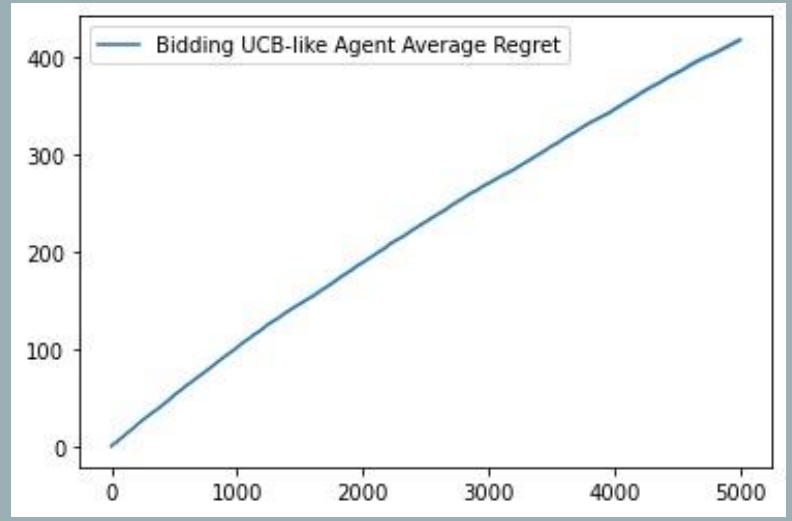
NUMBER OF INTERACTIONS	500
NUMBER OF CUSTOMERS	50
PRODUCT COST	0.1
NUMBER OF TRIALS	10



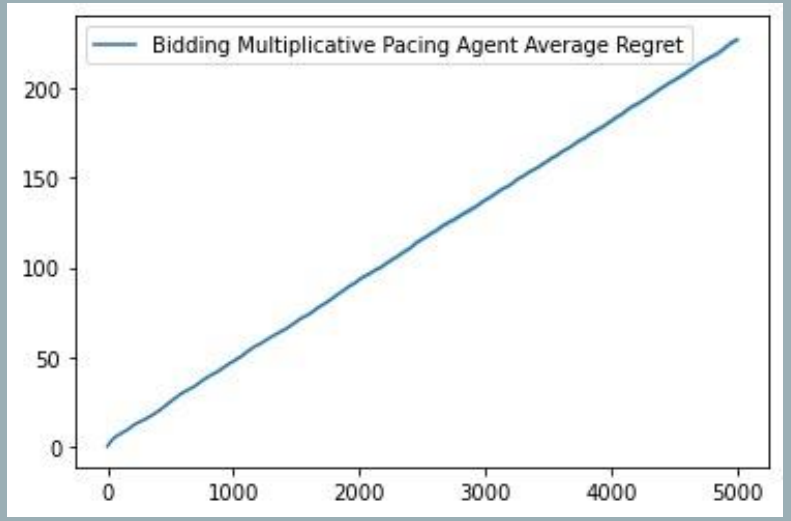
GPUCB

BIDDING ONLY

NUMBER OF INTERACTIONS	5000
BUDGET	1000
VALUATION	1
CTRs	[0.8 0.5 0.9 1]
NUMBER OF TRIALS	10



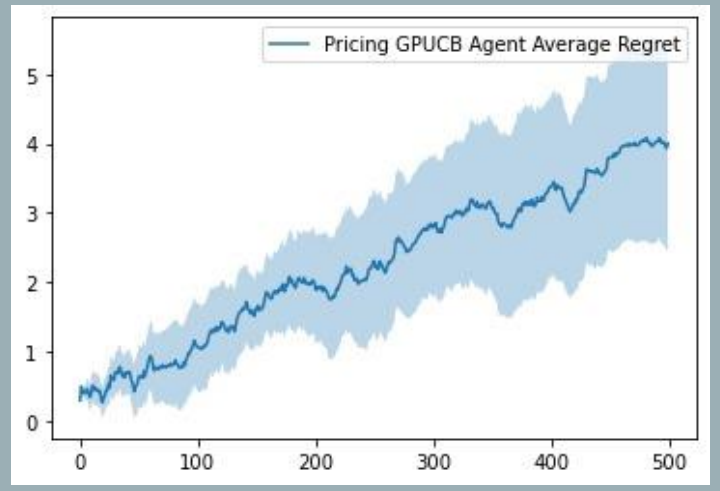
UCB-like

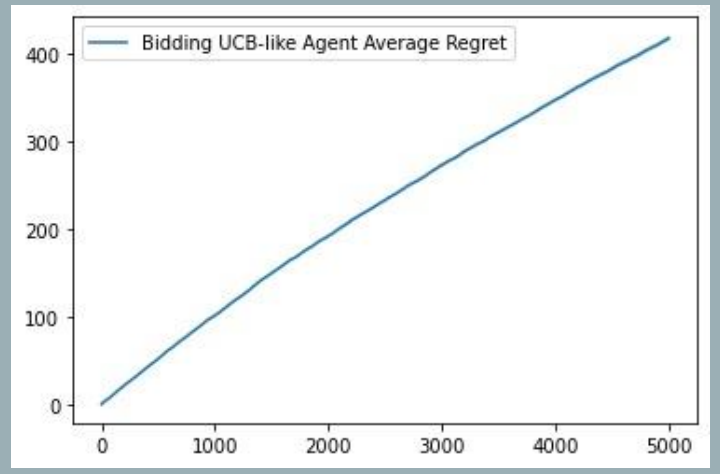


Multiplicative Pacing

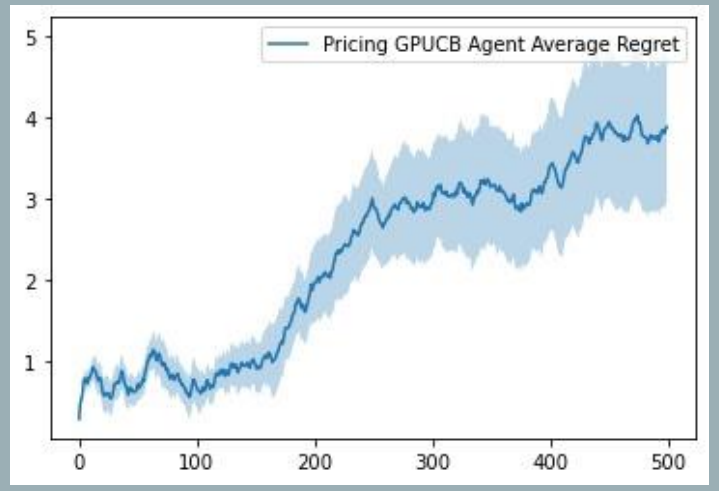
PRICING AND BIDDING

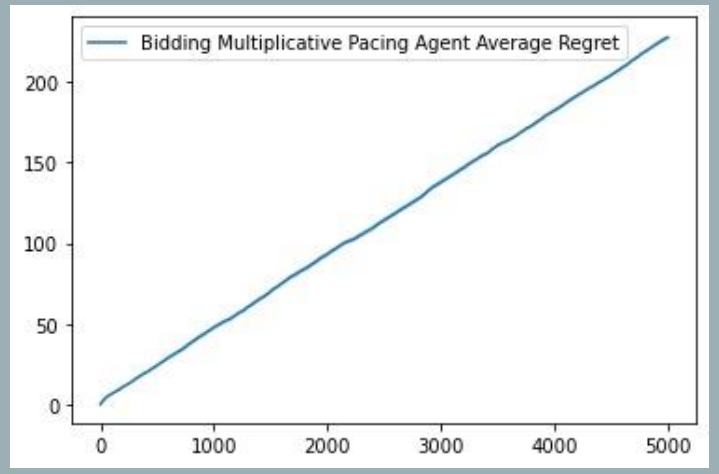
UCB-LIKE BIDDER





PRIMAL-DUAL BIDDER





REQUIREMENT 2

BIDDING ENVIRONMENT (COMPETITORS)	• Adversarial with uniformely distributed bids • Extremes of the interval sampled from a beta distribution each round
PRICING ENVIRONMENT	Adversarial with: • demands sampled from a binomial distribution with success probability given by the conversion rate $(1 - \theta * p)$ • θ sampled from a beta distribution each round • number of trials = number of visits obtained in the bidding interaction
AUCTION TYPE	Generalized First-Price (Non-truthful)
BIDDING AGENTS	Primal-Dual (Multiplicative Pacing)
PRICING AGENT	EXP3 (discretized prices)

