Learning to Bid in Discriminatory Auctions with Budget Constraints

Negin Golrezaei and **Sourav Sahoo**

MIT

AISTATS 2026

Model

Model

- Discriminatory (pay-as-bid or PAB) auctions are used in Treasury auctions and electricity markets.

Model

- Discriminatory (pay-as-bid or PAB) auctions are used in Treasury auctions and electricity markets.
- **Auction Format:**
  - K identical units of a single item.
  - Bidders can demand for multiple units.
  - **Allocation Rule:** Bidders with top K bids ‘win’ the auction.
  - **Payment Rule:** Each bidder pays **their own bid** for each unit they win.

Model

- Discriminatory (pay-as-bid or PAB) auctions are used in Treasury auctions and electricity markets.
- **Auction Format:**
  - K identical units of a single item.
  - Bidders can demand for multiple units.
  - **Allocation Rule:** Bidders with top K bids 'win' the auction.
  - **Payment Rule:** Each bidder pays **their own bid** for each unit they win.
- Buyers are **quasilinear utility maximizers**.

Utility = Value - α.Payment, where $\alpha \in [0,1]$ is the cost of capital.

Model

- Discriminatory (pay-as-bid or PAB) auctions are used in Treasury auctions and electricity markets.
- **Auction Format:**
  - K identical units of a single item.
  - Bidders can demand for multiple units.
  - **Allocation Rule:** Bidders with top K bids 'win' the auction.
  - **Payment Rule:** Each bidder pays **their own bid** for each unit they win.
- Buyers are **quasilinear utility maximizers**.

Utility = Value - α.Payment, where $\alpha \in [0,1]$ is the cost of capital.

If $K = 1$, PAB auctions = First-price auctions.

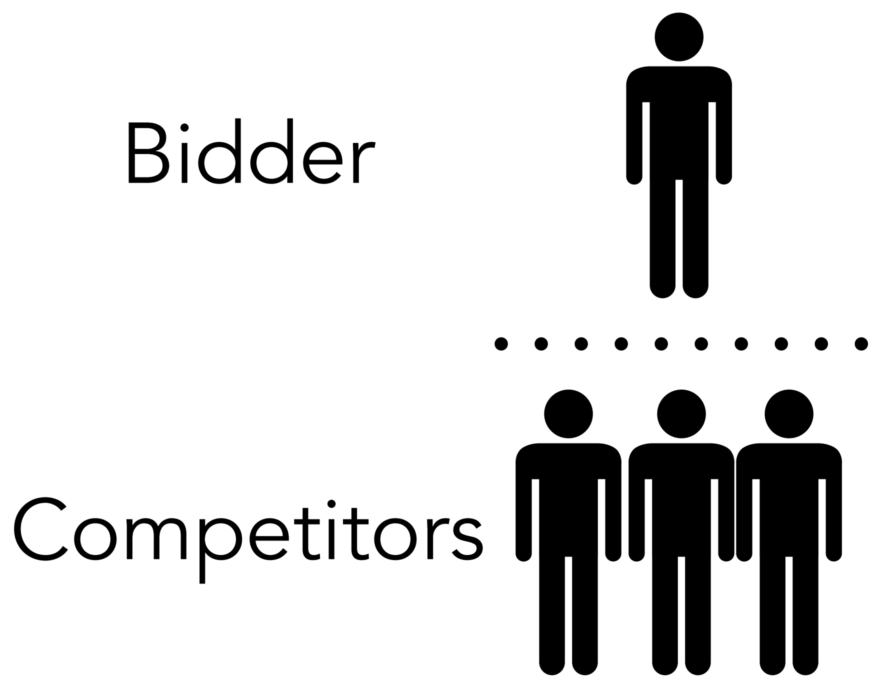
Problem Setting

Problem Setting

- Repeated setting over T rounds.
- The bidder has a total budget $B = \rho KT$. Once budget is exhausted, the game stops.
- In each round $t \in [T]$, the bidder observes $\mathbf{v}_t$. Assume $\mathbf{v}_t \in [0,1]^K$.

