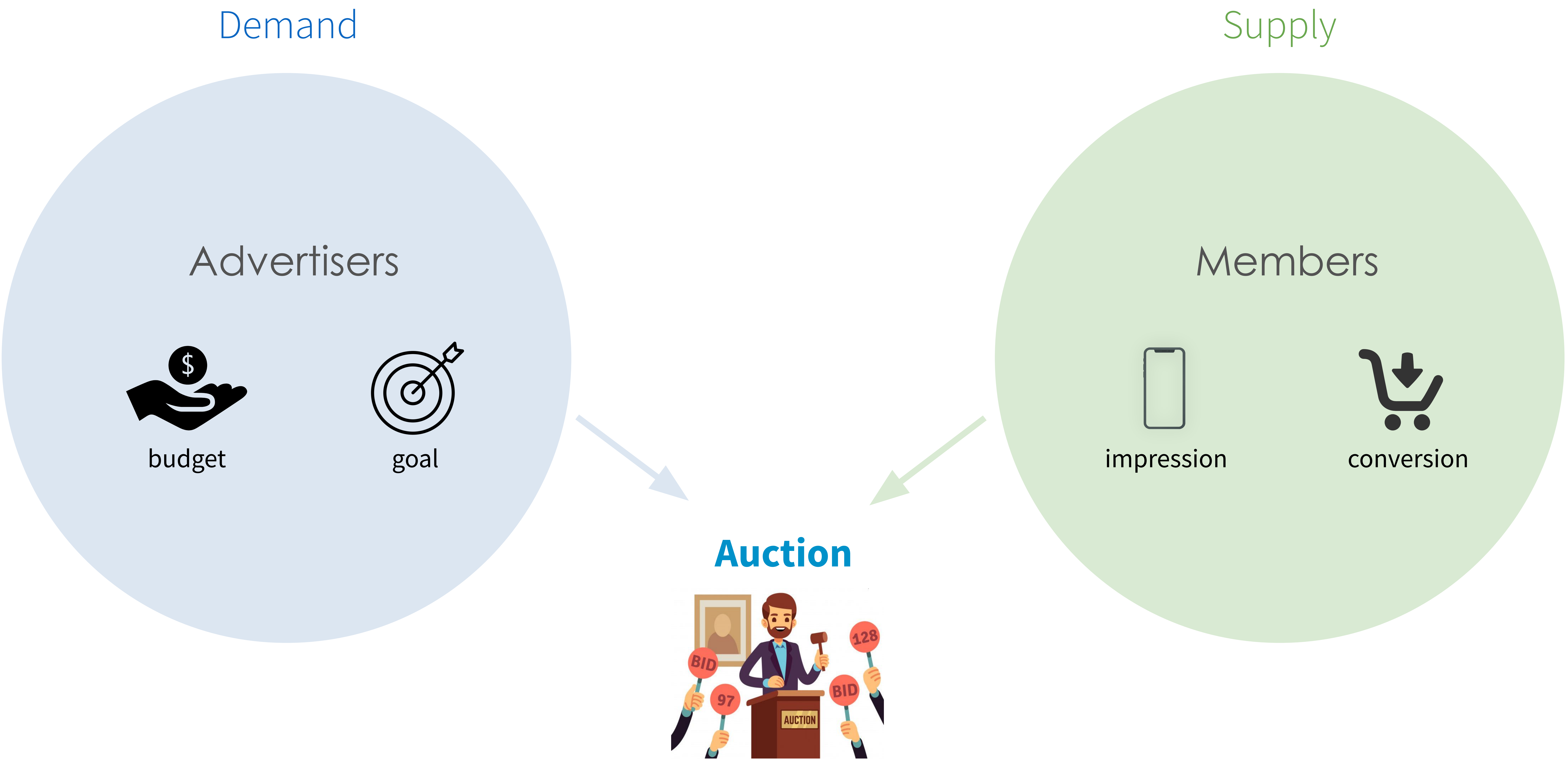
AdKDD 2022

Bidding Agent Design in the LinkedIn Ad Marketplace

Yuan Gao (joint work with Kaiyu Yang, Yuanlong Chen, Min Liu and Noureddine El Karoui)