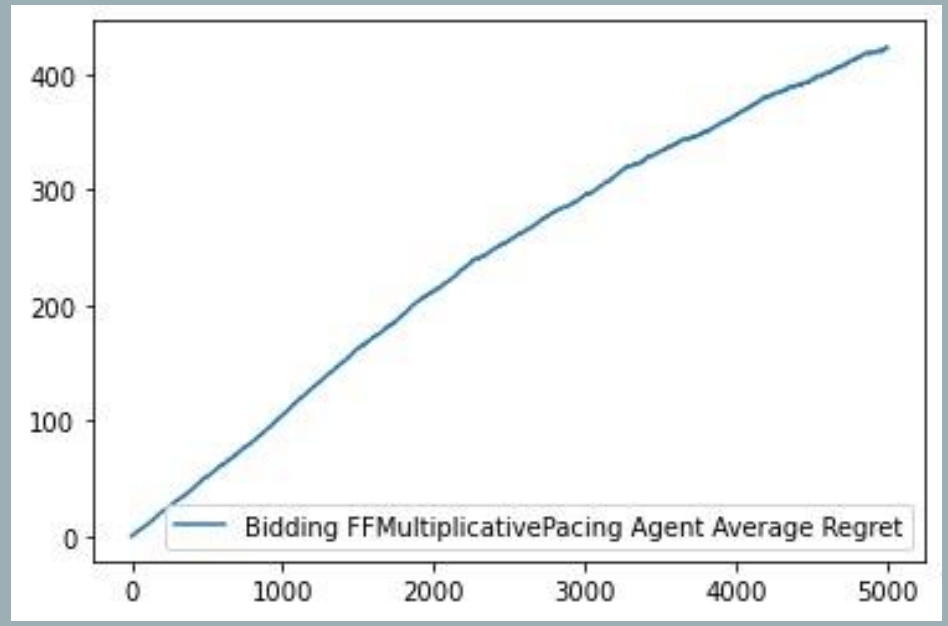
PRICING ONLY

NUMBER OF INTERACTIONS	500
NUMBER OF CUSTOMERS	50
PRODUCT COST	0.1
NUMBER OF TRIALS	10

EXP3

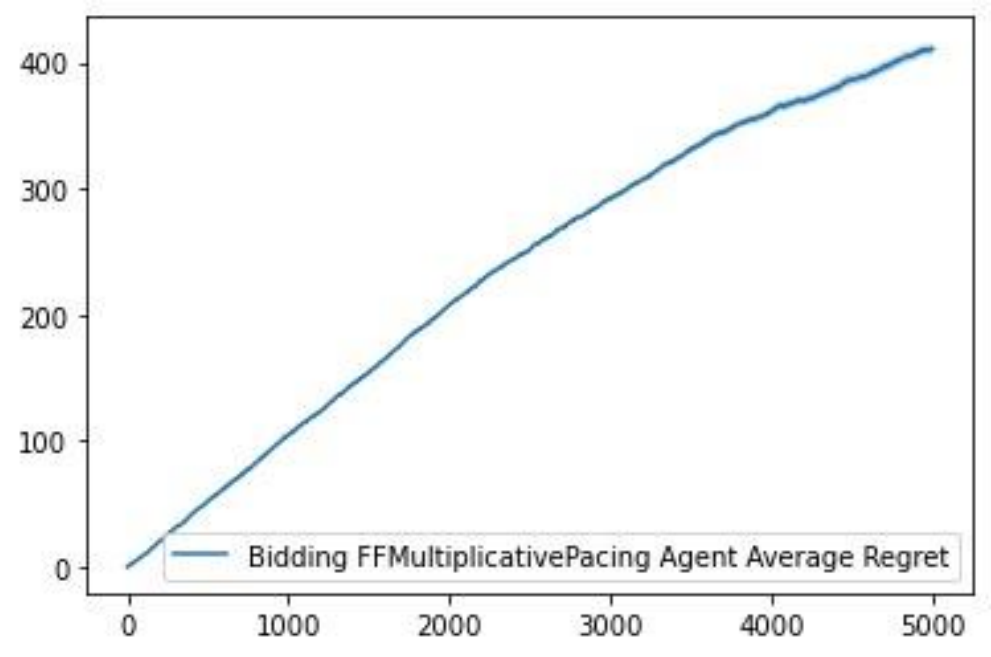
BIDDING ONLY

NUMBER OF INTERACTIONS	5000
BUDGET	1000
VALUATION	1
CTRs	[0.8 0.5 0.9 1]
LAMBDAs	[1 0.9]
NUMBER OF TRIALS	10



FF Multiplicative Pacing

PRICING AND BIDDING



REQUIREMENT 3

- Focus on the pricing problem
- Non-stationary environment
  - Abrupt changes
  - Noisy demand curve changes each interval

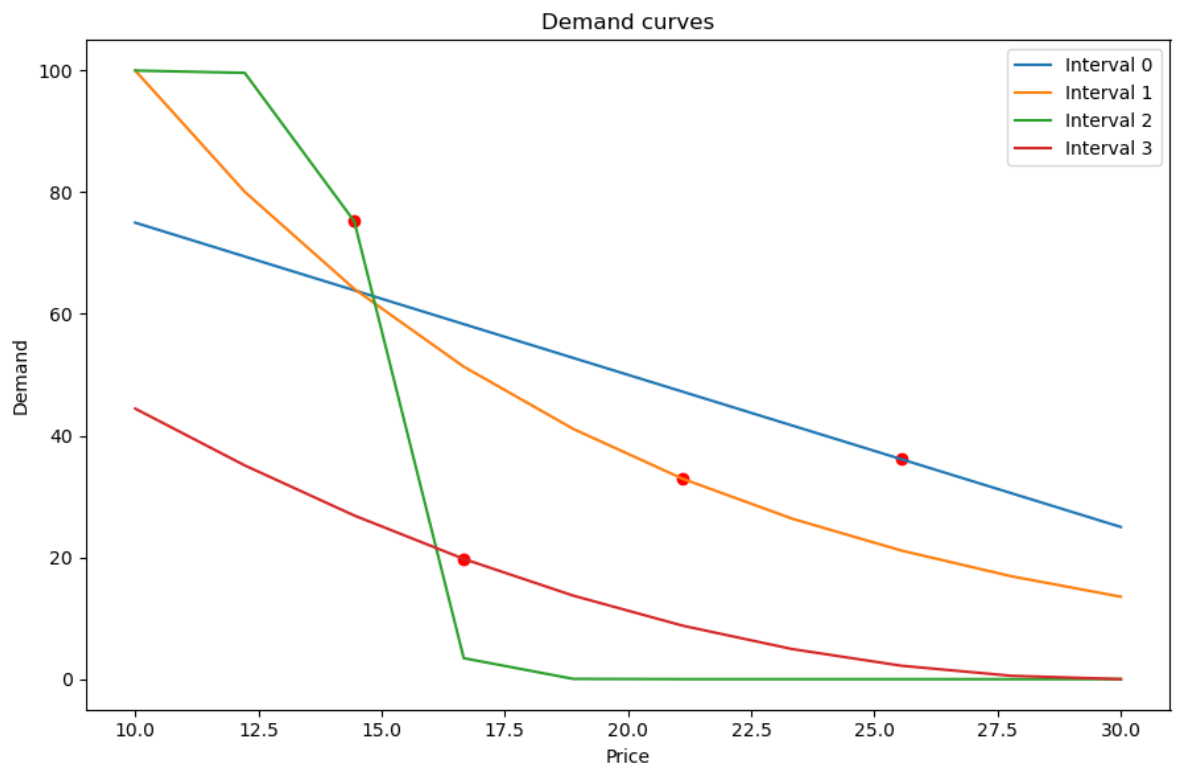
GOAL

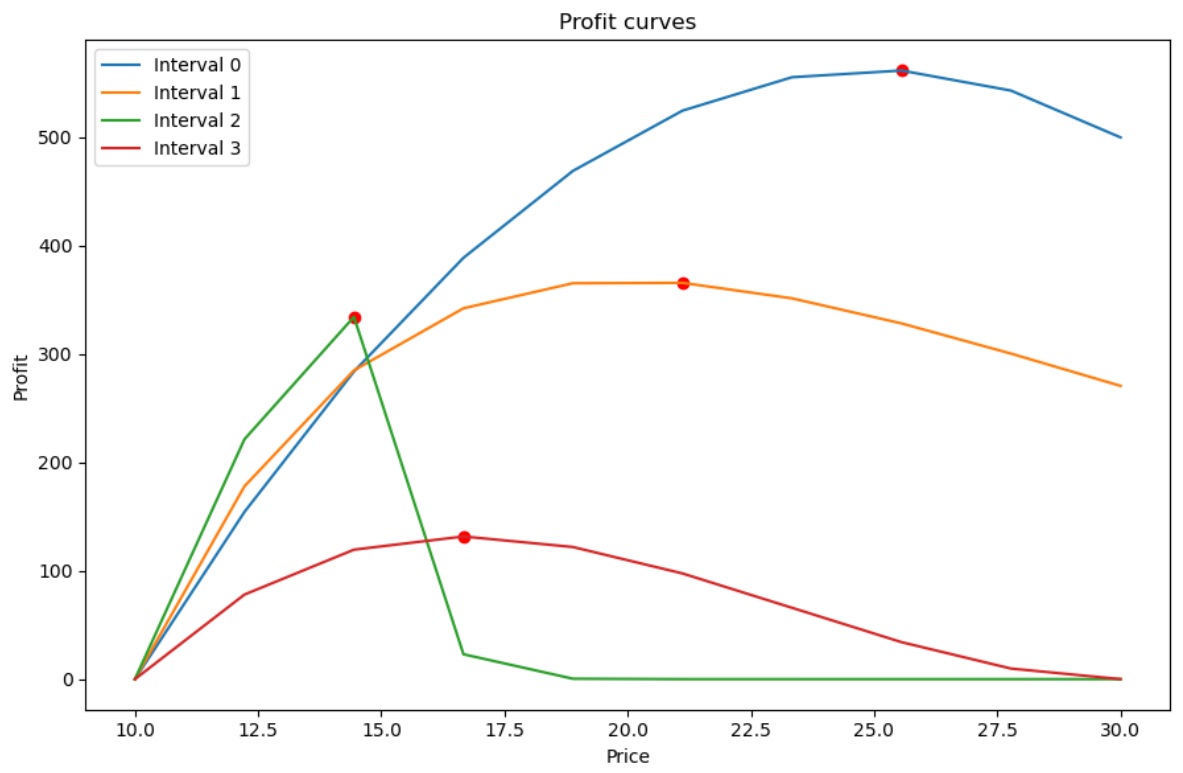
Pricing strategy for discretized set of prices $p \in [0,1]$ using:

- Sliding Window
- CUSUM

SETUP

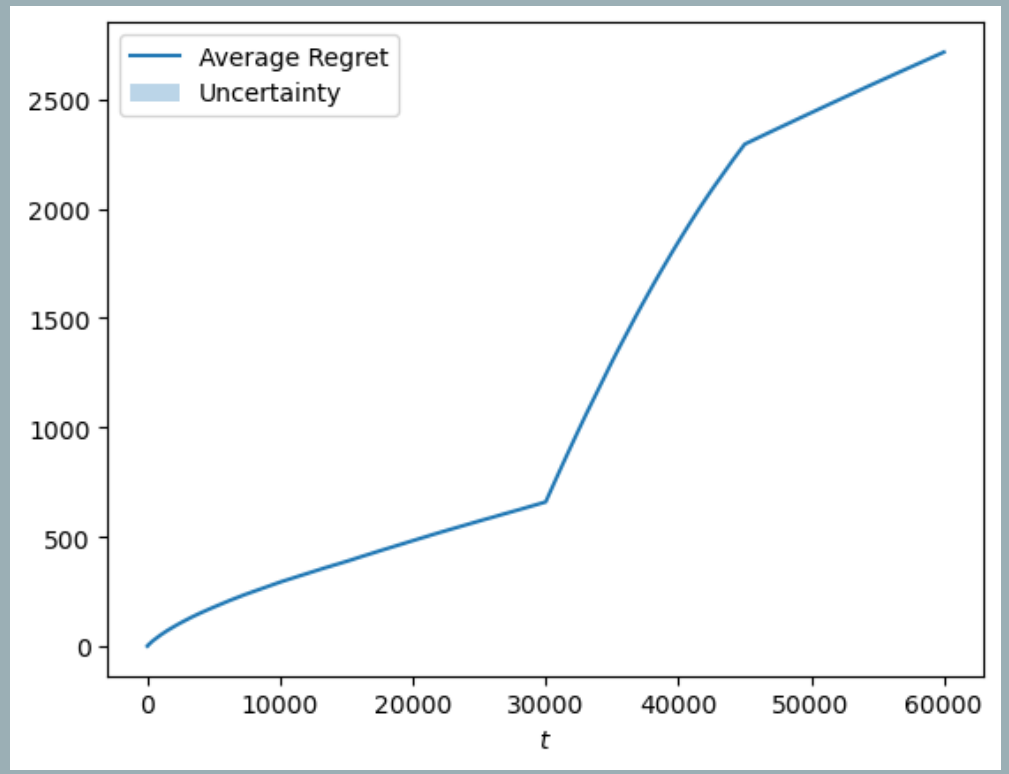
- Defined demand curves for each interval
- Derived respective profit curves

