# Advancements in the Control of Dynamic Matching Markets

Ali Aouad (LBS)

Dec 2024, CIRM

# Online Stochastic Matching

- Problem statement: sequentially matching agents (supply resources, customer queries) in a graph with stochastic arrivals

# Online Stochastic Matching

- Problem statement: sequentially matching agents (supply resources, customer queries) in a graph with stochastic arrivals
- Expansive literature in computer science / operations research
  - Online matching: KVV [’90], MGS [‘12], JL [‘14], PPSW[‘21]…
  - Matching queues: CK [‘09], GW [‘14], AAGK [‘17], TX [‘17]…

# Online Stochastic Matching

- Online matching problem: sequentially matching supply (resources) with demand (customer queries) across a bipartite graph

- Expansive literature in computer science / operations research
  - Online matching: KVV ['90], MGS ['12], JL ['14], PPSW['21]…
  - Matching queues: CK ['09], GW ['14], AAGK ['17], TX ['17]…

Optimization-based matching policies for dynamic processes

1. Dynamic arrival/departure process
2. Correlated arrival process

# Dynamic Stochastic Matching Under Limited Time

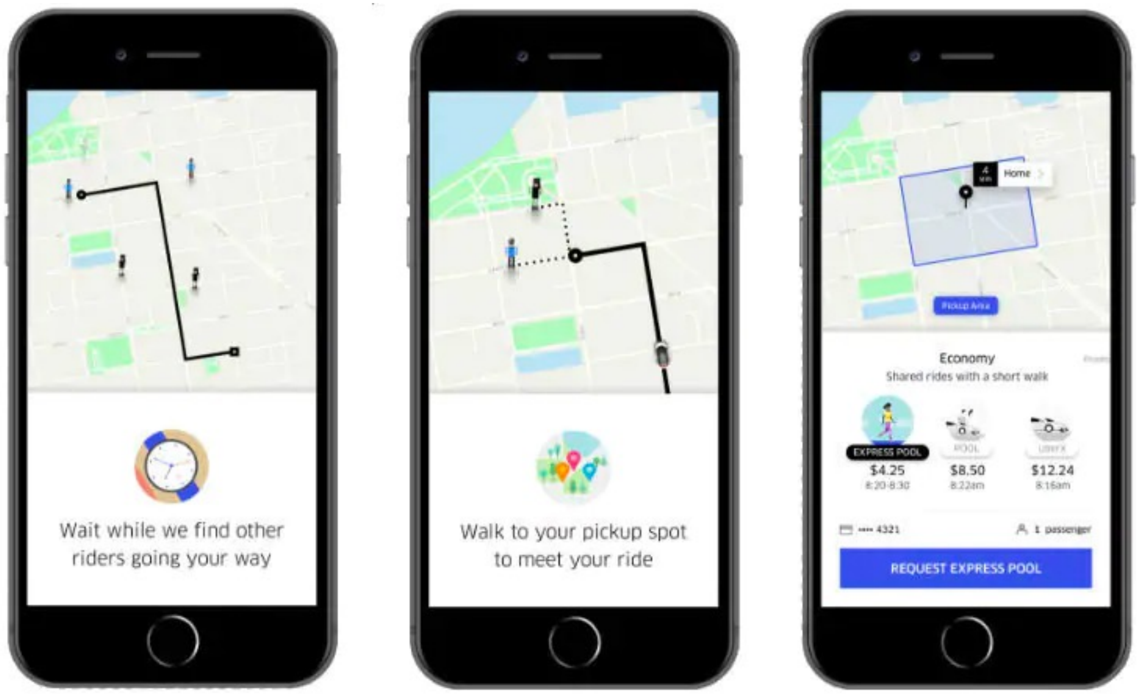
Joint work with Omer Saritaç (LBS)

# Role of “Timing” in Matching Platforms

- Matching markets are dynamic (= new agents enter or abandon the market)

# Role of "Timing" in Matching Platforms

- Matching markets are dynamic (= new agents enter or abandon the market)




E.g., Car-Pooling [1]

*"Longer initial wait times enabled the app to make more efficient matches"*

[1] Korolko et al. ['19], HBS case study: The Launch of Express Pool ['18]

# Contributions—what's coming

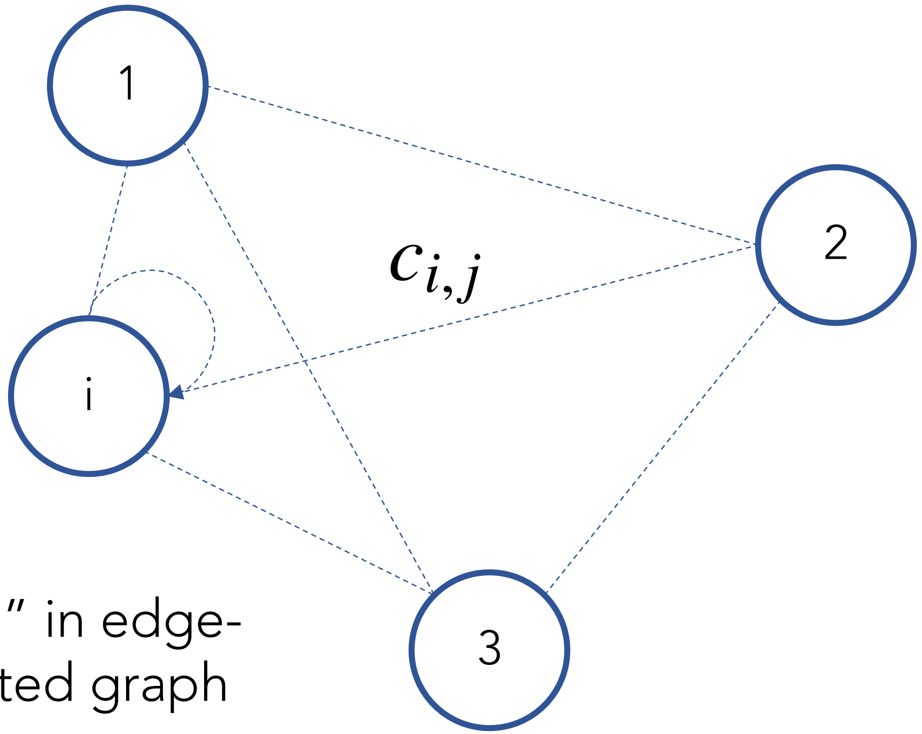
1. Modeling approach: Dynamic stochastic matching

2. Simple provably good matching algorithms
   - New mathematical programming relaxations
   - Threshold policies, online correlated rounding

3. Numerical simulations-car-pooling system

# Dynamic Stochastic Matching Model

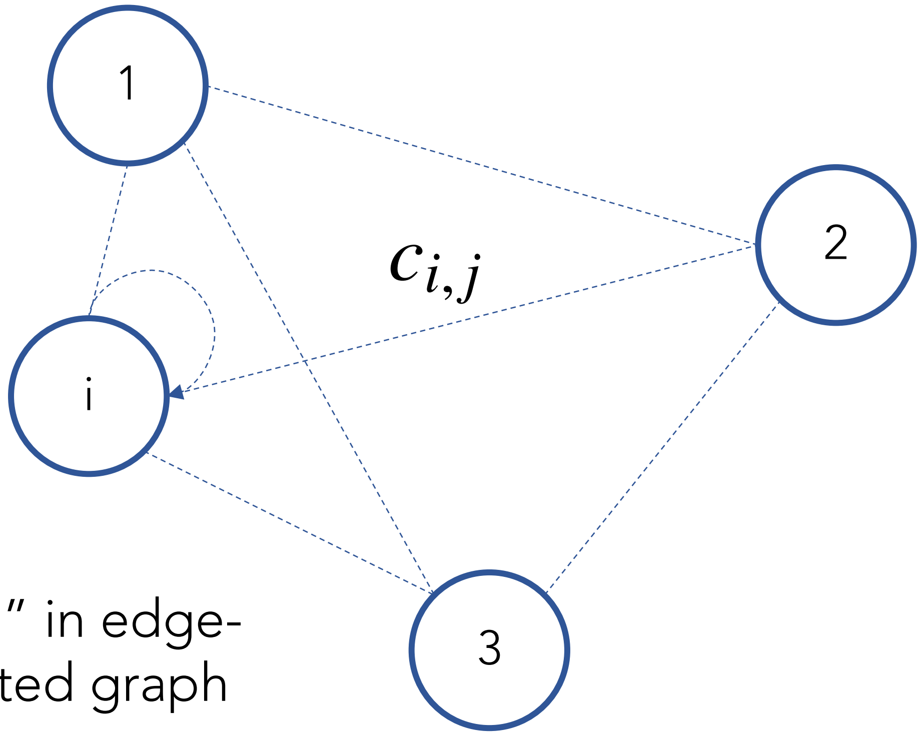
# Dynamic Stochastic Matching Model




"Types" in edge-weighted graph

1 Network

# Dynamic Stochastic Matching Model

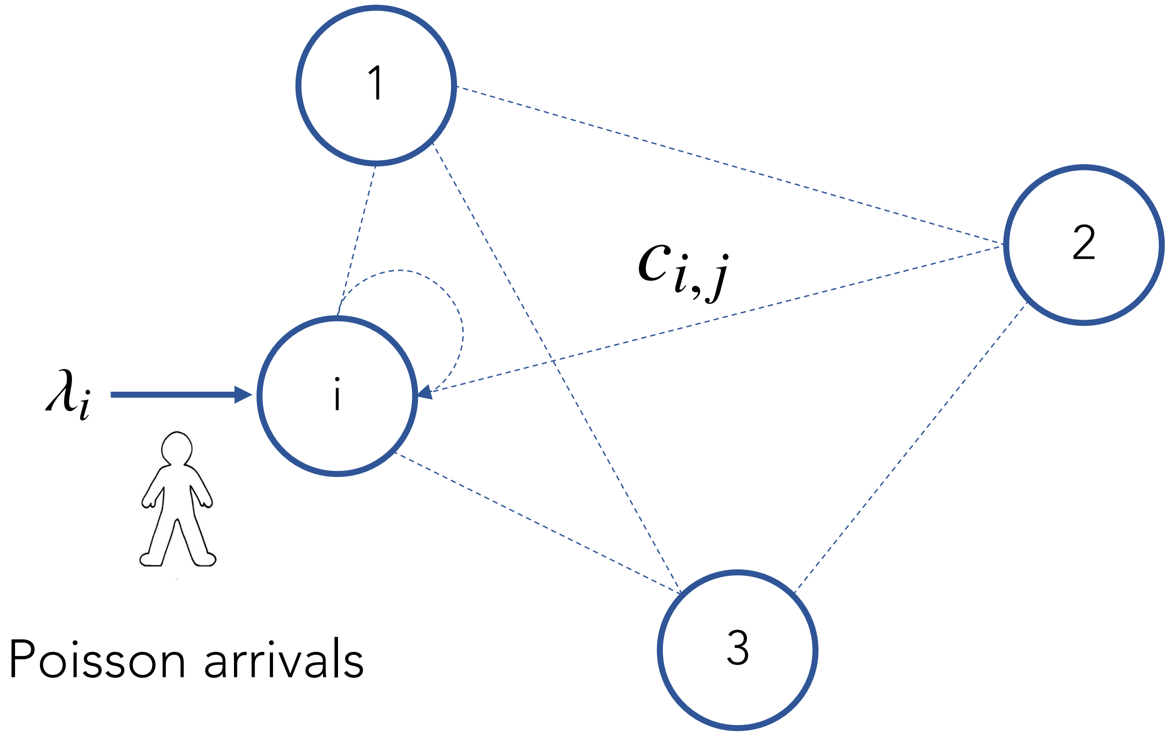



"Types" in edge-weighted graph

1 Network

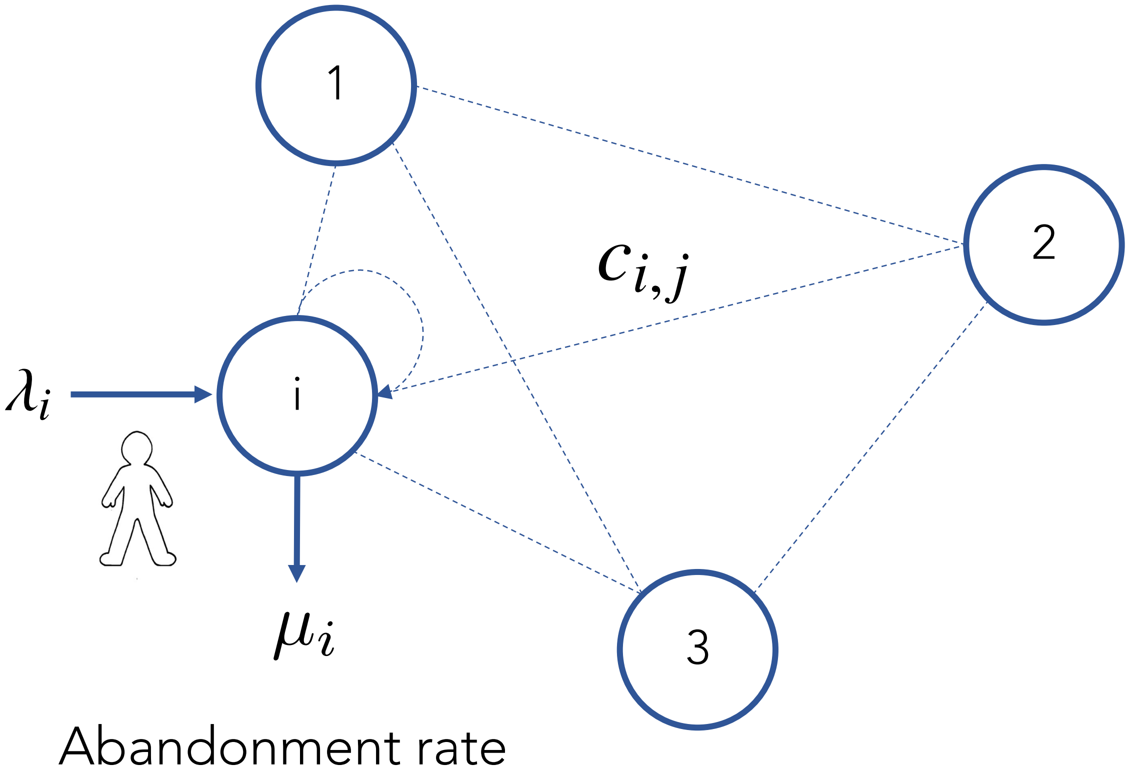
E.g., Car-pooling
Type = (Origin, Destination, "Patience")