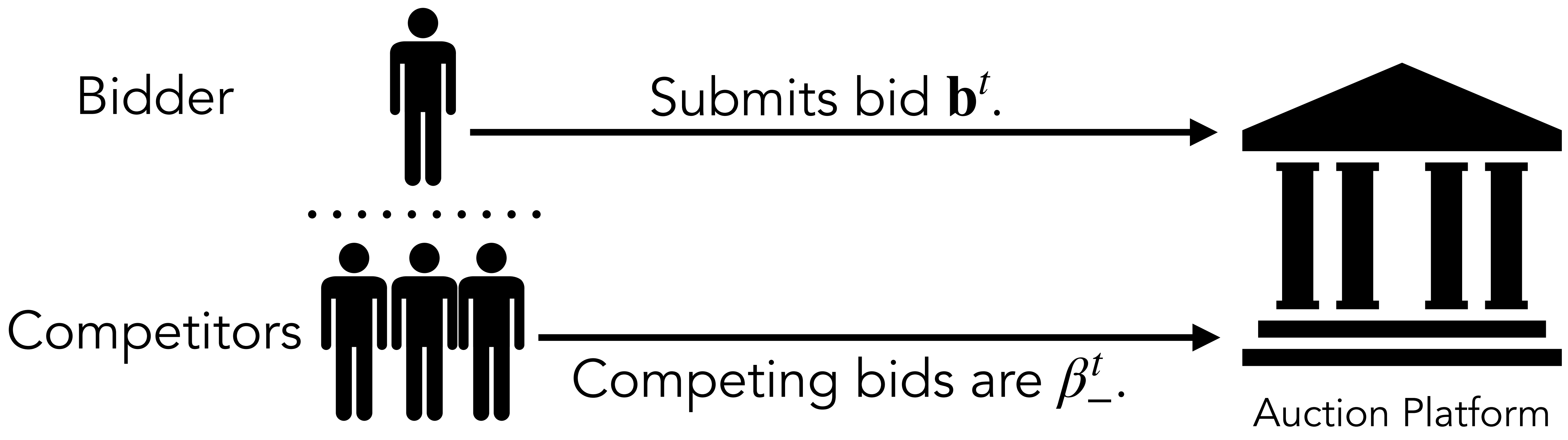
Problem Setting

- Repeated setting over T rounds.
- The bidder has a total budget $B = \rho K T$. Once budget is exhausted, the game stops.
- In each round $t \in [T]$, the bidder observes $\mathbf{v}_t$. Assume $\mathbf{v}_t \in [0,1]^K$.



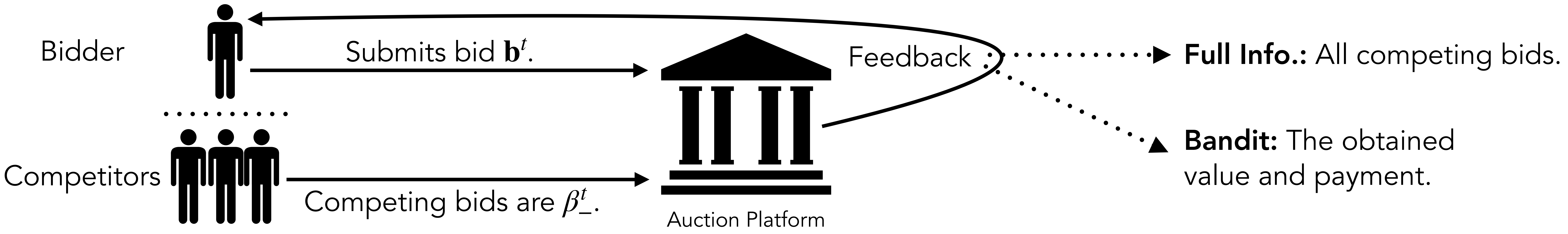
Problem Setting

- Repeated setting over T rounds.
- The bidder has a total budget $B = \rho KT$. Once budget is exhausted, the game stops.
- In each round $t \in [T]$, the bidder observes $\mathbf{v}_t$. Assume $\mathbf{v}_t \in [0,1]^K$.