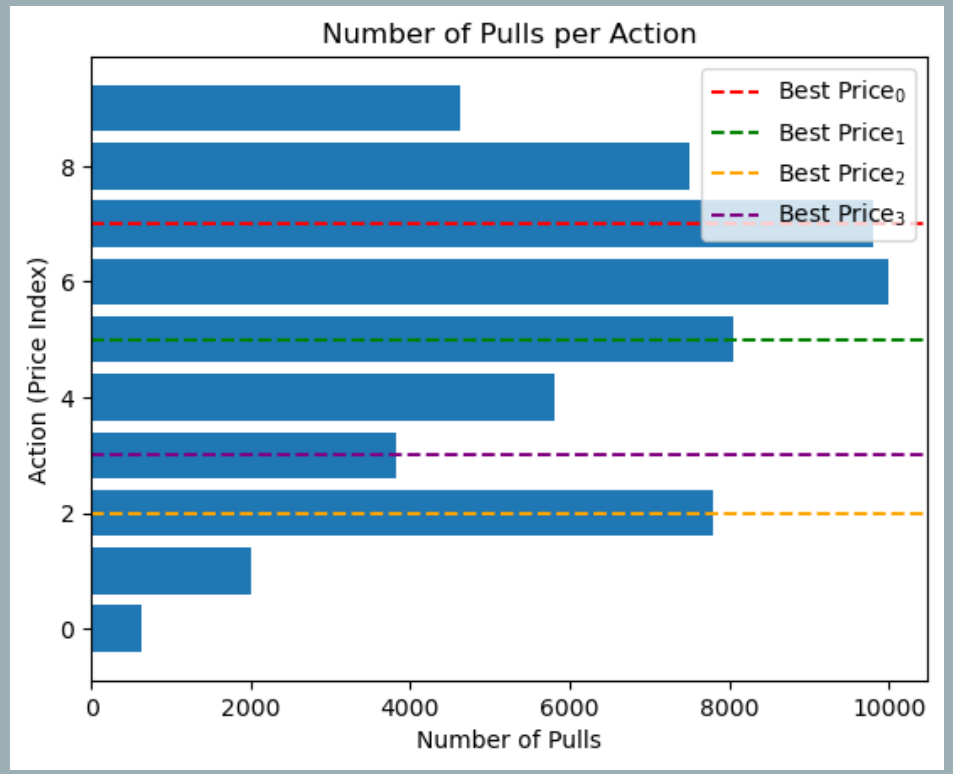




NON-STATIONARITY CHECK

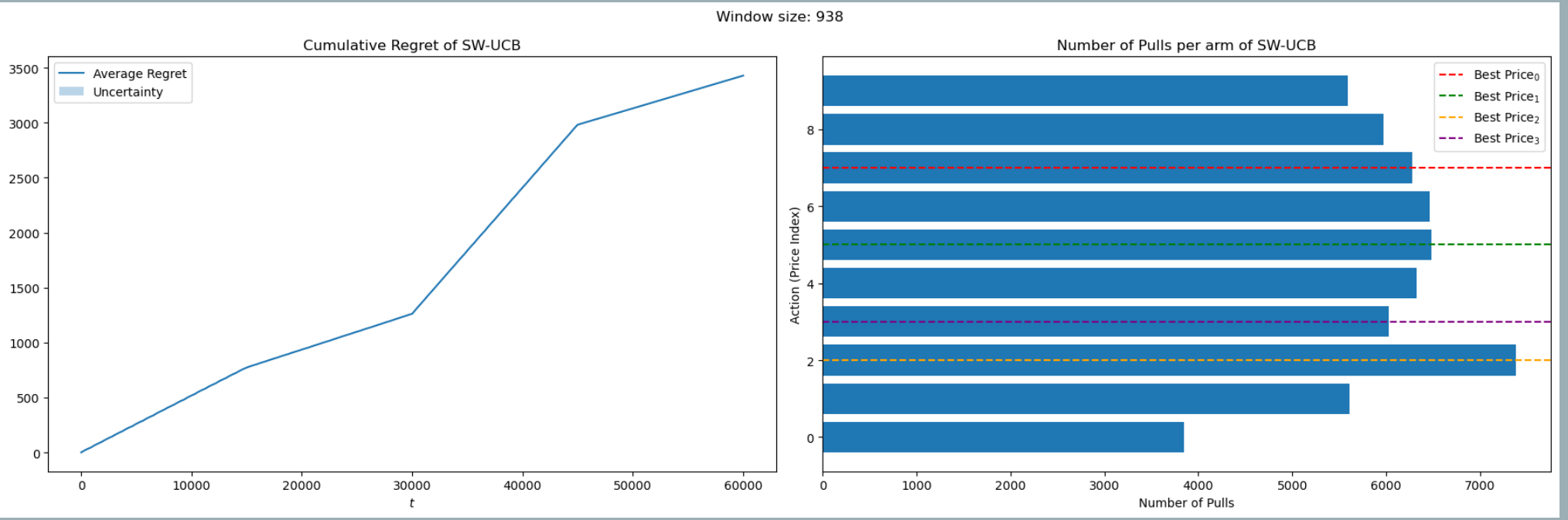
- Tested UCB1 on the environment
- Results indicate presence of non-stationarity