# Dynamic Stochastic Matching Model



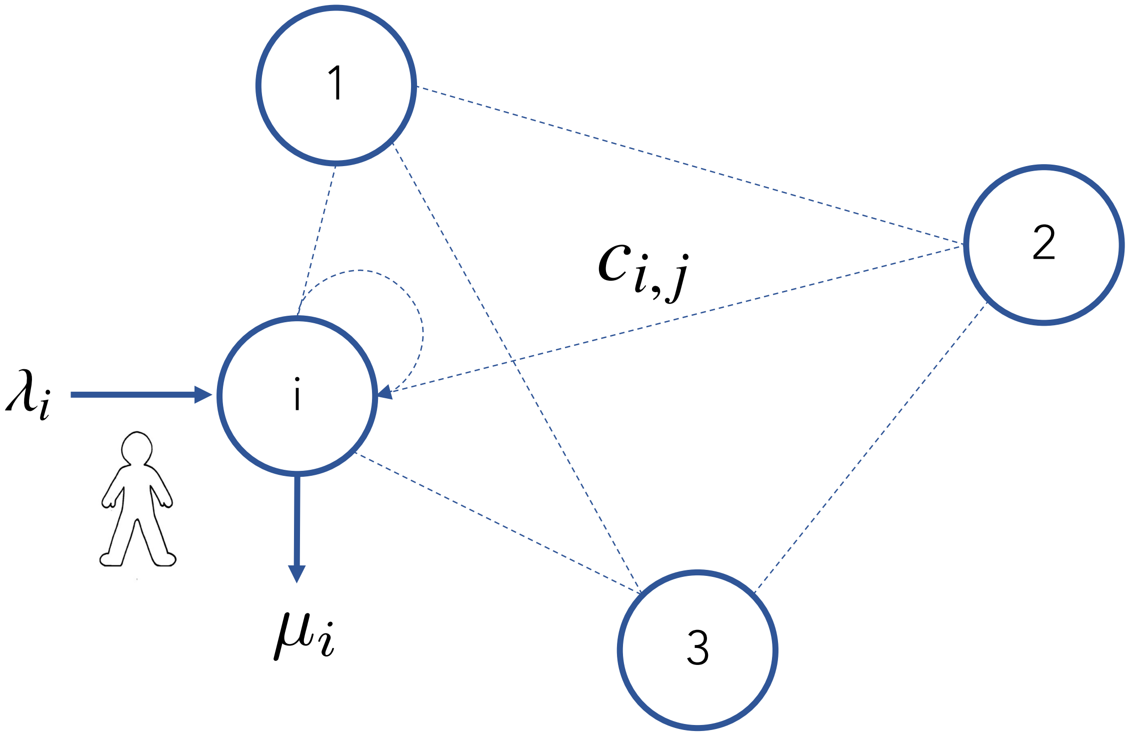

2 Market Dynamics

# Dynamic Stochastic Matching Model




2 Market Dynamics

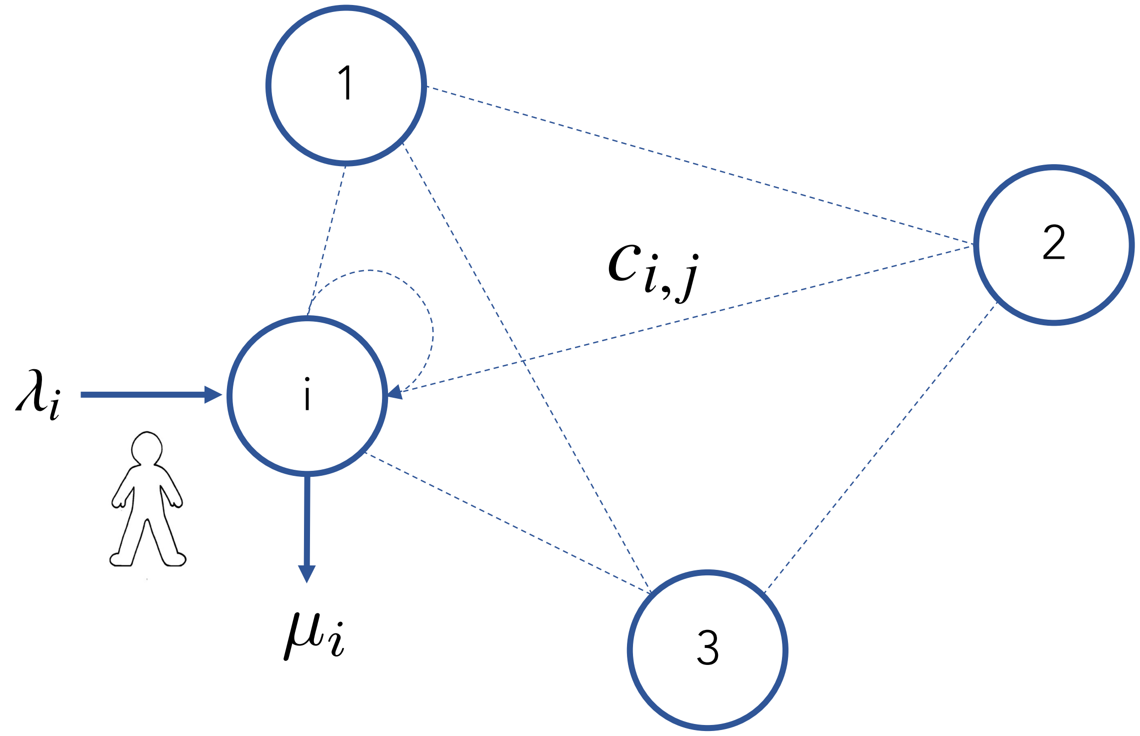
# Dynamic Stochastic Matching Model




Abandonment cost: $c_a(i)$

2 Market Dynamics

# Dynamic Stochastic Matching Model




Abandonment cost: $c_a(i)$

Cost Minimization Problem

$$\inf_{\pi} \lim_{t\to\infty} \frac{\mathbb{E}[C^{\pi}(t)]}{t}$$

Reward Maximization Problem

$$\sup_{\pi} \lim_{t\to\infty} \frac{\mathbb{E}[R^{\pi}(t)]}{t}$$

# Classical Matching vs. Dynamic Matching

| | "Classical" Online Matching | Our Setting |
|---|---|---|
| When agents arrive? | Online/Offline | Dynamic (Poisson) |
| When to match? | Immediately | Stopping time problem |
| Horizon? | Finite | Steady-state (avg. cost) |

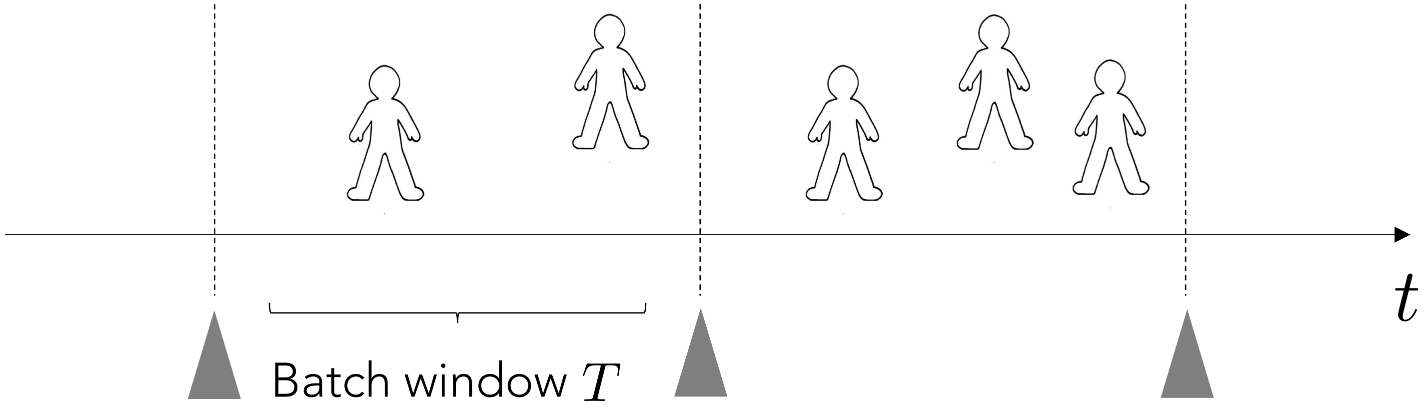
# Classical Matching vs. Dynamic Matching

| | "Classical" Online Matching | Our Setting |
|---|---|---|
| When agents arrive? | Online/Offline | Dynamic (Poisson) |
| When to match? | Immediately | Stopping time problem |
| Horizon? | Finite | Steady-state (avg. cost) |
| *Algorithm design?* | Competitive algorithms | ? |

# Performance of Batching Policies?

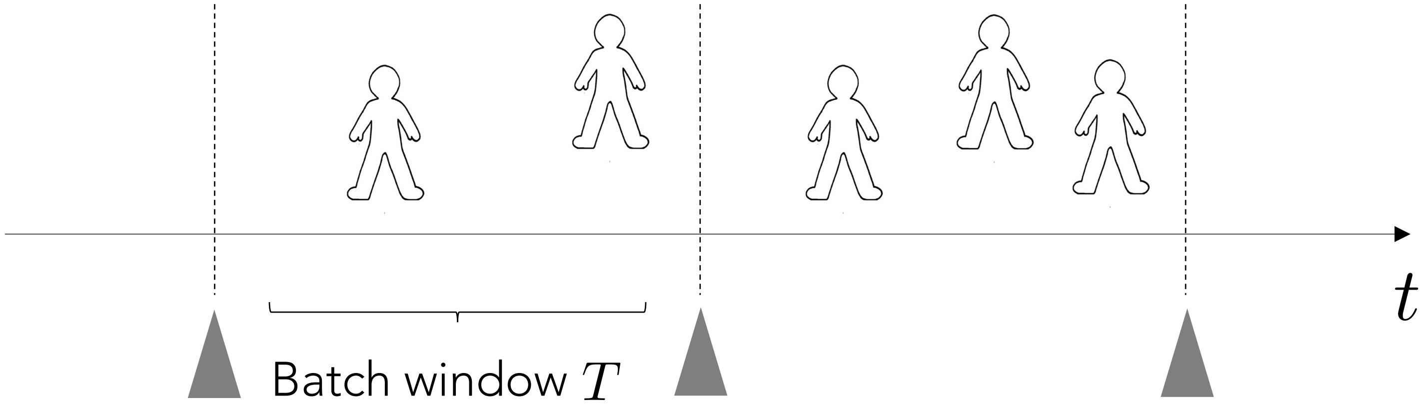
# Performance of Batching Policies?