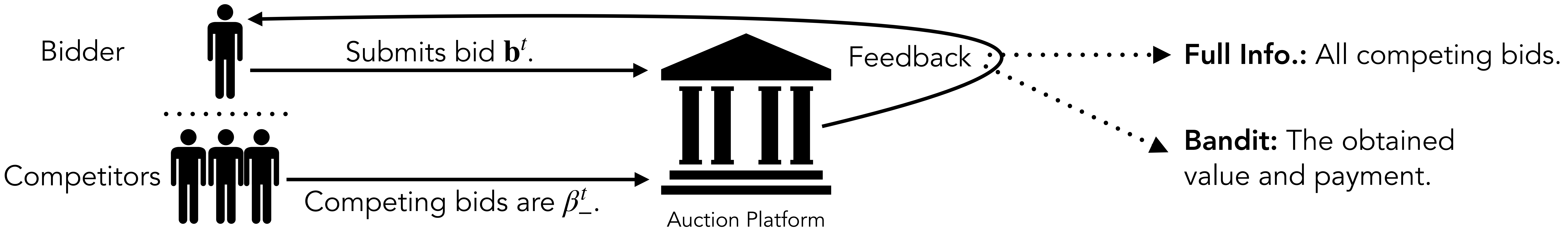
Problem Setting

- Repeated setting over T rounds.
- The bidder has a total budget $B = \rho KT$. Once budget is exhausted, the game stops.
- In each round $t \in [T]$, the bidder observes $\mathbf{v}_t$. Assume $\mathbf{v}_t \in [0,1]^K$.



Problem Setting

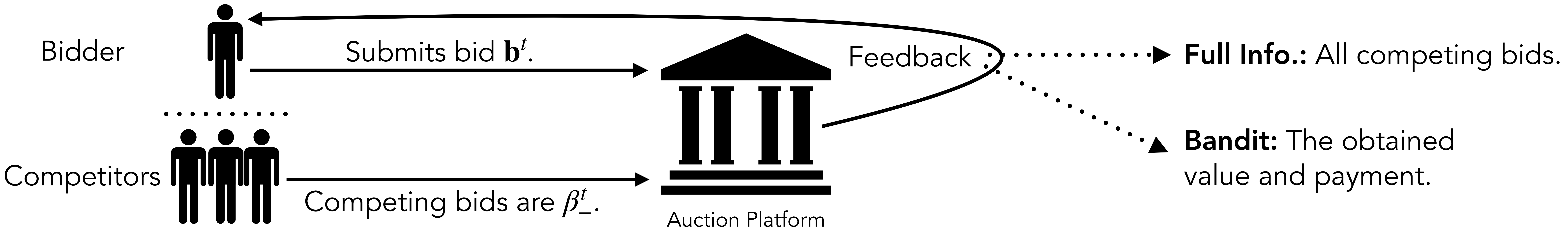
- Repeated setting over T rounds.
- The bidder has a total budget $B = \rho KT$. Once budget is exhausted, the game stops.
- In each round $t \in [T]$, the bidder observes $\mathbf{v}_t$. Assume $\mathbf{v}_t \in [0,1]^K$.



- Bidder follows **no-overbidding (NOB)** strategies to ensure that Utility$_t(\mathbf{v}_t) \geq 0$.

Problem Setting

- Repeated setting over T rounds.
- The bidder has a total budget $B = \rho KT$. Once budget is exhausted, the game stops.
- In each round $t \in [T]$, the bidder observes $\mathbf{v}_t$. Assume $\mathbf{v}_t \in [0,1]^K$.

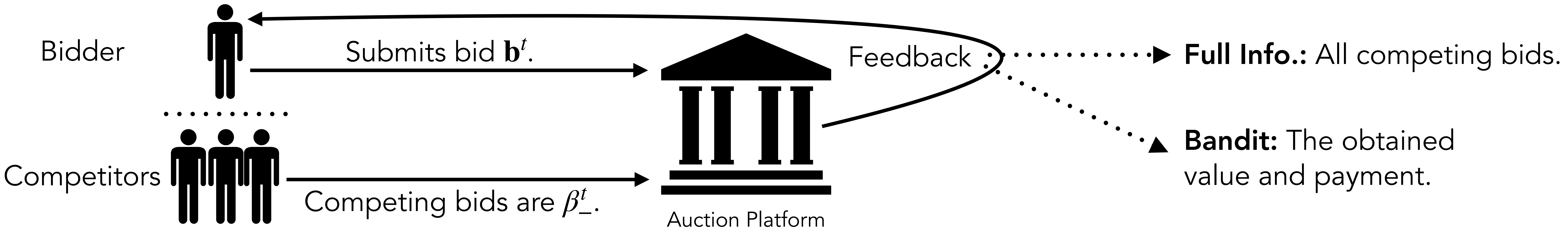


- Bidder follows **no-overbidding (NOB)** strategies to ensure that Utility$_t(\mathbf{v}_t) \geq 0$.

Performance Metric:

Problem Setting

- Repeated setting over T rounds.
- The bidder has a total budget $B = \rho KT$. Once budget is exhausted, the game stops.
- In each round $t \in [T]$, the bidder observes $\mathbf{v}_t$. Assume $\mathbf{v}_t \in [0,1]^K$.