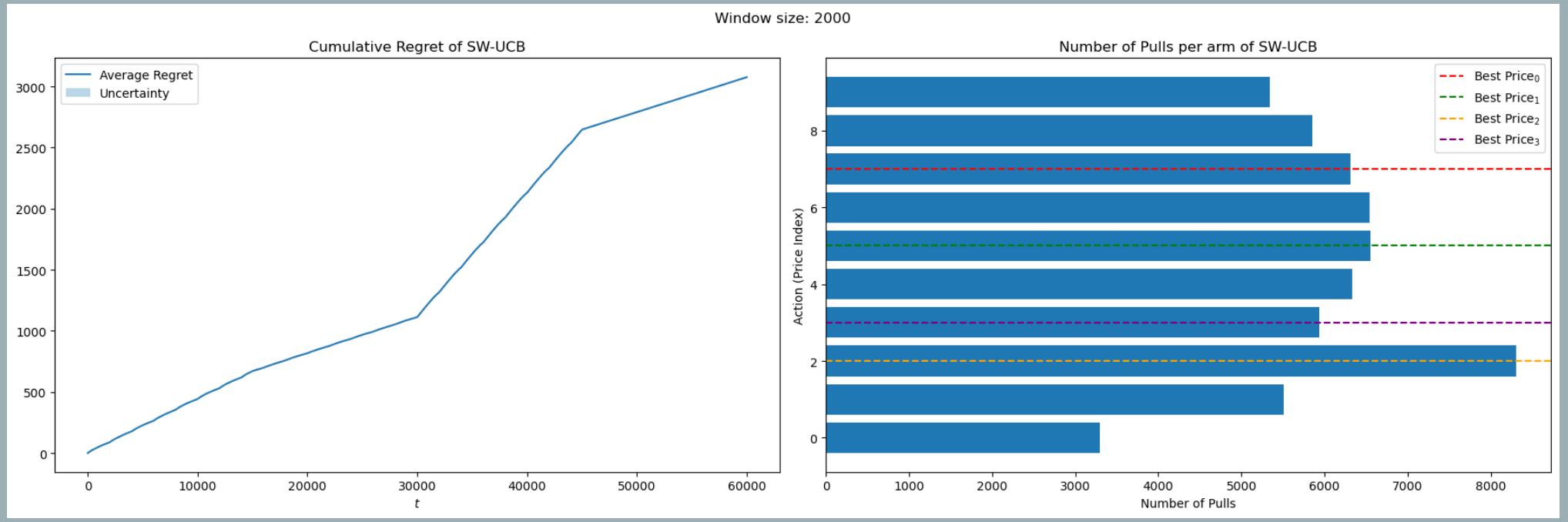


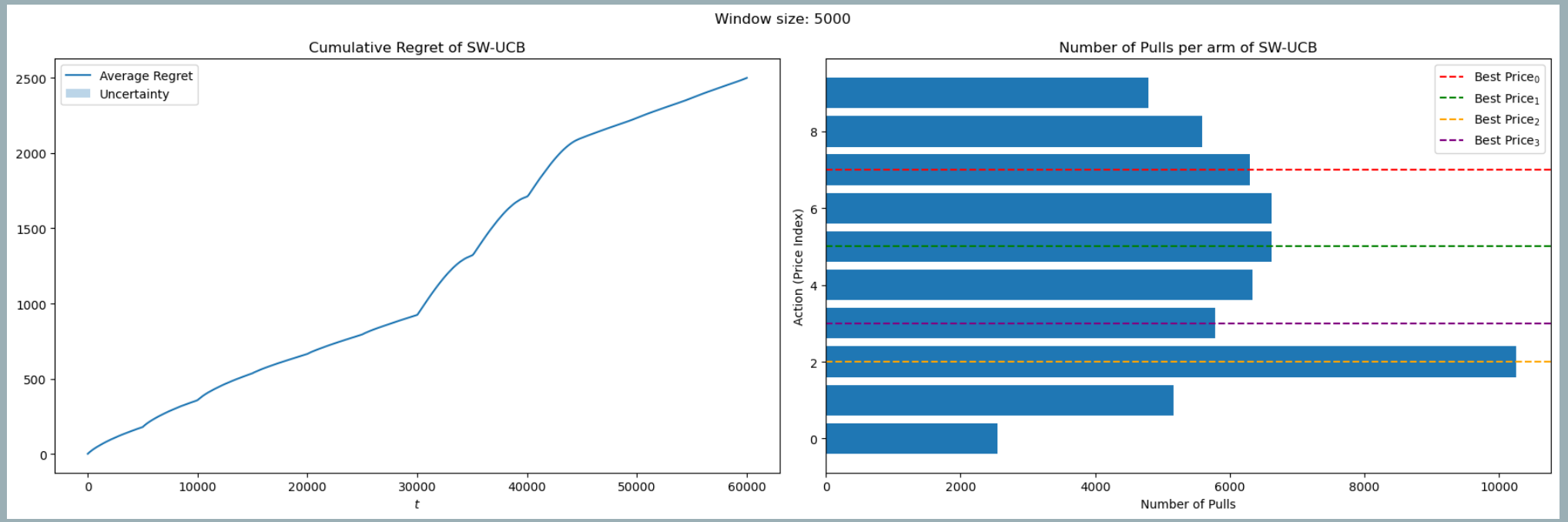


SLIDING WINDOW

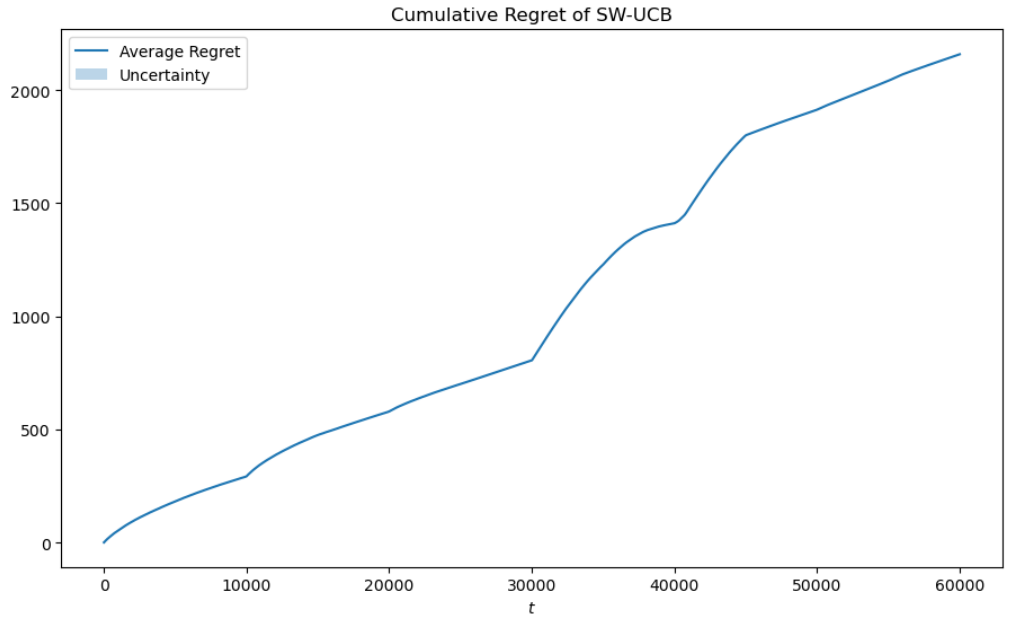
- Sliding Window UCB
- Initial trial failed
  - Decided to test different window sizes

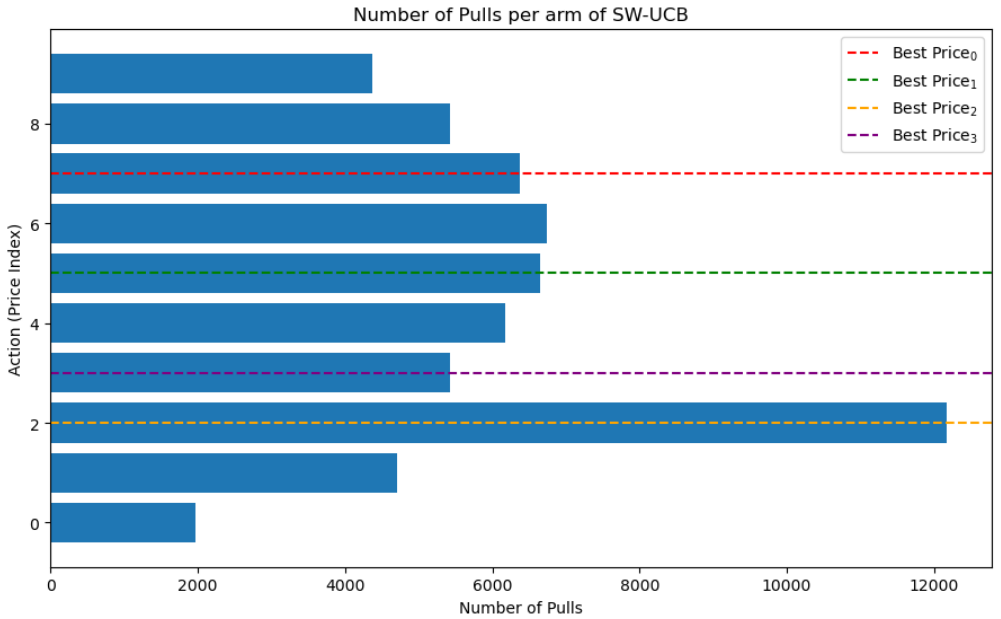



Window size: 2000
Cumulative Regret of SW-UCB
Average Regret
Uncertainty
t
Number of Pulls per arm of SW-UCB
Best Price0
Best Price1
Best Price2
Best Price3
Action (Price Index)
Number of Pulls

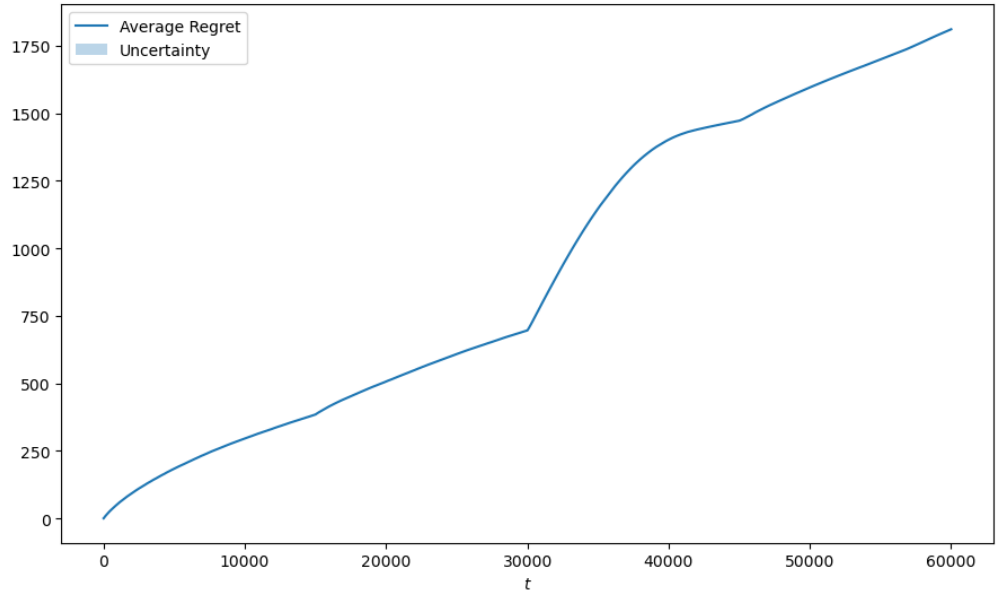

Window size: 5000
Cumulative Regret of SW-UCB
Average Regret
Uncertainty
t
Number of Pulls per arm of SW-UCB
Best Price0
Best Price1
Best Price2
Best Price3
Action (Price Index)
Number of Pulls