Theory & practice: ABDJSS ['18], YZKW ['19]



# Performance of Batching Policies?

Theory & practice: ABDJSS ['18], YZKW ['19]

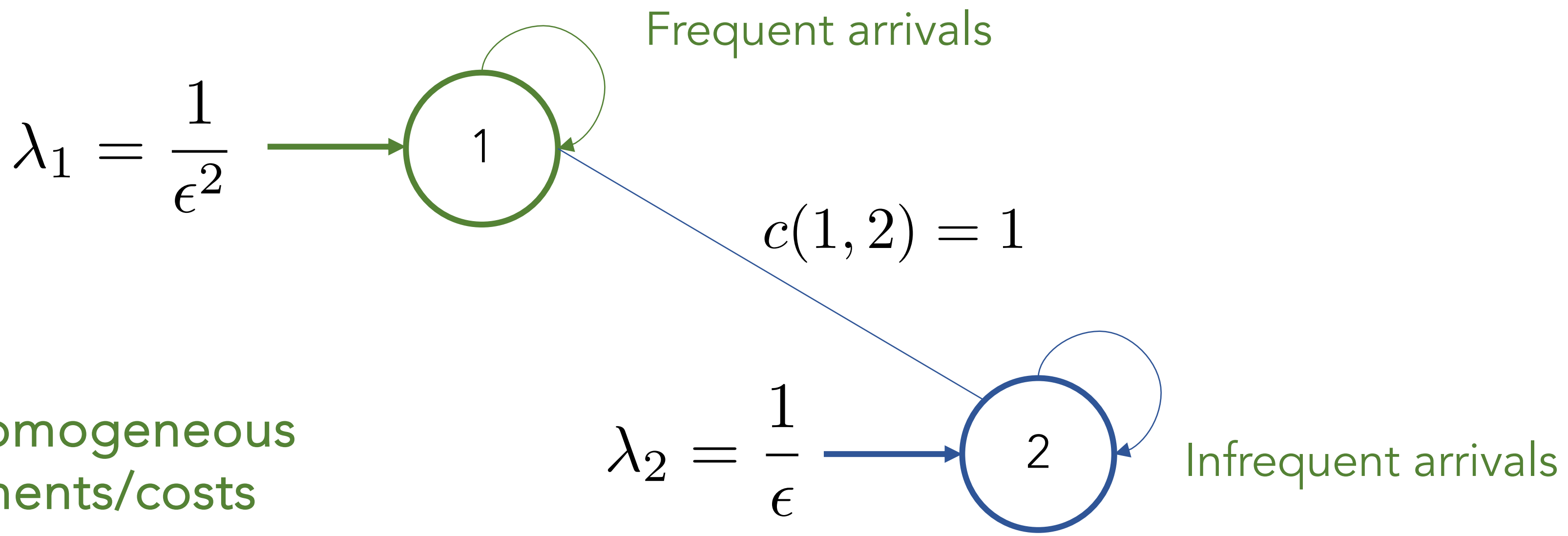



**Informal Theorem [A. and Saritac '20]:** In the min-cost setting, the performance guarantee of the batching policy with optimal batch window can be arbitrarily bad.

# Performance of Batching Policies?

**Informal Theorem [A. and Saritac '20]:** In the min-cost setting, the performance guarantee of the batching policy with optimal batch window can be arbitrarily bad.

*Proof construction:*




Even with homogeneous abandonments/costs

# Linear Programming Relaxation

# Linear Programming Relaxation

- Passive vs. active vertices: passive = matched with vertex arriving earlier
- Decision variable: $x_{i,j}$ match rate of active type-i with passive type-j vertices

# Linear Programming Relaxation

- Passive vs. active vertices: passive = matched with vertex arriving earlier
- Decision variable: $x_{i,j}$ match rate of active type-i with passive type-j vertices

$$L^* = \min_{x_{i,j}, x_{i,a} \geq 0} \quad \sum_i c_a(i) \cdot x_{i,a} + \sum_{(i,j)} c(i,j) \cdot x_{i,j}$$

# Linear Programming Relaxation

- Passive vs. active vertices: passive = matched with vertex arriving earlier
- Decision variable: $x_{i,j}$ match rate of active type-i with passive type-j vertices

$$
\begin{aligned}
L^* \quad = \quad \min_{x_{i,j}, x_{i,a} \geq 0} \quad & \sum_i c_a(i) \cdot x_{i,a} + \sum_{(i,j)} c(i,j) \cdot x_{i,j} \\
\text{s.t.} \quad & \sum_j x_{j,i} + \sum_j x_{i,j} + x_{i,a} = \lambda_i \ , \qquad \forall i
\end{aligned}
$$

# Linear Programming Relaxation

- Passive vs. active vertices: passive = matched with vertex arriving earlier
- Decision variable: $x_{i,j}$ match rate of active type-i with passive type-j vertices

$$L^* = \min_{x_{i,j}, x_{i,a} \geq 0} \sum_i c_a(i) \cdot x_{i,a} + \sum_{(i,j)} c(i,j) \cdot x_{i,j}$$

$$\text{s.t.} \quad \sum_j x_{j,i} + \sum_j x_{i,j} + x_{i,a} = \lambda_i \ , \qquad \forall i$$

$$\frac{\mu_i}{\lambda_j} \cdot x_{i,j} \leq x_{i,a} \ , \qquad \forall (i,j)$$

"Minimal" level abandonment

# Performance Metrics (Refresher)

- Competitive Ratio: Performance relative to "optimum offline"

$$\max_{\mathcal{I}} \frac{c^{\text{alg}}(\mathcal{I})}{c^{\text{off}(\mathcal{I})}(\mathcal{I})}$$

Benchmark knows all arrivals and sojourn times!

- Approximation Ratio: Performance relative to "optimum online"

$$\max_{\mathcal{I}} \frac{c^{\text{alg}}(\mathcal{I})}{c^{*}(\mathcal{I})}$$

Realistic benchmark
= best implementable policy

# Value of Dynamic Information

- Approximation Ratio: Relative performance vs. optimal policy
- Competitive Ratio: Relative performance vs. full-information policy

**Informal Theorem [A. and Saritac '20]:** For the min-cost problem, no algorithm achieves a positive constant-factor competitive ratio.

# Approximation Result for Cost-Minimization

Spatial graphs: The costs $\{c(i,j), c_a(i)\}_{i,j}$ satisfy the triangle inequality

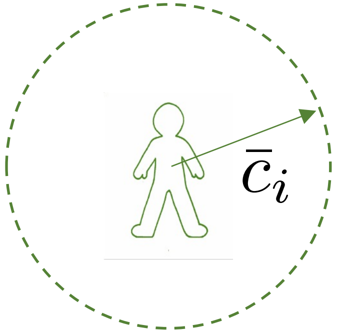
**Theorem 1 [A. and Saritac '20]:** The cost minimization dynamic matching problem admits a polynomial-time factor-3 approximation on spatial graphs with uniform $\mu$ -s.

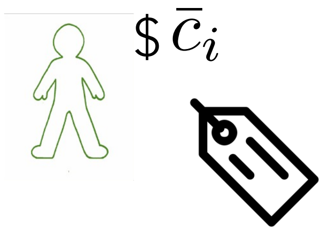
# Approximation Result for Cost-Minimization

Spatial graphs: The costs $\{c(i,j), c_a(i)\}_{i,j}$ satisfy the triangle inequality

**Theorem 1 [A. and Saritac '20]:** The cost minimization dynamic matching problem admits a polynomial-time factor-3 approximation on spatial graphs with uniform $\mu$-s.

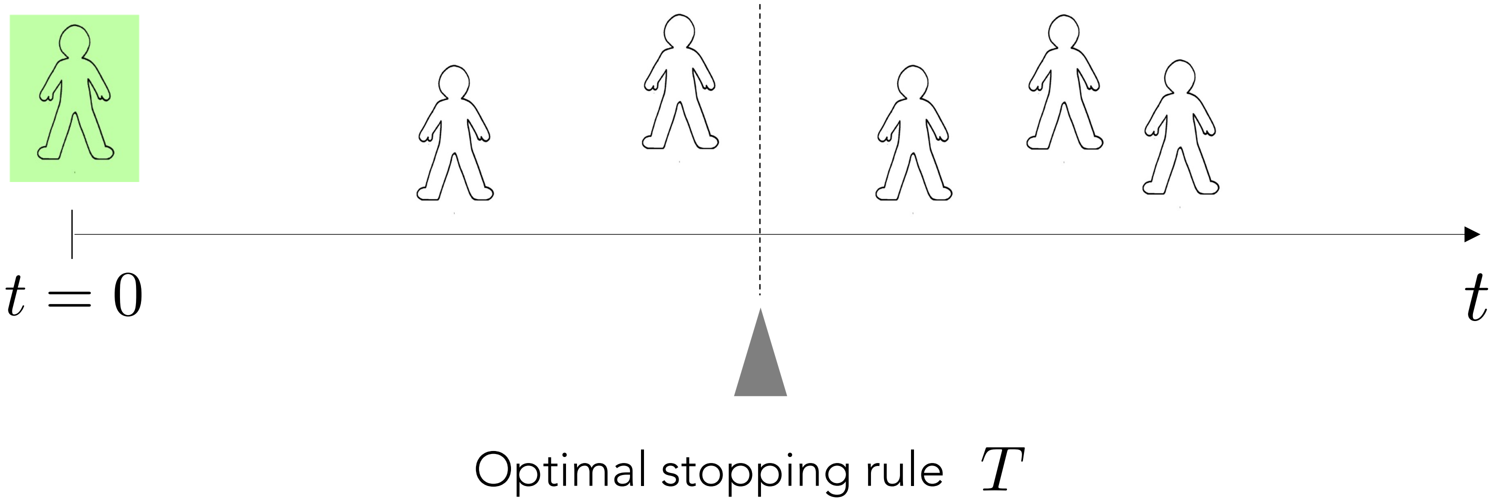
Matching policy: threshold-based or additive-approximation of value function



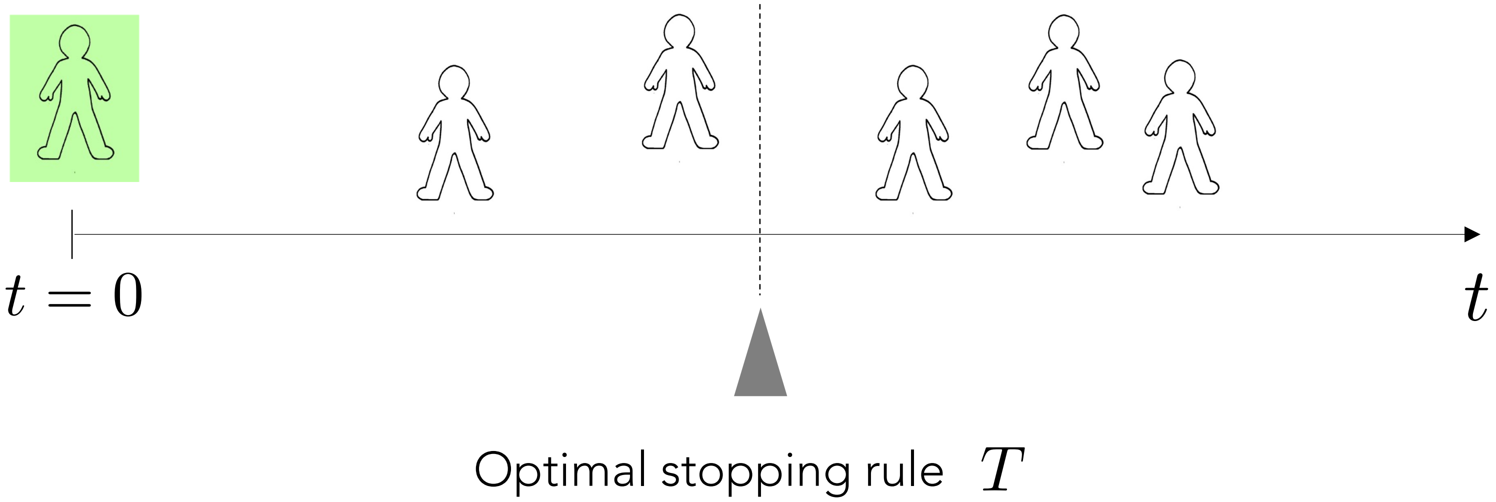
# Auxiliary Stopping Time Problem

Focus on a single active vertex (ignore competition)



# Auxiliary Stopping Time Problem

Focus on a single active vertex (ignore competition)



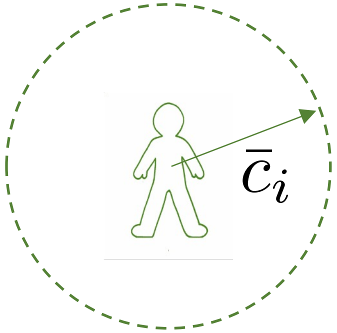

**Lem. [A. and Saritac '20]:** The optimal stopping rule is threshold-based. The optimal threshold $\bar{c}_i$ is independent of the current state and can be computed in polynomial-time.

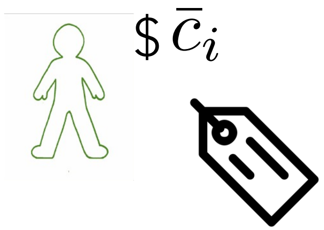
# Approximation Result for Cost-Minimization

Spatial graphs: The costs $\{c(i,j), c_a(i)\}_{i,j}$ satisfy the triangle inequality

**Theorem 1 [A. and Saritac '20]:** The cost minimization dynamic matching problem admits a polynomial-time factor-3 approximation on spatial graphs with uniform $\mu$-s.