LinkedIn

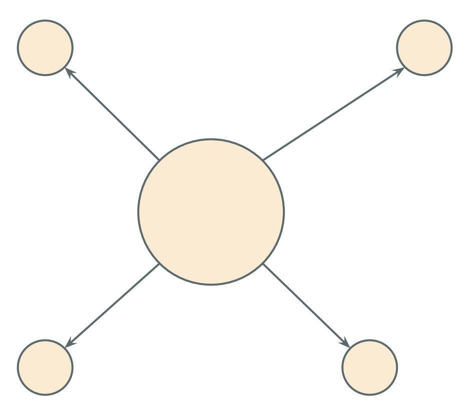
LinkedIn Ad Marketplace



Design Philosophy

Centralized v.s. Decentralized

Centralized



Formulate the Optimization Problem from Seller-defined Objective

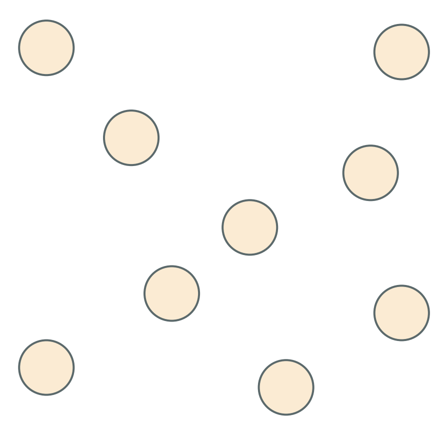
One Giant Optimization Problem

"Online Resource Allocations"

Mehta et al. (2007); Abrams et al. (2008); Feldman et al. (2010); Chen et al. (2011); Grigas et al. (2017); Aggarwal et al. (2019); Balseiro et al. (2020);

Bidding strategy as a by-product of global optimization

Decentralized



Formulate the Optimization Problem from Individual Buyer's Objective

Massive Small Optimization Problems

Market Equilibrium as a by-product of bidding strategy

General Formulation

$$\max_{b_t} \sum_{t=1}^{T} v_t G_t(b_t), \qquad s.t. \sum_{t=1}^{T} H_t(b_t) \leq B.$$

- T: impression opportunities
- v_t: value for the t-th impression
- $G_t(b_t)$: probability of winning
- $H_t(b_t)$: expected cost
- B: budget

Knapsack Problem: Modeling G_t and H_t as ***step*** functions [Chakrabarty et al. (2008); Zhou and Naroditskiy (2008)]

Fluid Approximation: Modeling G_t and H_t as ***smooth*** functions [Gallego and Van Ryzin (1994); Gummadi et al. (2013); Balseiro et al. (2015); Fernandez-Tapia et al. (2017)]

Surplus Maximization and Optimal Bidding Strategy

$$L(b_t, \lambda) = \sum_{t=1}^{T} [v_t G_t(b_t) - \lambda H_t(b_t)] + \lambda B.$$

Optimization in b_t amounts to surplus maximization:

$$\boxed{(v_t/\lambda)} G_t(b_t) - H_t(b_t)$$

budget adjusted value

First order optimality condition implies

$$b_t^*(\lambda) = \left(\frac{h_t}{g_t}\right)^{-1} \left(\frac{v_t}{\lambda}\right)$$

Surplus Maximization and Optimal Bidding Strategy

$$L(b_t, \lambda) = \sum_{t=1}^{T} [v_t G_t(b_t) - \lambda H_t(b_t)] + \lambda B.$$

Optimization in b_t amounts to surplus maximization:

$$\boxed{(v_t/\lambda)} G_t(b_t) - H_t(b_t)$$

budget adjusted value

First order optimality condition implies

$$b_t^*(\lambda) = \left(\frac{h_t}{g_t}\right)^{-1} \left(\frac{v_t}{\lambda}\right) \begin{cases} \dfrac{v_t}{\lambda} & \text{2nd price auction} \\ \left(I + \dfrac{G_t}{g_t}\right)^{-1} \left(\dfrac{v_t}{\lambda}\right) & \text{1st price auction} \end{cases}$$