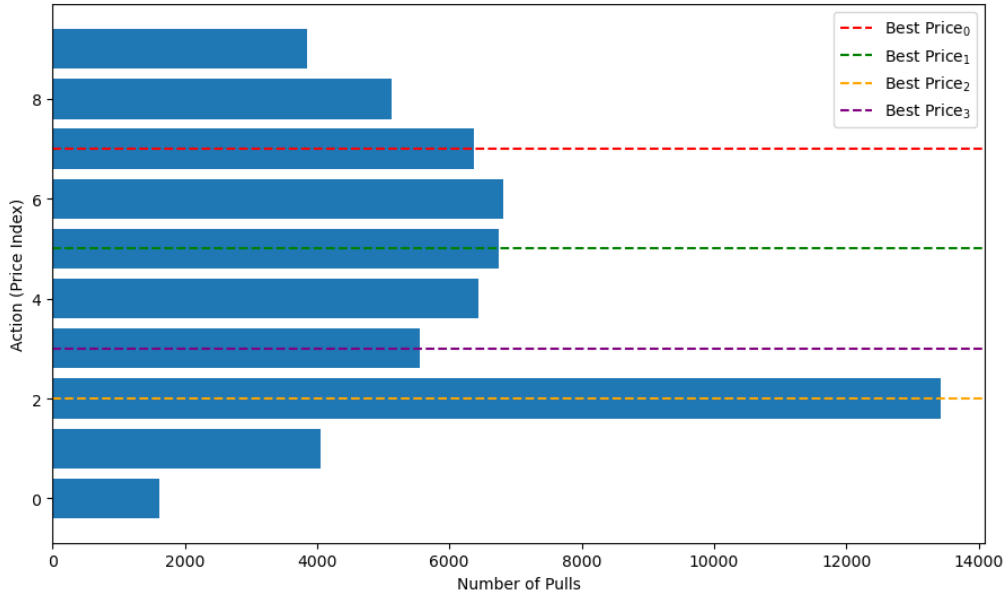
Window size: 10000





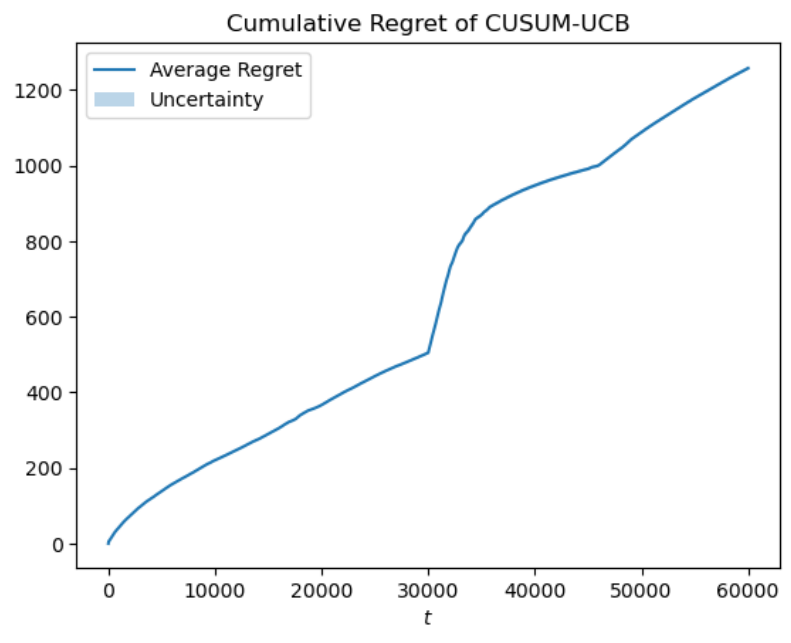
Window size: 15000

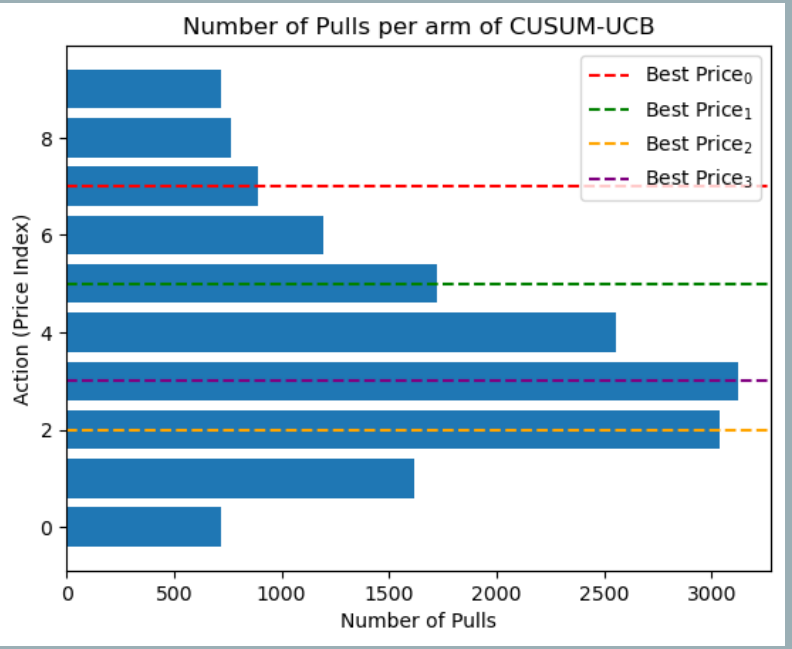


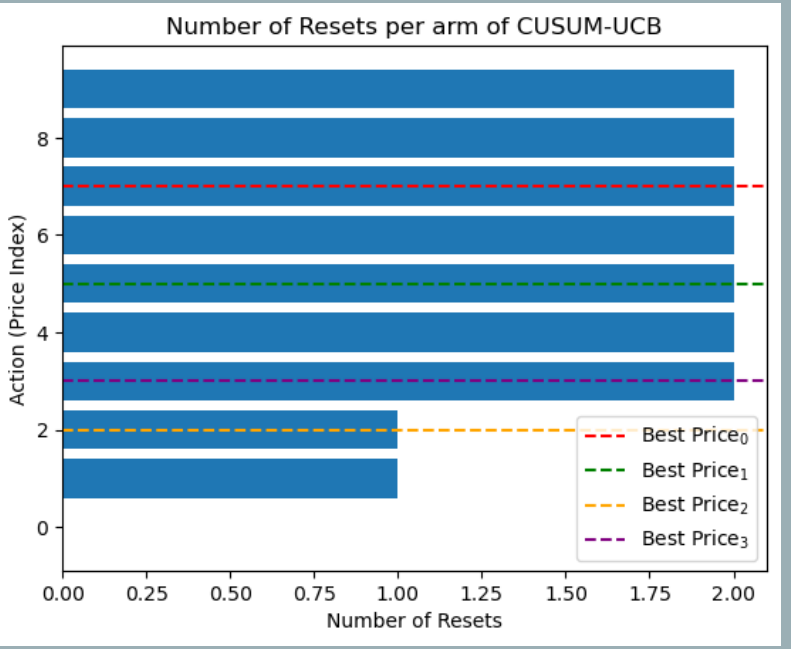


CHANGE DETECTION

- CUSUM UCB
- Best performing method overall



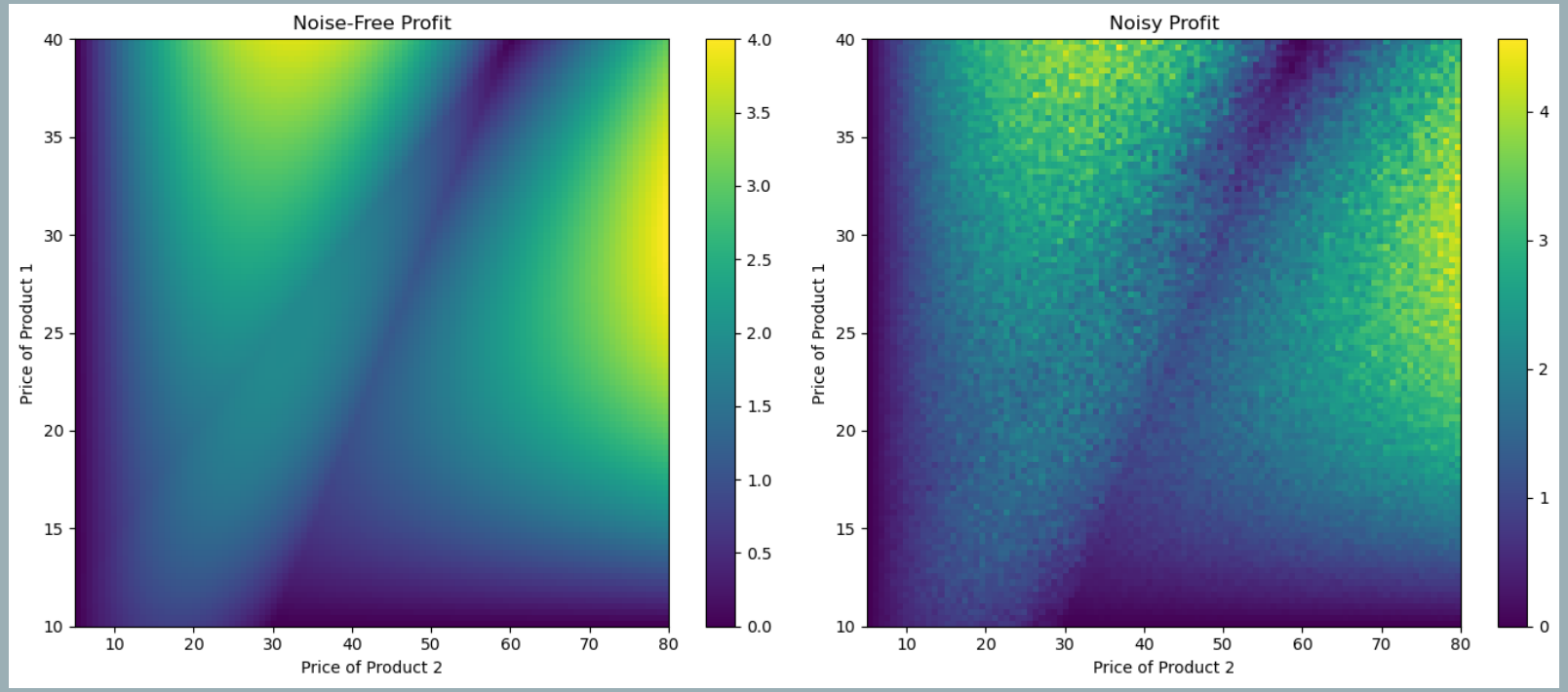




BONUS POINT

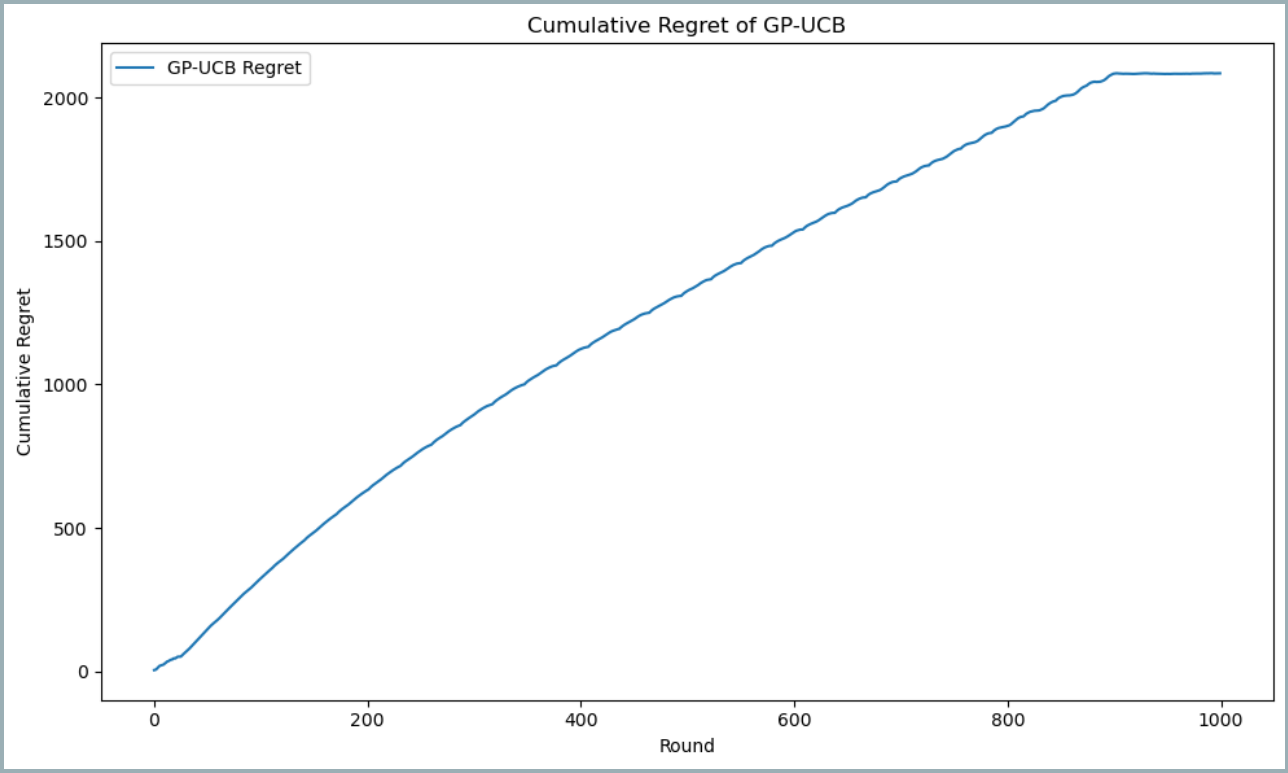
- Two-item stochastic pricing environment
- Noisy demand curve $D(p_1, p_2) + \eta$