Matching policy: threshold-based or additive-approximation of value function



# Empirical Simulation---NY Taxi Demand

- We focus on four time windows that represents various market conditions
- Split the data into training and test sets
  - Define rider types and estimate their arrival rates

| Day of week | Time of day | Number of types $\lvert\mathcal{T}\rvert$ | Sample size: Training set | Sample size: Test set |
|---|---|---|---|---|
| Monday | 7:30 AM - 8:00 AM | 272 | 50988 | 9022 |
| | 11:00 AM - 11:30 AM | 272 | 48484 | 7064 |
| Saturday | 7:30 AM - 8:00 AM | 218 | 15036 | 2535 |
| | 5:30 AM - 6:00 PM | 307 | 70035 | 10275 |

*Table: Summary statistics of the data set generated for each time window*

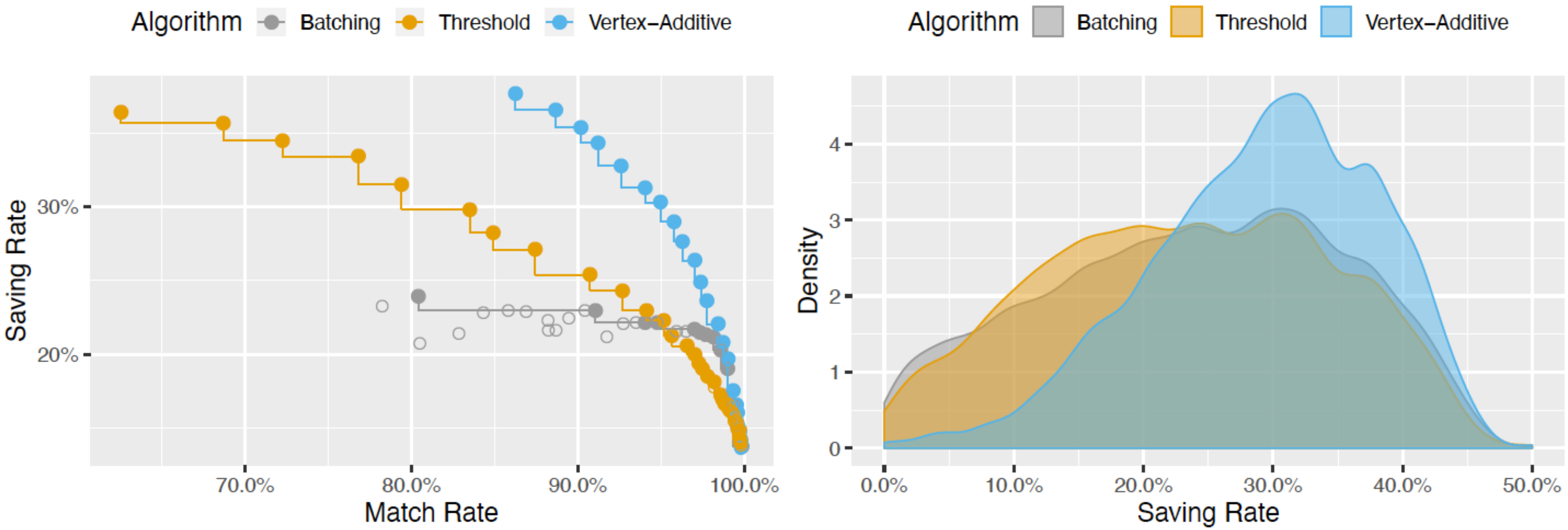
# Performance Metrics

Total Cost is an affine function of the two performance metrics

- Match Rate: Percentage of riders matched
- Saving Rate: Percentage of trip costs saved by pooling

$$\text{Total Cost} = \alpha - \beta_1 \cdot \text{Match Rate} - \beta_2 \cdot \text{Saving Rate}$$

# Numerical Results




Match rate = % matched before abandoning

Saving rate = % cost saved by pooling riders

# Main Result for General Networks

# Main Result for General Networks

**Theorem 2 [A. and Saritac '20]:** The reward maximization dynamic matching problem admits efficient constant-factor approximations:

- On general graphs, the approximation ratio is $\frac{1}{4} \cdot \left(1 - \frac{1}{e}\right)$.
- On bipartite graphs, the approximation ratio is $\frac{1}{2} \cdot \left(1 - \frac{1}{e}\right)$.
- On bipartite graphs with one impatient side, the approximation ratio is $\left(1 - \frac{1}{e}\right)$.

Our policy is a correlated rounding of the LP relaxation.

# LP Rounding Algorithm

Step 1: Solve a flow matching problem ("fluid relaxation")

Step 2: Randomization based on fractional flow ("rounding")

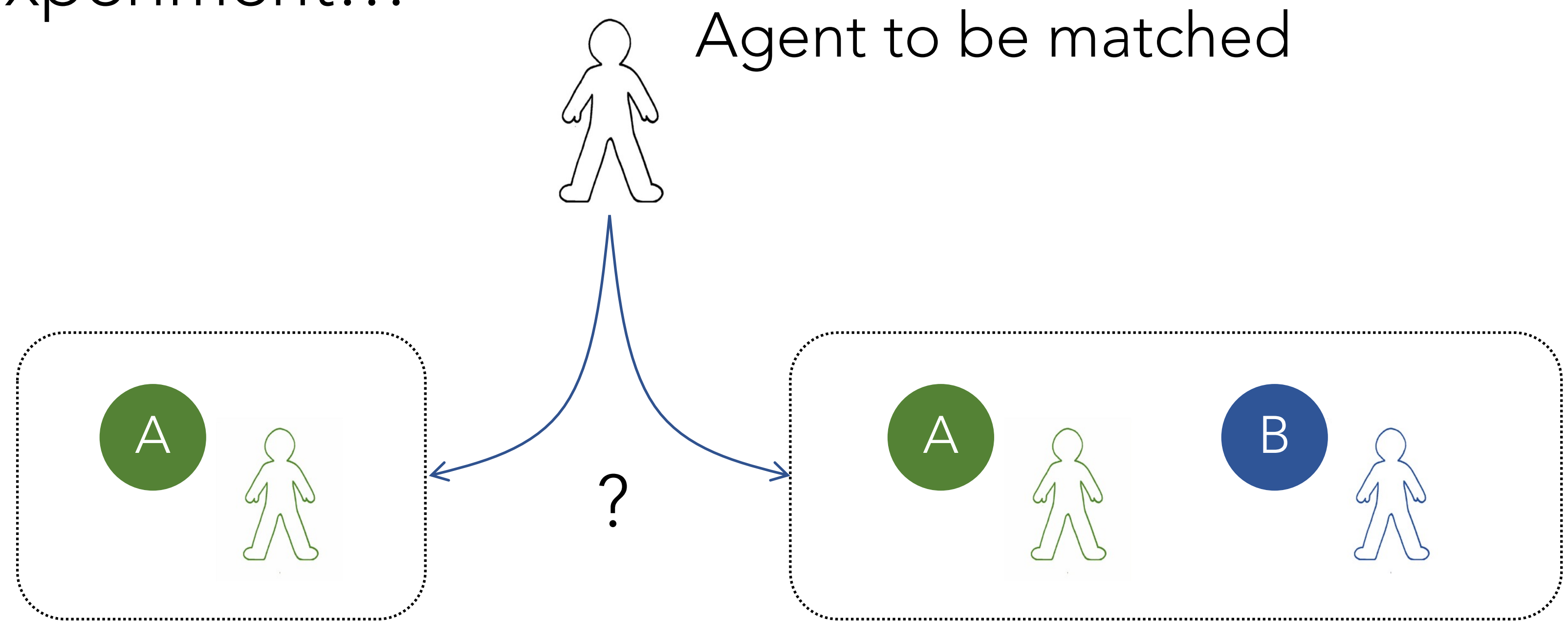
# LP Rounding Algorithm

Step 1: Solve a variant of our LP relaxation

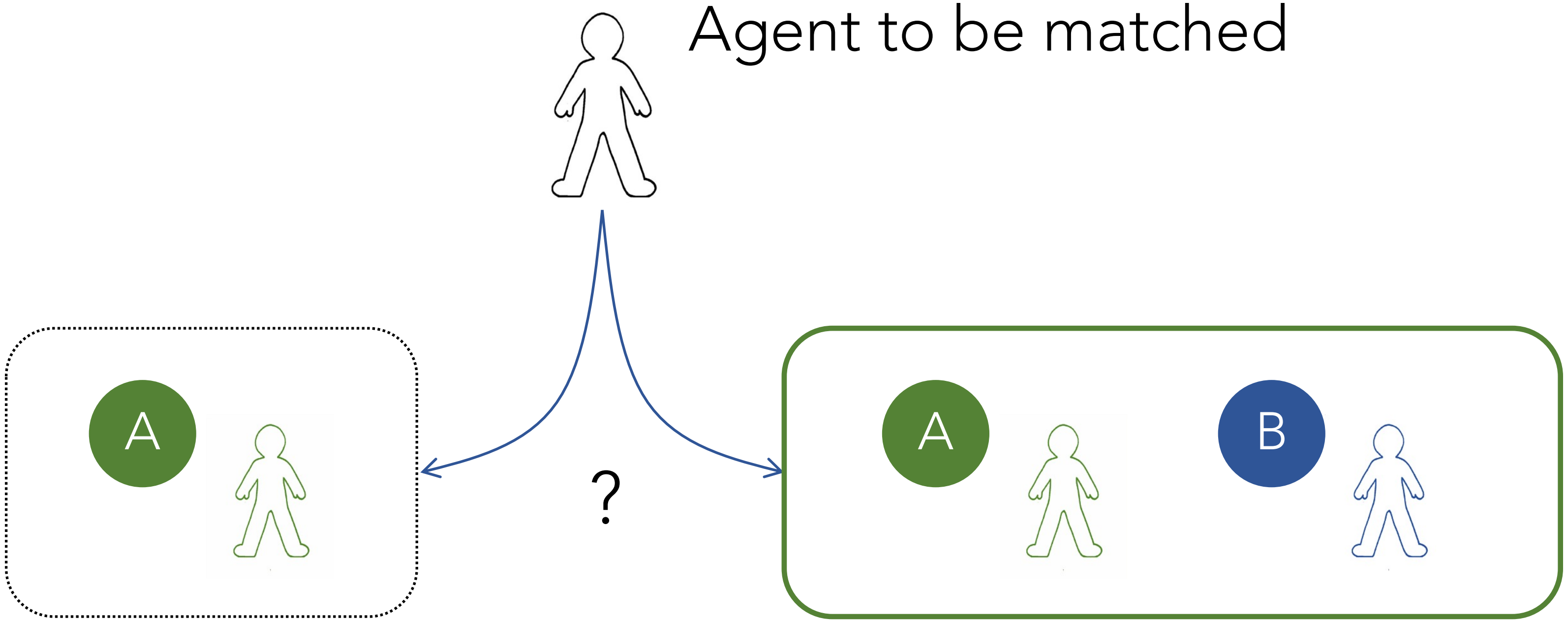
Step 2: Flexible randomization based on fractional flow