Online Methods in the Dual

KKT condition implies

$$\sum_{t=1}^{T} \boxed{H_t(b_t^*(\lambda^*))} = B$$

spend

The online loss function

$$L(b_t^*, \lambda) = \sum_{t=1}^{T} \boxed{L_t(b_t^*, \lambda) := [v_t G_t(b_t^*) - \lambda H_t(b_t^*)] + \lambda B/T.}$$

Derivative of the online loss

$$L_t'(b_t^*, \lambda) = \boxed{B/T - H_t(b_t^*)}$$

discrepancy between forecasted and actual spend

Follow-the-leader: Find $\lambda_{t+1} = \mathrm{argmin}_{\lambda \geq 0} \sum_{\tau=1}^{t} L_\tau(b_\tau^*, \lambda).$

Follow-the-regularized-leader: *Online Gradient Descent*, *Online Mirror Descent*, etc.

Other Types of Constraints

Cost Control (ROI target)

$$\max_{b_t} \sum_{t=1}^{T} v_t G_t(b_t), \quad s.t. \sum_{t=1}^{T} H_t(b_t) \leq \min\left(B, C\sum_{t=1}^{T} v_t G_t(b_t)\right).$$

$$b_t^*(\lambda, \mu) = \left(\frac{h_t}{g_t}\right)^{-1} \left(\frac{1+\mu C}{\lambda+\mu} \cdot v_t\right)$$

Budget Delivery Control

$$\max_{b_t} \sum_{t=1}^{T} v_t G_t(b_t), \; s.t. \sum_{t=1}^{T} H_t(b_t) \leq B, \sum_{t \in T_k} H_t(b_t) \leq B_k, \forall k.$$

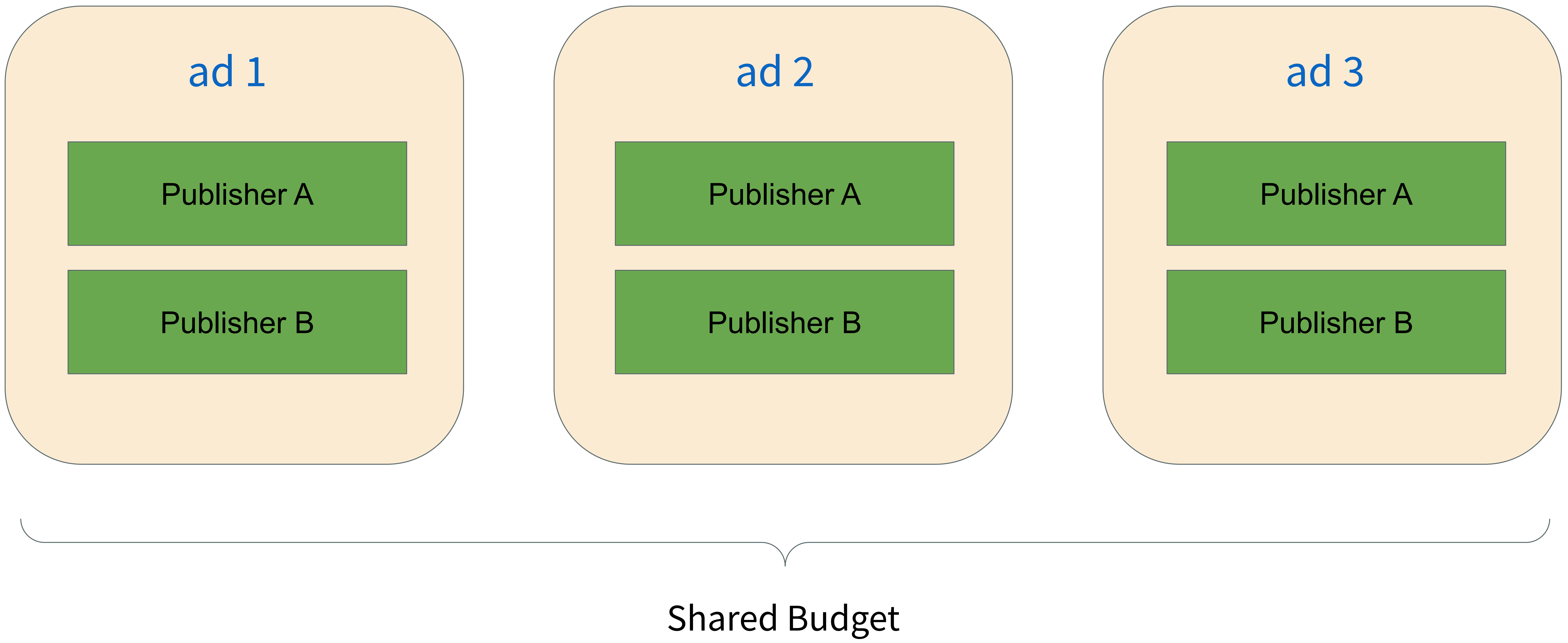
$$b_t^*(\lambda, \lambda_k) = \left(\frac{h_t}{g_t}\right)^{-1} \left(\frac{v_t}{\lambda+\lambda_k}\right)$$