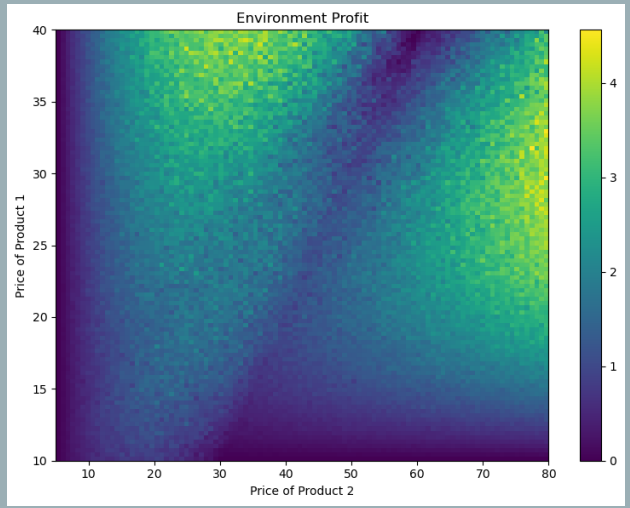
GOAL Build a regret minimizer for the continuous action set $[0,1]^2$ using 2D Gaussian Processes

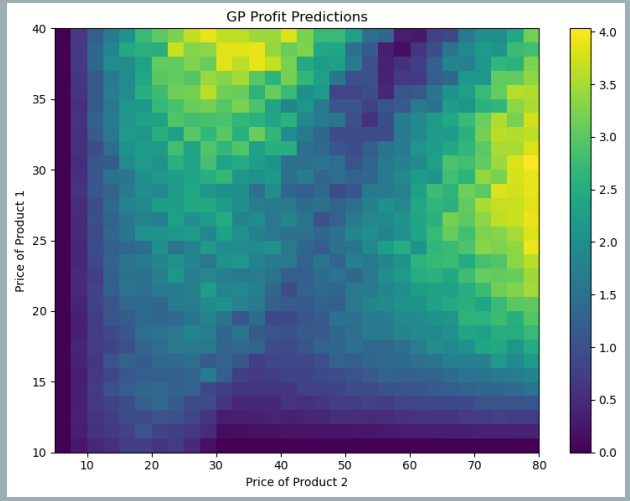


RESULTS

- GP UCB
- Great results obtained







REQUIREMENT 4

BIDDING ENVIRONMENT	Adversarial with full feedback
AUCTION TYPE	Generalized First-Price (Non-truthful)
BIDDING AGENTS	• Primal-Dual for Truthful (Multiplicative Pacing) • Primal-Dual for Non-truthful (Multiplicative Pacing with Hedge) • UCB-like (UCB1) • UCB-like Updating ρ (our version of the algorithm)
GOAL	Compare the performances of different algorithms under various setups