# Role of Pooling Effects

Thought experiment…



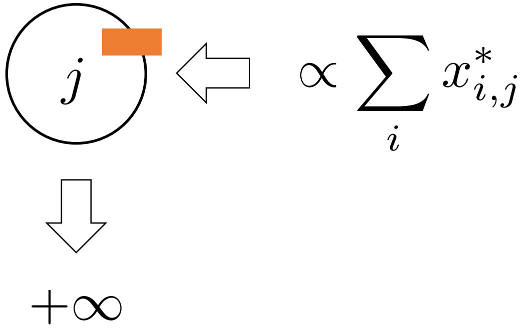
# Role of Pooling Effects



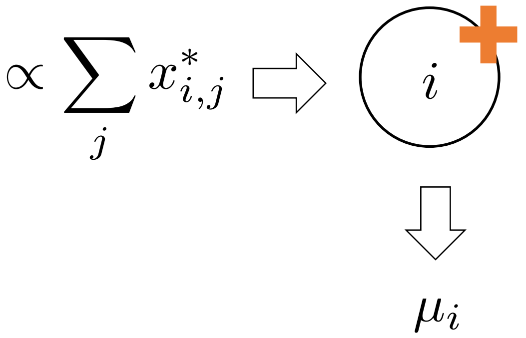

By pooling A + B, we can minimize waiting times + abandonments

# Step 1: Flexible Randomization

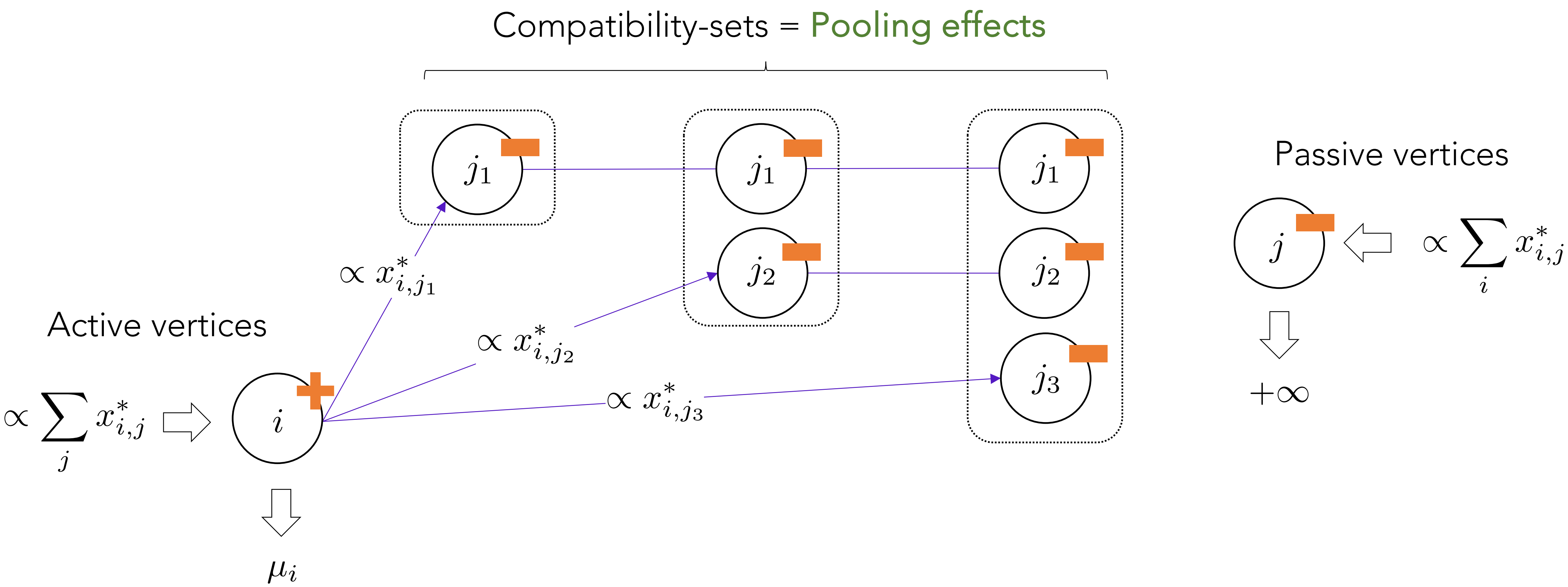
Passive vertices




Active vertices



# Step 1: Flexible Randomization



# Analysis outline

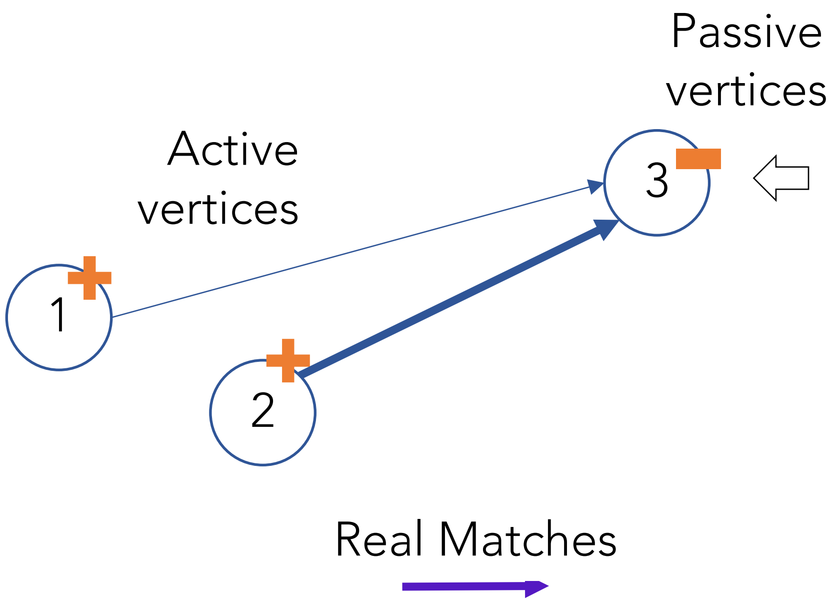
1. Flow decomposition : each arrival → randomly assigned "active" or "passive" and "compatibility set"
2. Lower bound on the availability rates of active types: virtual Markov chain
3. PASTA property: relating the lower bound on reward rates to original LP solution

# Analysis outline

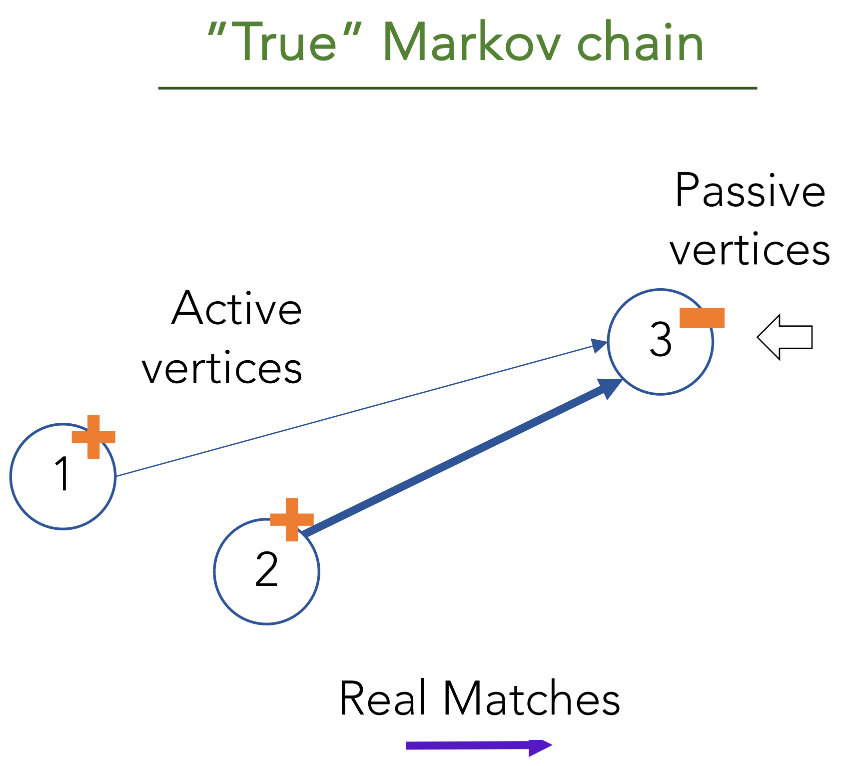
1. Flow decomposition : each arrival → randomly assigned "active" or "passive" and "compatibility set"
2. Lower bound on the availability rates of active types: virtual Markov chain
3. PASTA property: relating the lower bound on reward rates to the original LP solution

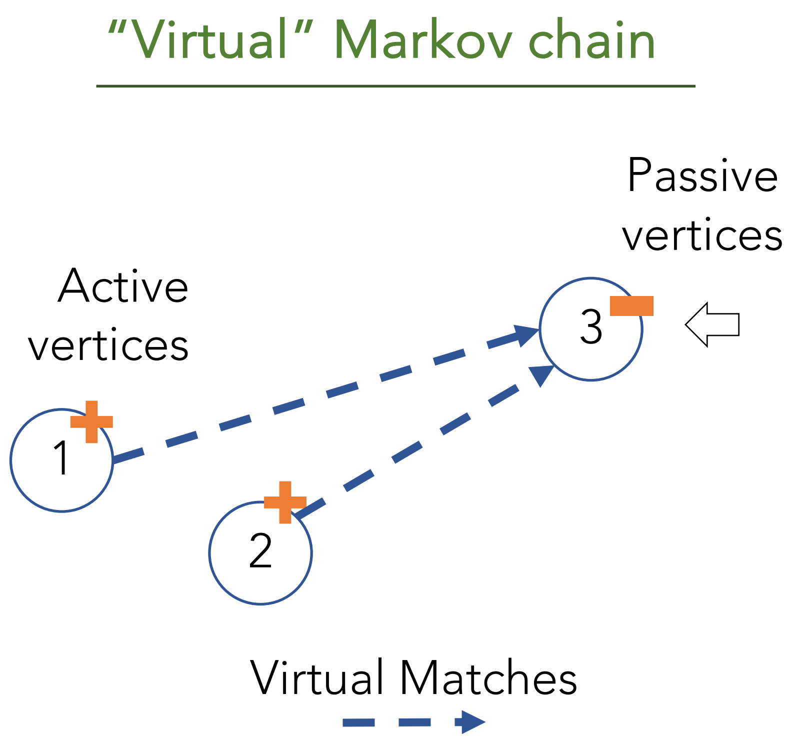
# Step 2: Lower bound via virtual Markov chain

"True" Markov chain



# Step 2: Lower bound via virtual Markov chain




*Create virtual copies of passive vertices to satisfy all active vertices*

London Business School

# A Nonparametric Framework for Online Stochastic Matching with Correlated Arrivals

Joint work with Will Ma (Columbia GSB)

# Outline

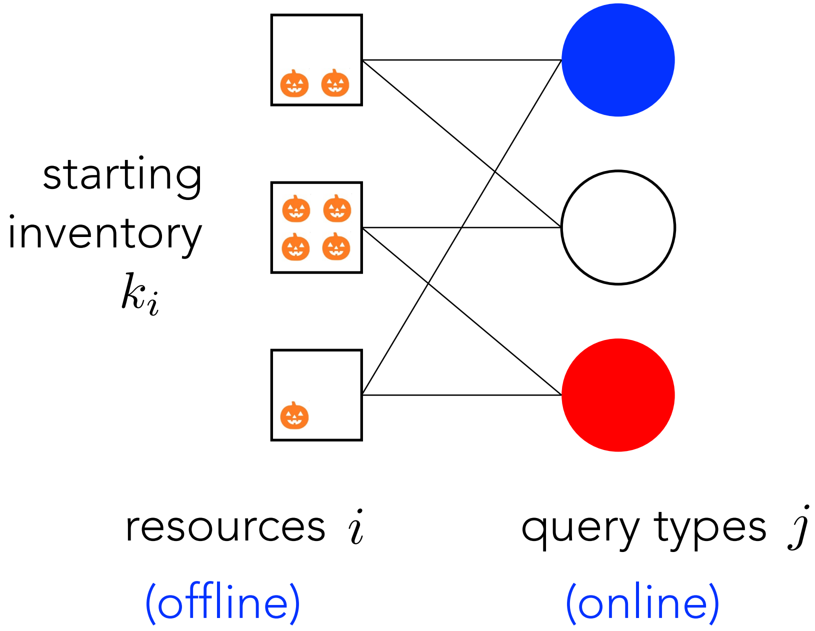
1. Nonparametric models with correlated arrivals
2. New matching algorithms with optimal competitive/approximation ratios

# Classic offline/online model

bipartite graph:

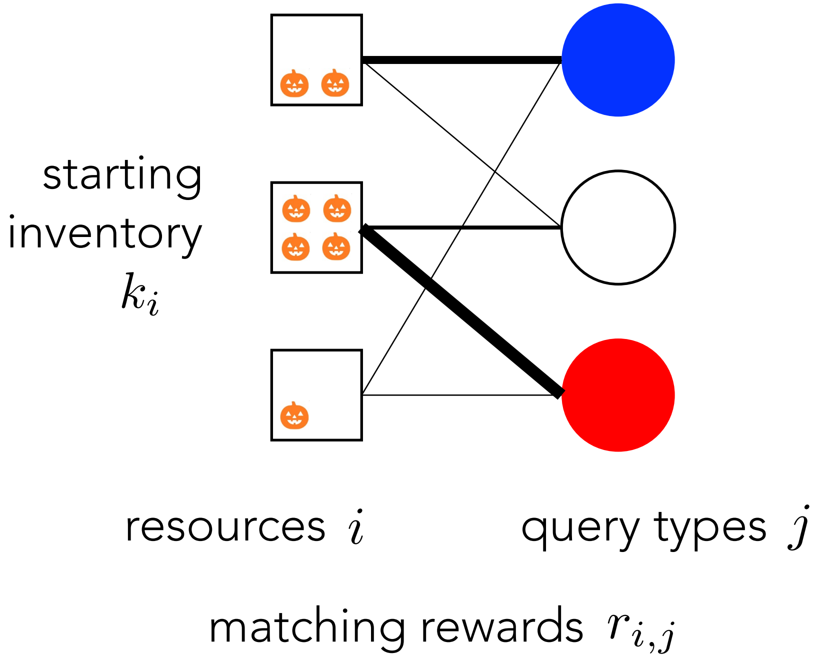
# Classic offline/online model

bipartite graph:

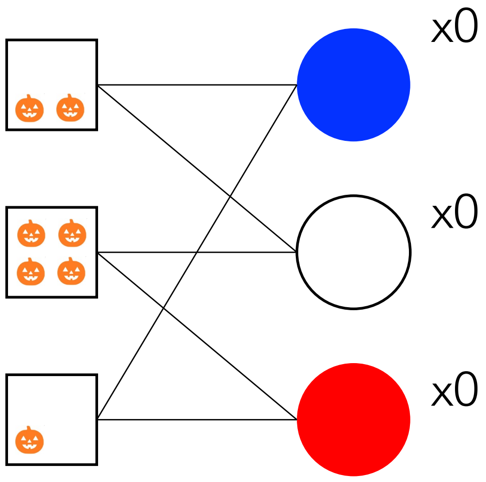


# Classic offline/online model

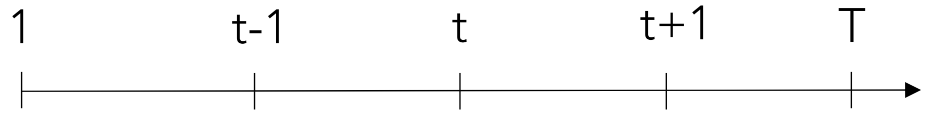
bipartite graph:



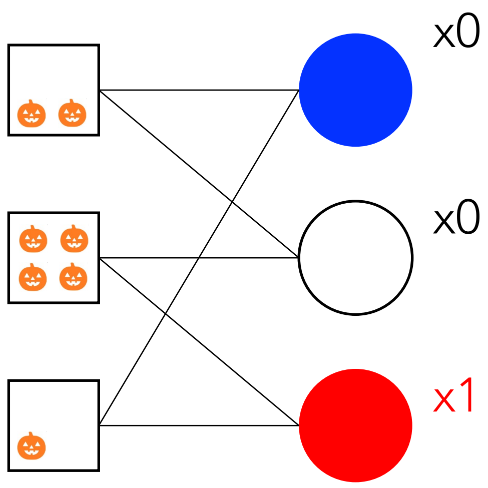
# Classic offline/online model



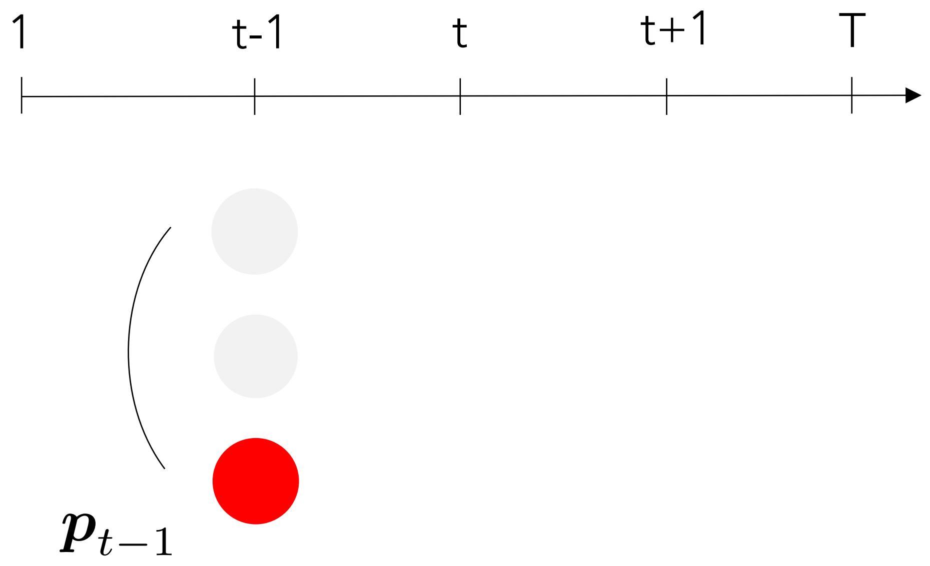

arrival process:



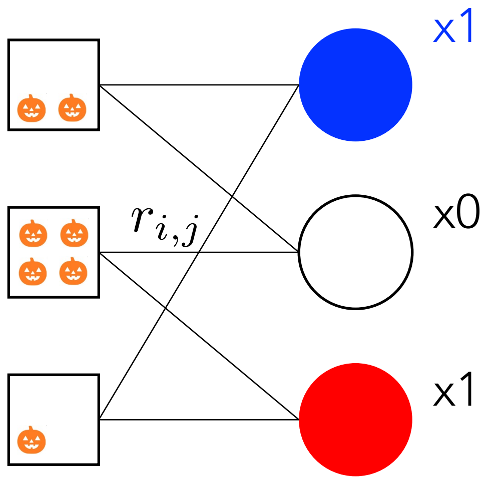
# Classic offline/online model



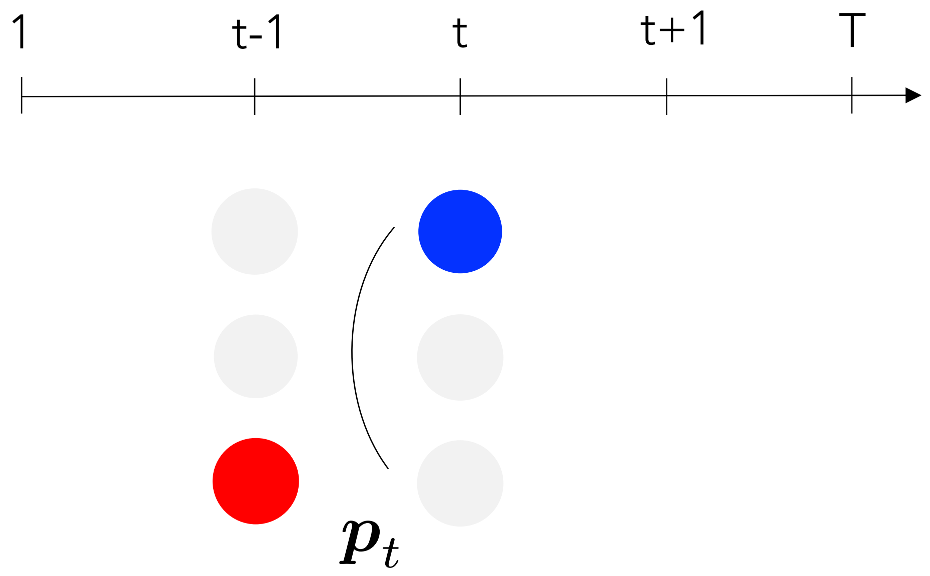

arrival process:



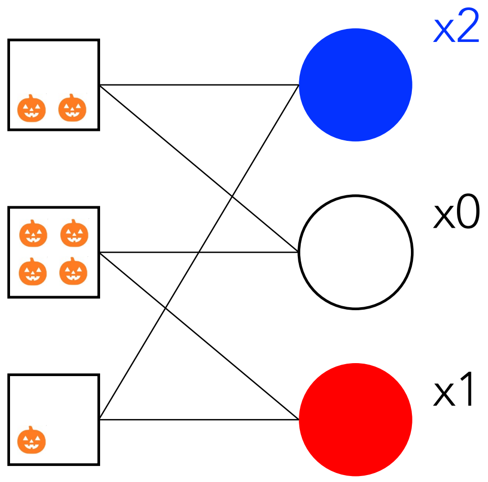
# Classic offline/online model



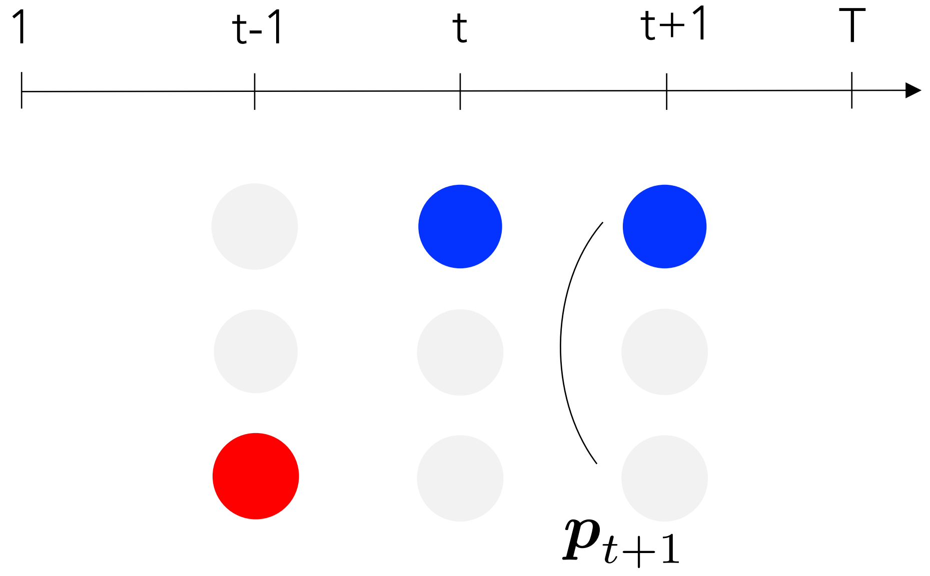

arrival process:



# Classic offline/online model

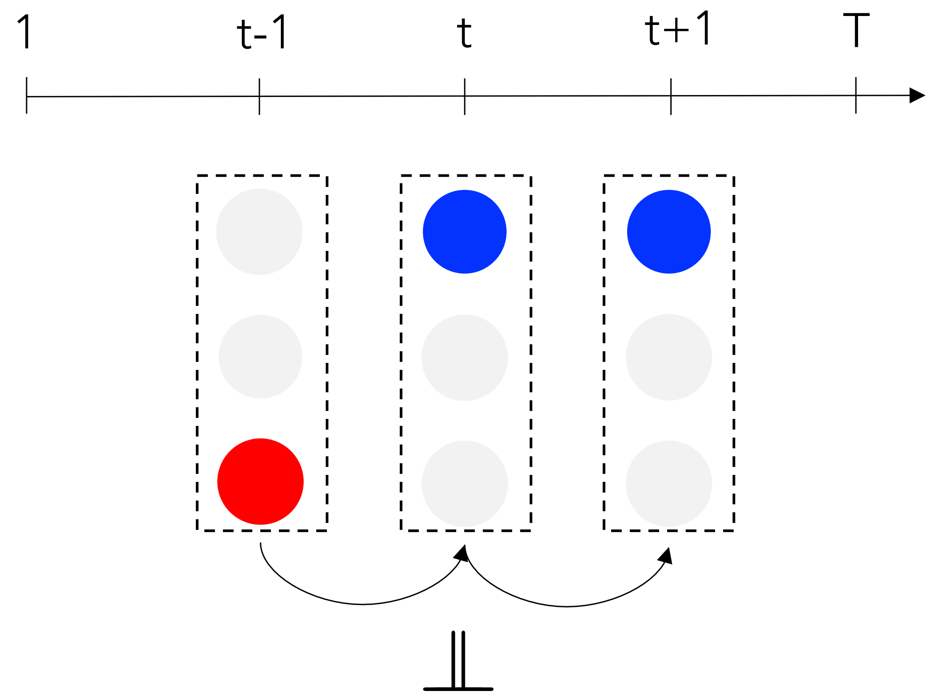



arrival process:



# Classic offline/online model

arrival process:




serial independence
(i.e., no correlations)

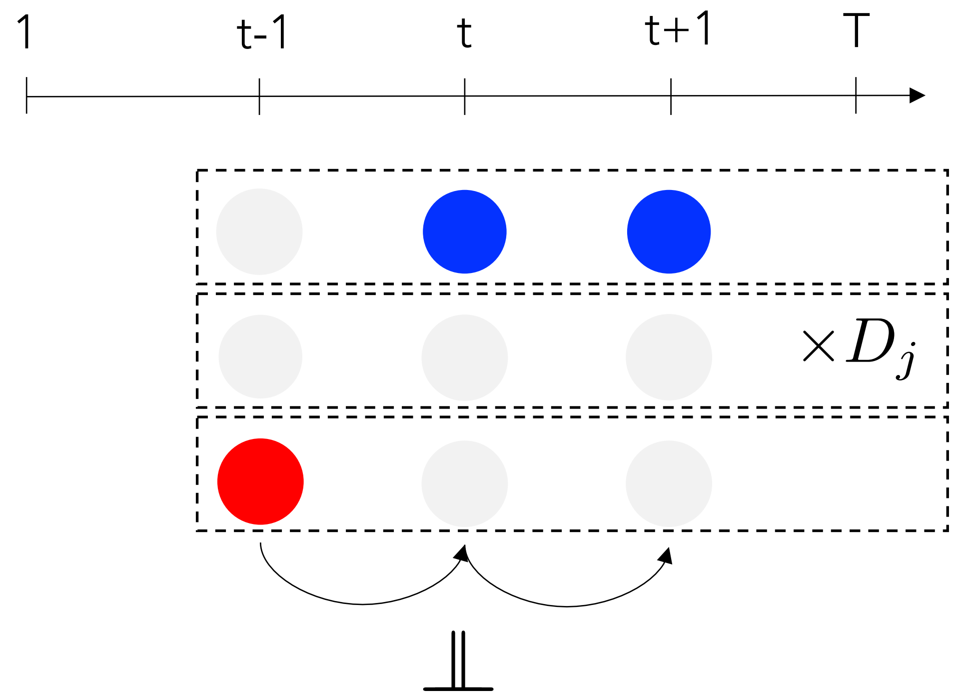
# Limitations of “serial independence” assumption