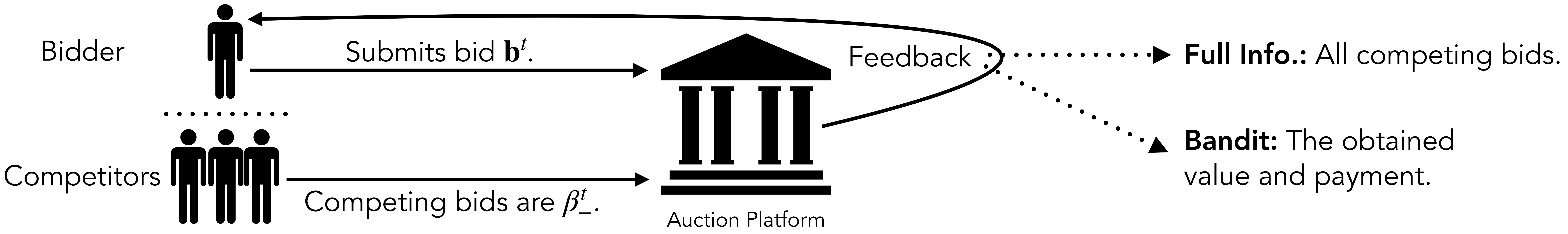
- Bidder follows **no-overbidding (NOB)** strategies to ensure that Utility$_t(\mathbf{v}_t) \geq 0$.

Performance Metric:

- OPT = expected utility by the hindsight optimal policy under budget constraint ($\rho < 1$).
- OPT$_{nb}$= the no budget constraint analogue ($\rho \geq 1$).

Problem Setting

- Repeated setting over T rounds.
- The bidder has a total budget $B = \rho KT$. Once budget is exhausted, the game stops.
- In each round $t \in [T]$, the bidder observes $\mathbf{v}_t$. Assume $\mathbf{v}_t \in [0,1]^K$.



  - Bidder follows **no-overbidding (NOB)** strategies to ensure that $\text{Utility}_t(\mathbf{v}_t) \geq 0$.

Performance Metric:

- OPT = expected utility by the hindsight optimal policy under budget constraint ($\rho < 1$).
- OPT_{nb}= the no budget constraint analogue ($\rho \geq 1$).

$$\text{REG} = \text{OPT}_{nb} - \sum_{t=1}^{T} \mathbb{E}[\text{Utility}_t(\mathbf{v}_t)] \quad \Big| \quad \rho \cdot \text{REG} = \rho \cdot \text{OPT} - \sum_{t=1}^{T} \mathbb{E}[\text{Utility}_t(\mathbf{v}_t)]$$

Key Challenges and Results Summary

Key Challenges and Results Summary

context = *valuation*

- The number of NOB strategies is **exponential in** K.
- In bandit setting, a naive implementation incurs an **additional** $O(\sqrt{\text{\# of contexts}})$ **factor** in the regret bound.
- Under budget constraints, bidders face the classical **save-or-spend dilemma**.

Key Challenges and Results Summary

context = valuation

- The number of NOB strategies is **exponential in K**.
- In bandit setting, a naive implementation incurs an **additional $O(\sqrt{\text{\# of contexts}})$ factor** in the regret bound.
- Under budget constraints, bidders face the classical **save-or-spend dilemma**.

- Reduce the learning problem to an **online shortest path problem over DAGs.** Results in a **polynomial time algorithm** that yields **sublinear regret.**
- Leverage **complete cross-learning** in the bandit setting: the utility of the chosen action (bid) for a given context reveals its utility for all other contexts.
- **Primal-dual framework** to handle budgets: dual-adjusted utility for primal algorithm updates, OGD to update the dual variable. Yields **ρ-approximate sublinear regret.**

Key Challenges and Results Summary

context = valuation

- The number of NOB strategies is **exponential in K**.
- In bandit setting, a naive implementation incurs an **additional $O(\sqrt{\text{\# of contexts}})$ factor** in the regret bound.
- Under budget constraints, bidders face the classical **save-or-spend dilemma**.

- Reduce the learning problem to an **online shortest path problem over DAGs.** Results in a **polynomial time algorithm** that yields **sublinear regret.**
- Leverage **complete cross-learning** in the bandit setting: the utility of the chosen action (bid) for a given context reveals its utility for all other contexts.
- **Primal-dual framework** to handle budgets: dual-adjusted utility for primal algorithm updates, OGD to update the dual variable. Yields **ρ-approximate sublinear regret.**

Thank You!