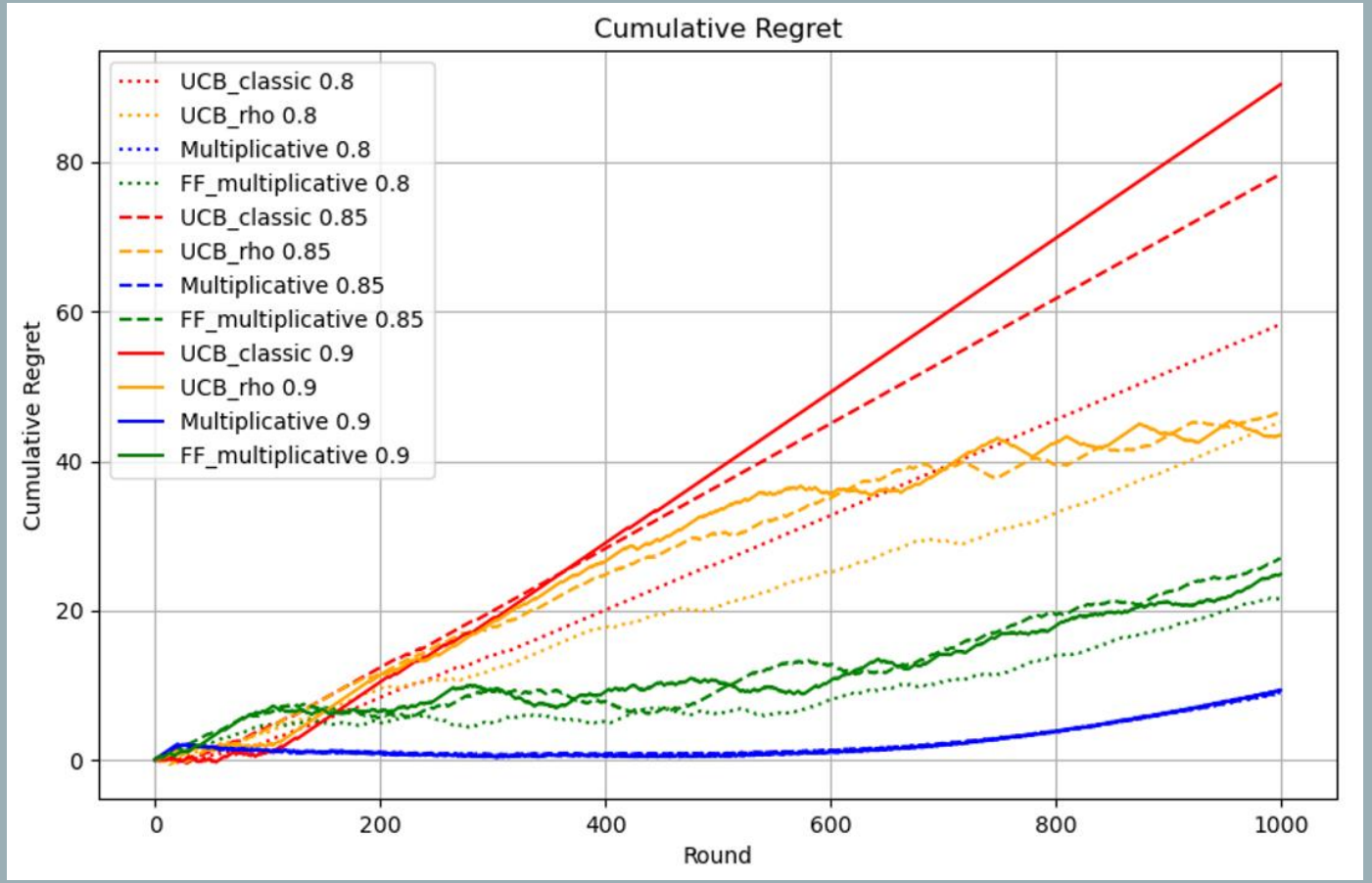
STANDARD CASE

SETUP

BIDDERS PER TYPE	3
VALUATIONS	[0.80 0.85 0.90]
NUMBER OF SLOTS	3
SLOT PROMINENCE	[0.8 0.9 1.0]
BUDGET	250

OBSERVATION

- UCB_classic has linear regret
- Multiplicative pacing has the smoothest curve



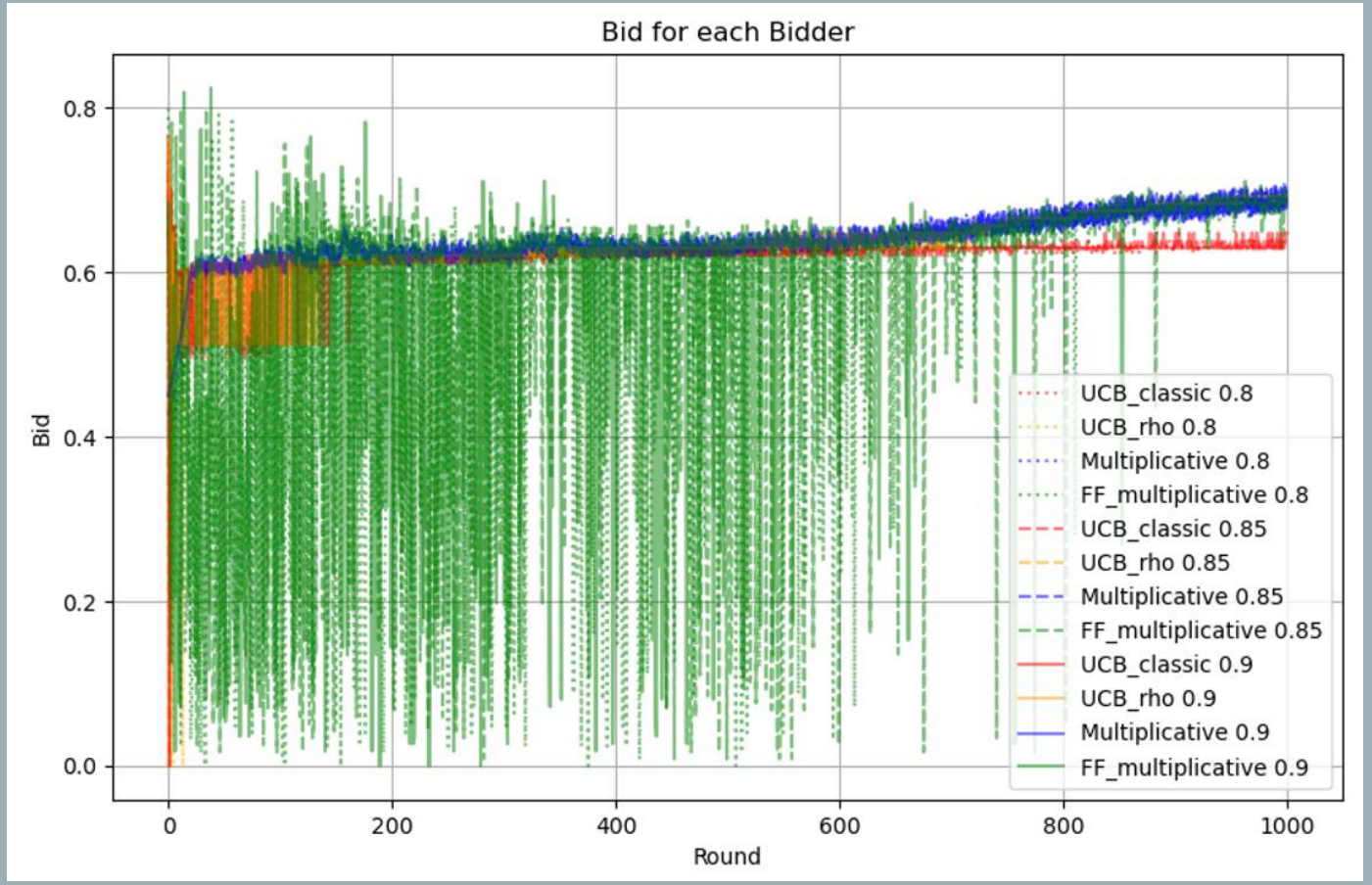
STANDARD CASE

SETUP

BIDDERS PER TYPE	3
VALUATIONS	[0.80 0.85 0.90]
NUMBER OF SLOTS	3
SLOT PROMINENCE	[0.8 0.9 1.0]
BUDGET	250

OBSERVATION

- Bids increase over rounds
- UCB_classic stops winning after round 500
- FF_multiplicative needs time to be consistent with bids



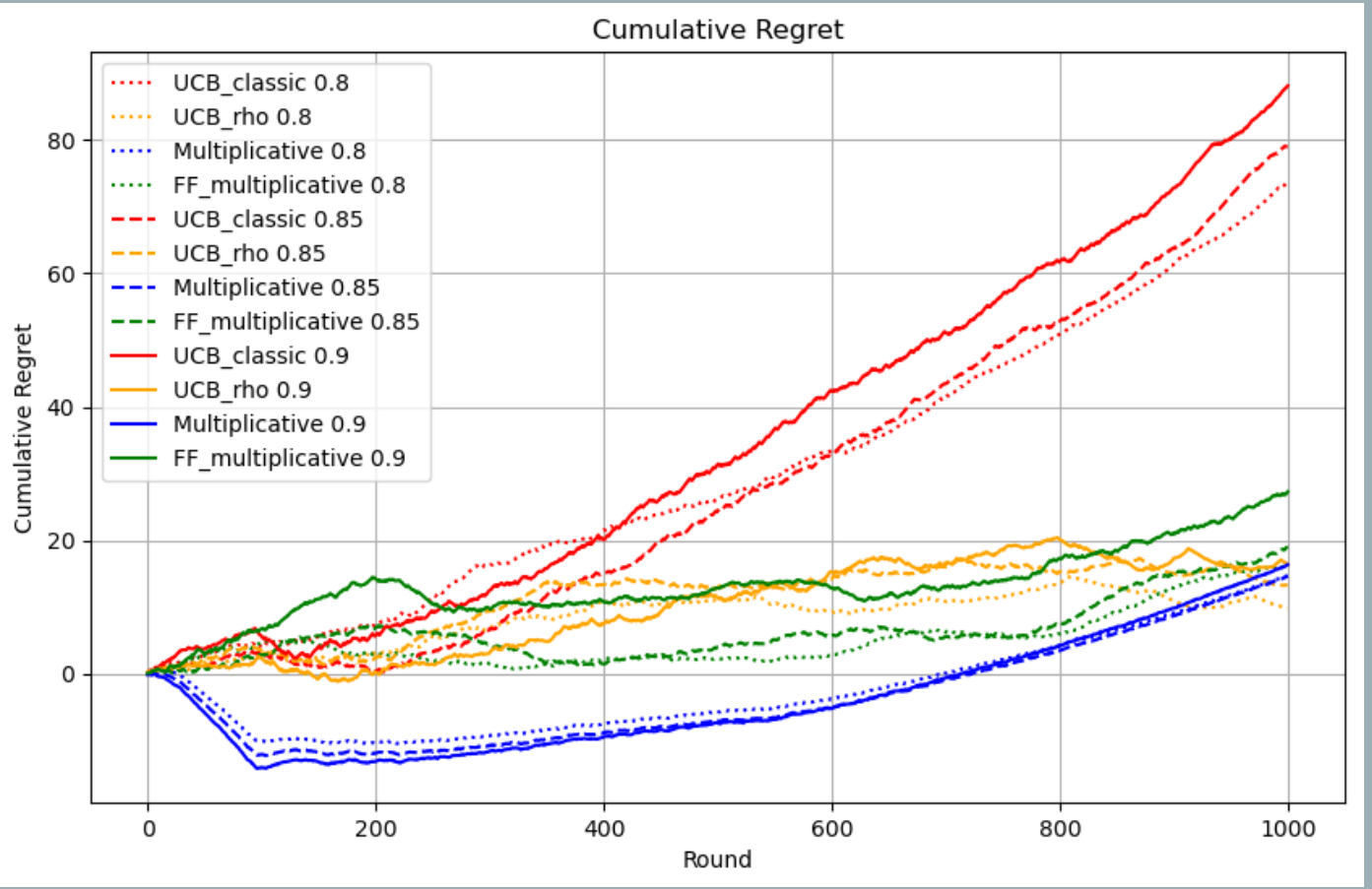
MANY SLOTS CASE

SETUP

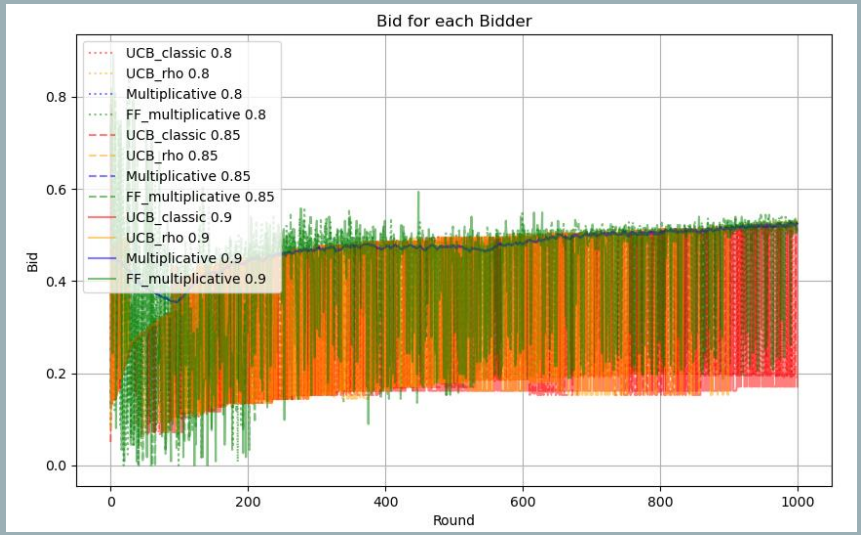
BIDDERS PER TYPE	3
VALUATIONS	[0.80 0.85 0.90]
NUMBER OF SLOTS	10
SLOT PROMINENCE	[0.1 … 1.0]
BUDGET	250

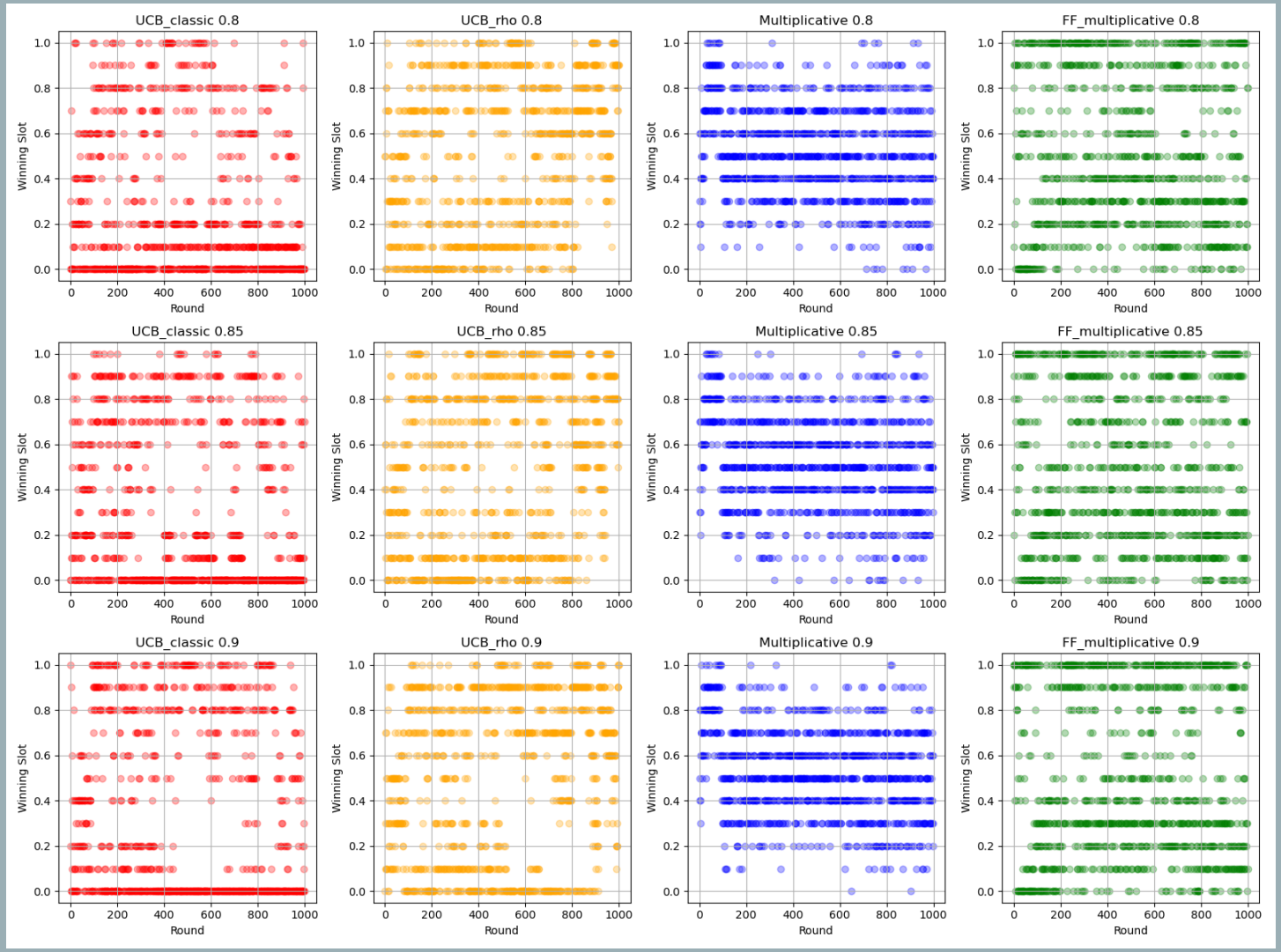
OBSERVATION

- All algorithms have more or less the same regret except for standard UCB_like algorithm