- Estimation error $T, \boldsymbol{p}_t$

# Classic offline/online model

$$D_j \sim PoissonB\left(p_{1,j}, p_{2,j}, \ldots\right)$$

arrival process:



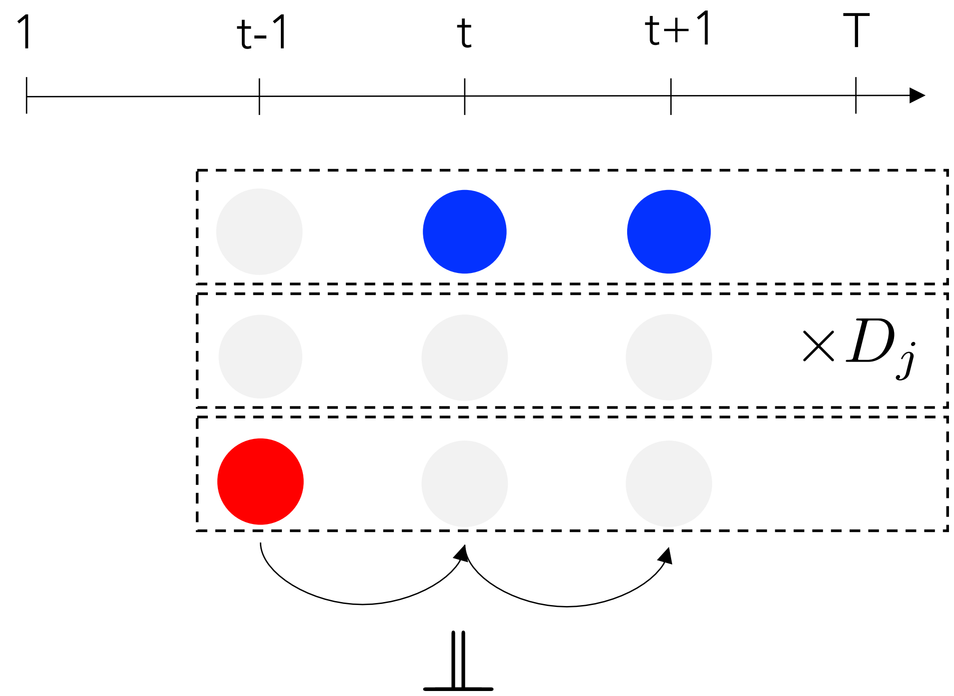
# Classic offline/online model

$$D_j \sim PoissonB\left(p_{1,j}, p_{2,j}, \ldots\right)$$

$$\Downarrow$$

$$\mathrm{Var}(D_j) \leq \mathrm{E}[D_j]$$

arrival process:



# Limitations of “serial independence” assumption

- Estimation error $T, \boldsymbol{p}_t$
- Textbook demand models, e.g., Gaussian

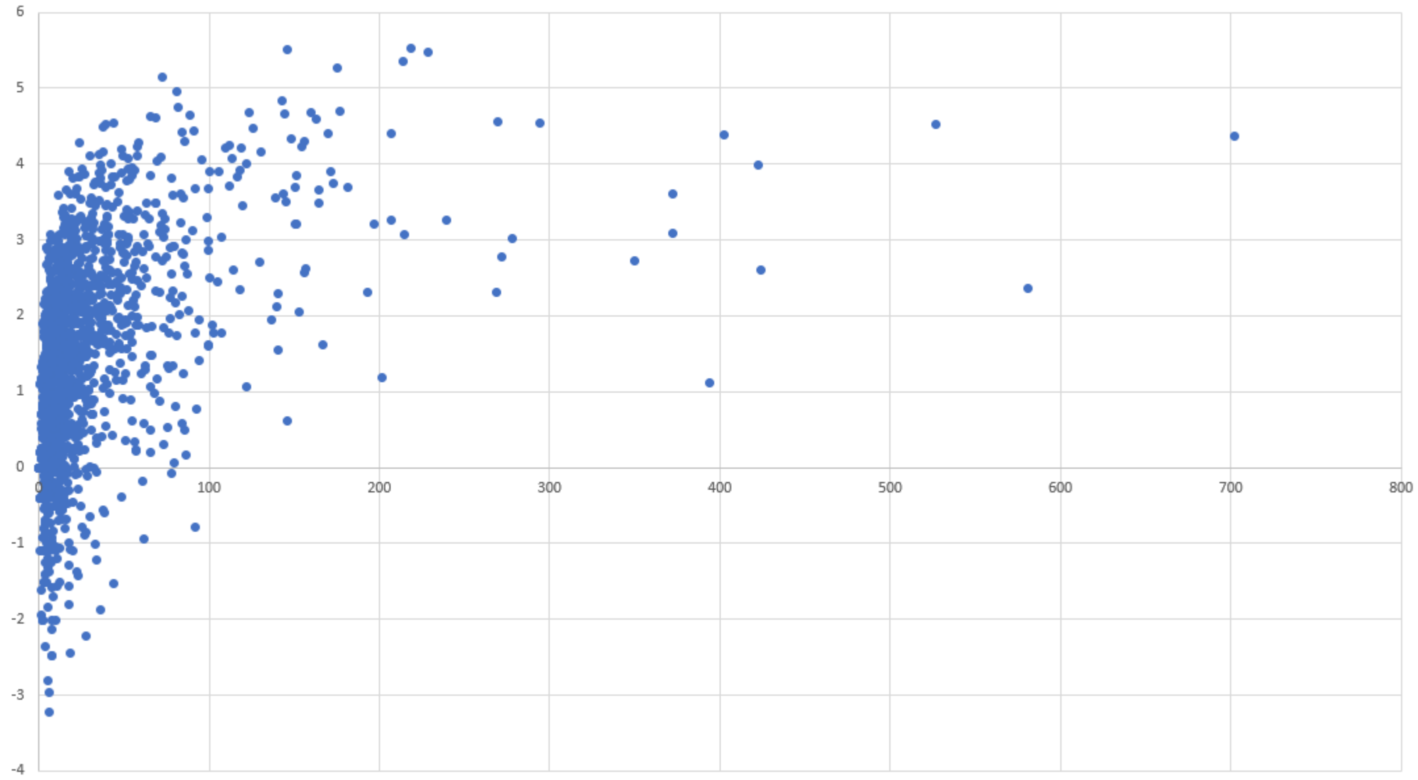
# Limitations of "serial independence" assumption

- Estimation error $T, \boldsymbol{p}_t$
- Textbook demand models, e.g., Gaussian
- A majority (70%+) of high-demand SKUs violate $\mathrm{Var}(D_j) \leq \mathrm{E}[D_j]$
  - JD.com e-commerce order data (M&SOM 2020)
  - Large fashion retailer (2014-2015 data), 200,000 SKUs

# Var($D_j$) < E[$D_j$] is unreasonably optimistic

JD.com Data[1]

$$\log\left(\frac{\hat{V}(D_j)}{\hat{\mu}(D_j)}\right)$$

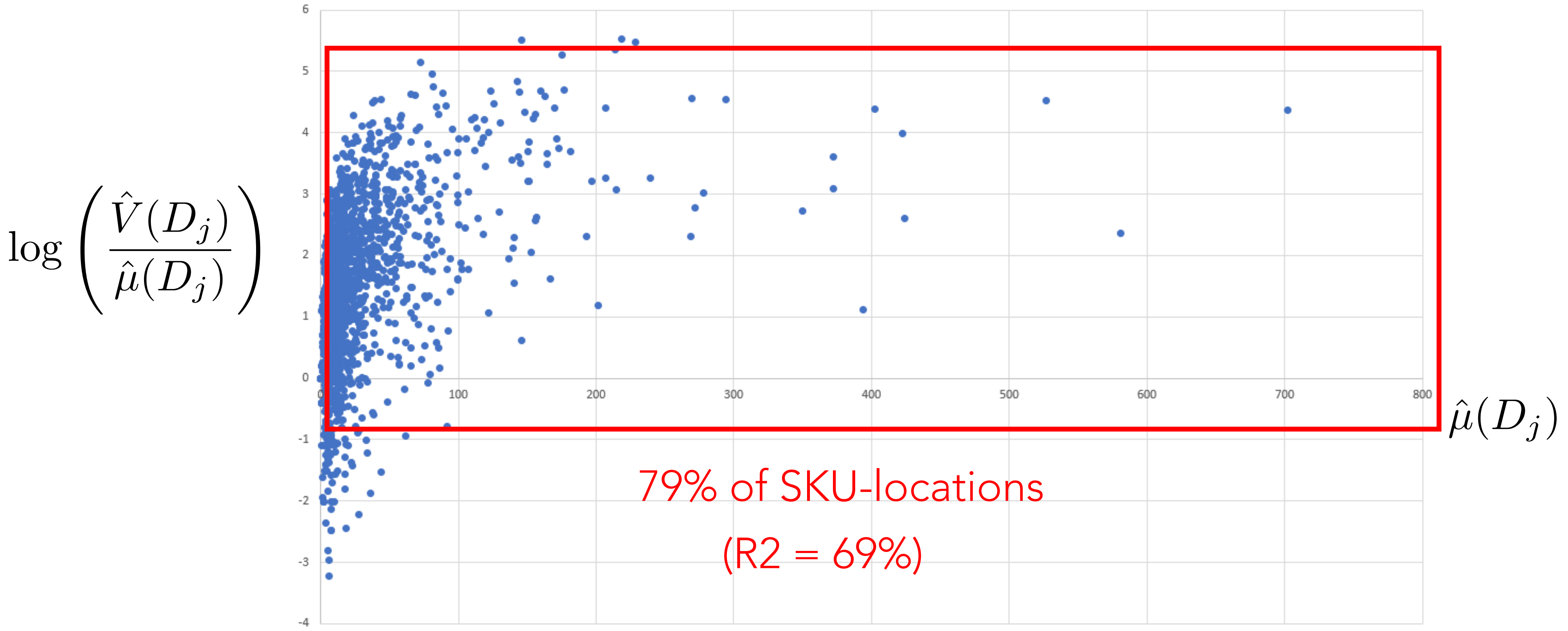



$\hat{\mu}(D_j)$

[1]Largest 40x40 (SKU,location) pairs, weekly level aggregates[1] (March 2018)

# Var($D_j$) < E[$D_j$] is unreasonably optimistic

JD.com data[1]




[1]Largest 40x40 (SKU,location) pairs, weekly level aggregates[1] (March 2018)

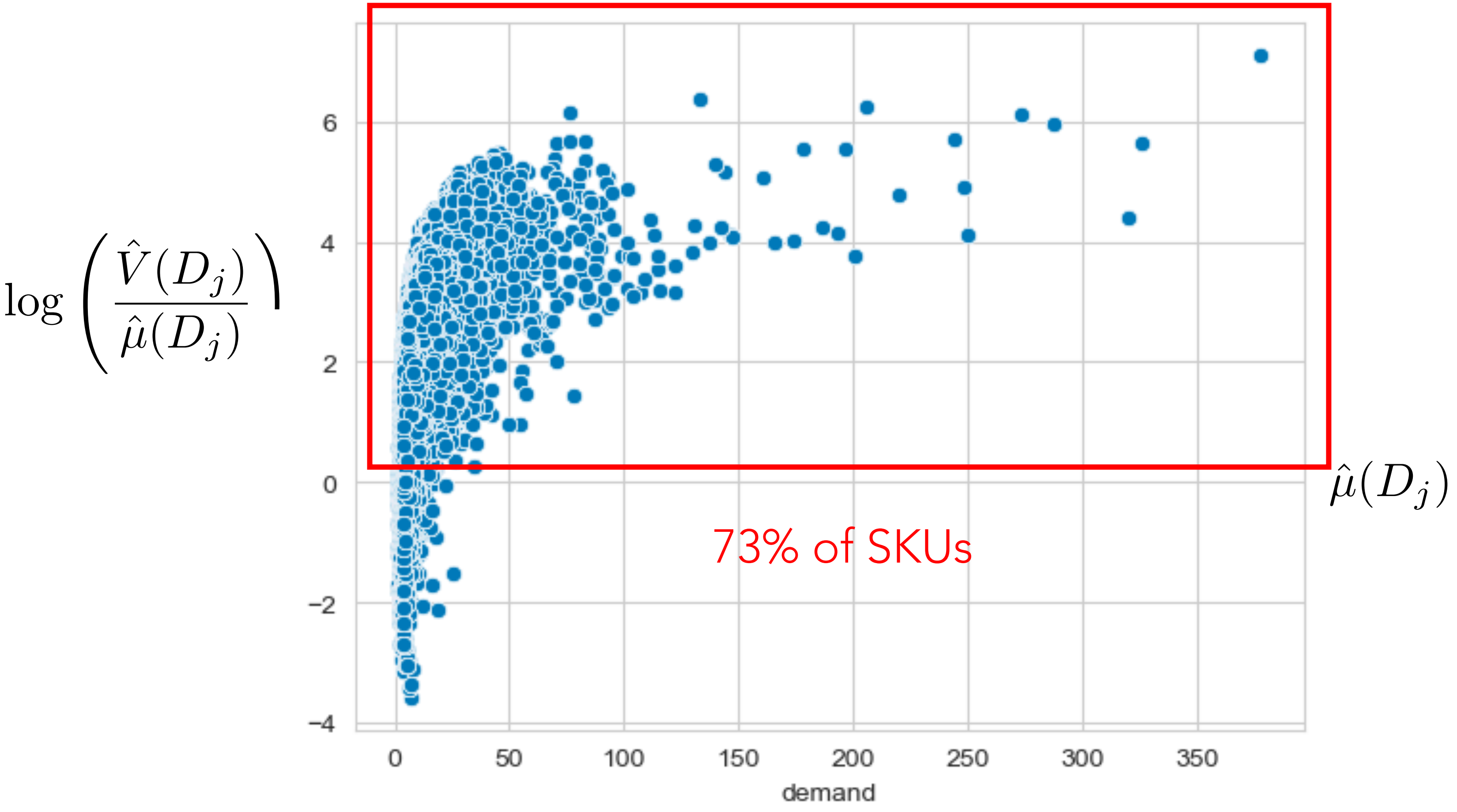
# $Var(D_j) < E[D_j]$ is unreasonably optimistic