Guaranteed Delivery

$$\max_{b_t} \sum_{t=1}^{T} v_t G_t(b_t), \; s.t. \sum_{t=1}^{T} H_t(b_t) \leq B, \sum_{t \in T_k} v_t G_t(b_t) \geq V_k, \forall k.$$

$$b_t^*(\lambda, \mu_k) = \left(\frac{h_t}{g_t}\right)^{-1} \left(\frac{1+\mu_k}{\lambda} \cdot v_t\right)$$

Multiple Placements, Group of Ads and Marginal ROI

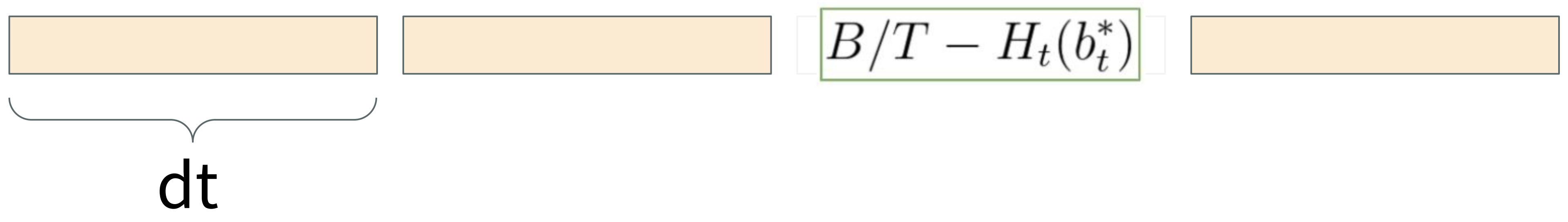


The proposed bidding framework ensures optimality, which implies

equality of marginal ROI among all ads across all placements

Engineering Considerations

- Batch online update



- Parameter Tuning
  - Normalization of λ
  - Empirical vs. Asymptotic Regret

- Forecast Error
  - Sensitivity Analysis on T
  - Relative Forecast vs. Total Forecast

- Model Predictive Control

Initialization

Theorem (Gao, Yang, Chen, Liu, El Karoui '22)

Competitor's bids $\sim Lognormal(\mu, \sigma^2)$

value $\sim Lognormal(\mu', (\sigma')^2)$

Then optimal λ^* is the solution to

$$e^{\mu+\frac{\sigma^2}{2}}\Phi\left(\frac{\mu' - \mu - \ln\lambda^* - \sigma^2}{\sqrt{(\sigma')^2 + \sigma^2}}\right) = \frac{B}{T}$$

Thank you!

LinkedIn