MANY SLOTS CASE





OBSERVATION

- Multiplicative algorithm wins middle slots
- FF_multiplicative prefers going all or nothing, winning the best slot more often

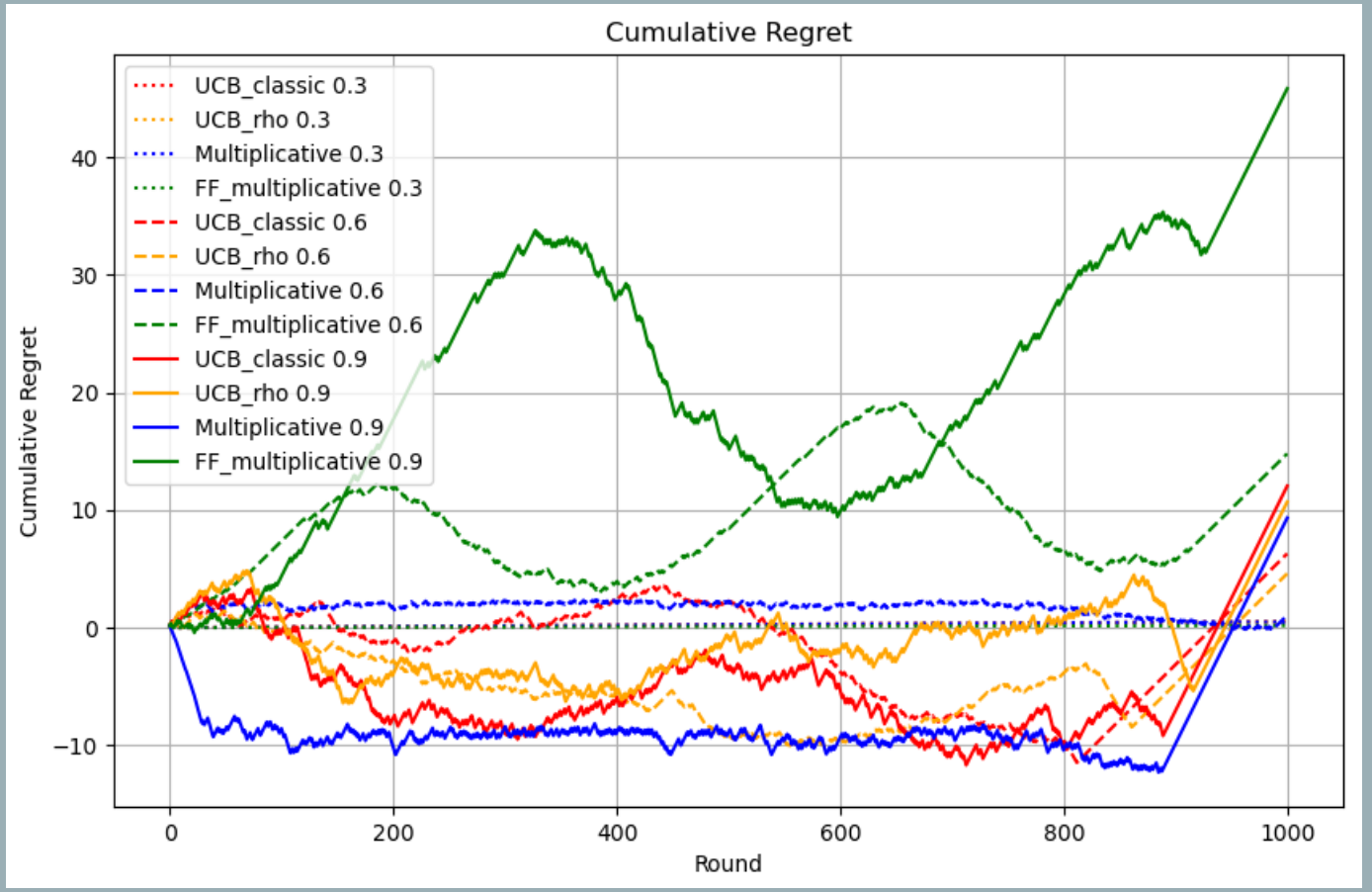
LOW BUDGET - DIFFERENT VALUATION

SETUP

BIDDERS PER TYPE	3
VALUATIONS	[0.3 0.6 0.9]
NUMBER OF SLOTS	3
SLOT PROMINENCE	[0.8 0.9 1.0]
BUDGET	100

OBSERVATION

- UCB_like performs better than FF_multiplicative
- Our modification to the UCB_like algorithm performs slightly better than the original
- Many agents deplete the budget about 100 rounds before the end



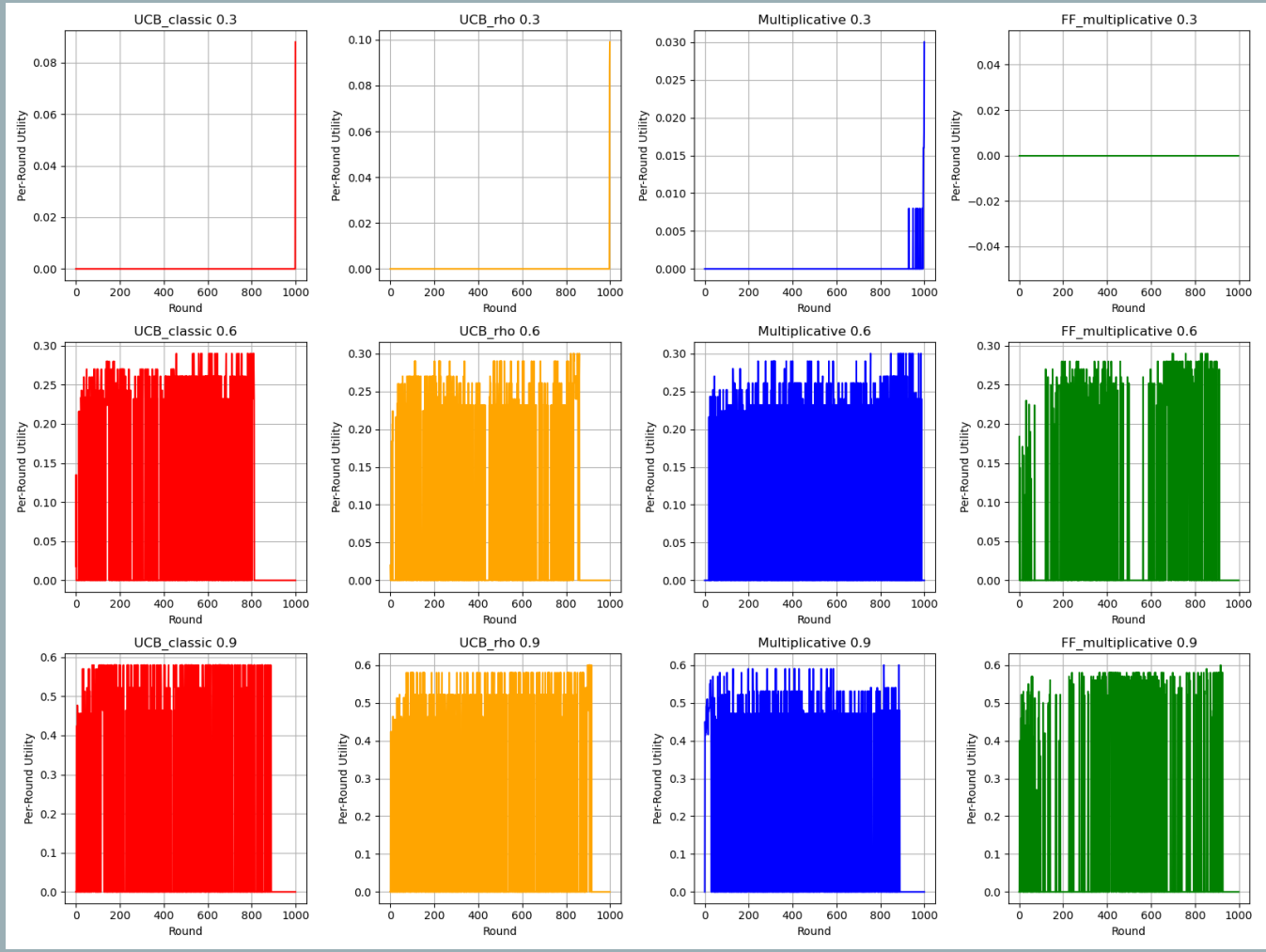
LOW BUDGET - DIFFERENT VALUATION

SETUP

BIDDERS PER TYPE	3
VALUATIONS	[0.3 0.6 0.9]
NUMBER OF SLOTS	3
SLOT PROMINENCE	[0.8 0.9 1.0]
BUDGET	100

OBSERVATION

- Bidder with valuation of 0.3 have regret 0 since minimum bid for winning a slot is 0.4
- Unclear why FF_multiplicative performs badly



REQUIREMENT 4

FINAL CONSIDERATIONS

- Best performing algorithm is:
  - Multiplicative, despite
  - FF_multiplicative having better theoretical guarantees

- Our version of UCB-like algorithm for expert feedback worked well:
  - Almost always much better than UCB1
  - Slightly better in a single case

THANK YOU

Source code available at this GitHub repository