Order fulfilment data[1]

[1]Largest 200,000 SKUs, August-December 2014, bi-weekly aggregation

# $Var(D_j) < E[D_j]$ is unreasonably optimistic

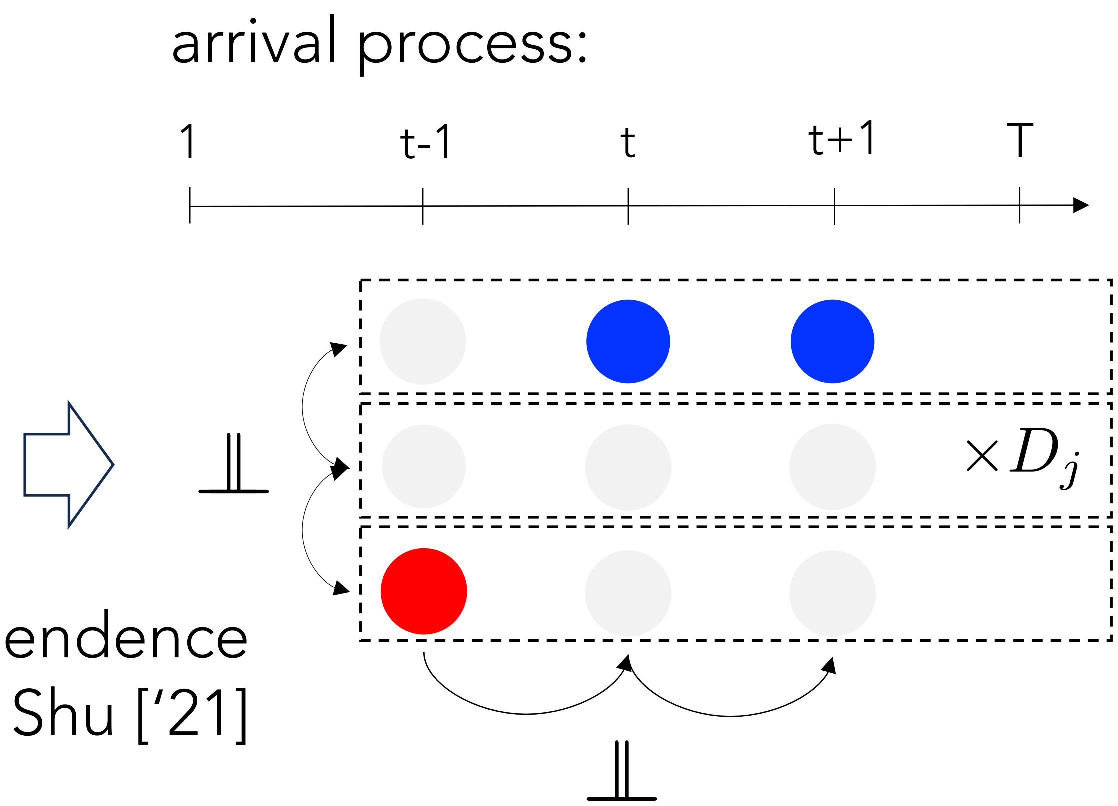
Order fulfilment data[1]




[1]Largest 200,000 SKUs, August-December 2014, bi-weekly aggregation

# Classic offline/online model

$$D_j \sim PoissonB\left(p_{1,j}, p_{2,j}, \ldots\right)$$

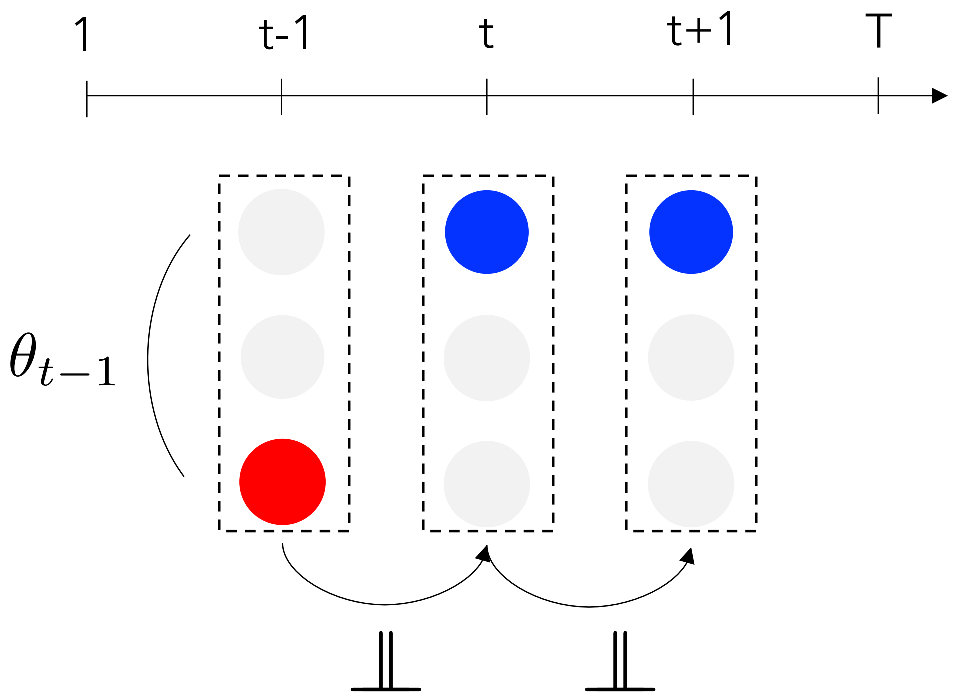



≈ cross-sectional independence
e.g., Huang and Shu ['21]

# Outline

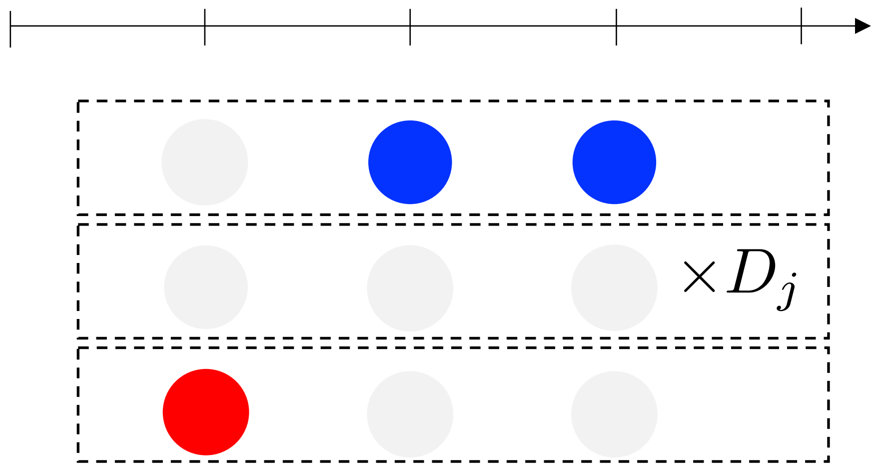
1. Nonparametric models with correlated arrivals
2. New matching algorithms with optimal competitive/approximation ratios

# Nonparametric models




- serial independence assumption
- modelling the arrival process (t)

## Nonparametric models

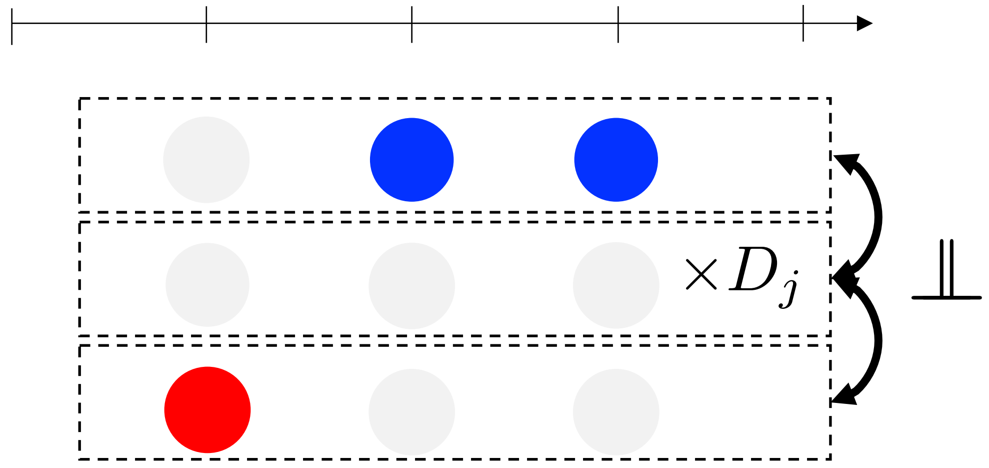



1. demand distribution $\boldsymbol{D}$
2. arrival ordering: ADV, RAND

# Nonparametric models

**INDEP model**

- Each type-demand $D_j$ follows an arbitrary (known) distribution
- But type-demands are independent $D_j \perp\!\!\!\perp D_k$



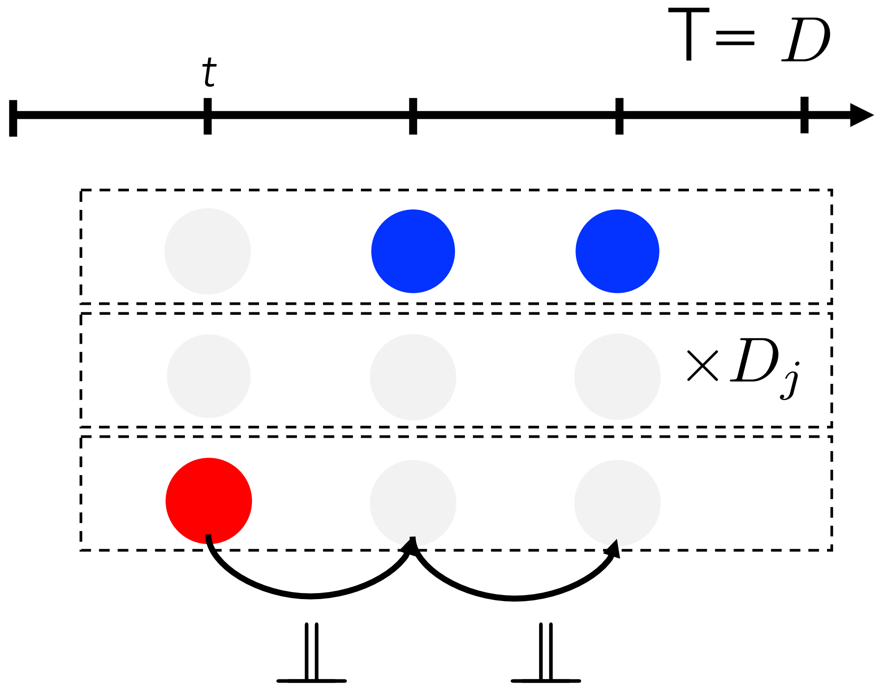

E.g., independent regions

# Nonparametric models

**CORREL model**

- The total demand $D = \sum_j D_j$ follows an arbitrary (known) distribution
- Conditional on T= D, the $t$-th query type independently sampled from $\boldsymbol{p}_t$

E.g., common shock across regions




conditional on D!

# Outline

1. Nonparametric models with correlated arrivals

2. New matching algorithms with optimal competitive/approx. ratios
   - Tighter polyhedral relaxations (≠ fluid relaxation)
   - Lossless rounding scheme

# LP benchmark for INDEP

$$
\begin{aligned}
\mathrm{LP}^{\mathrm{fluid}} = \max_{x} \quad & \sum_{i,j} r_{i,j} x_{i,j} \\
s.t. \quad & \sum_{i} x_{i,j} \leq k_i \\
& \sum_{j} x_{i,j} \leq \mathbb{E}[D_j] \\
& x_{i,j} \geq 0
\end{aligned}
$$

# LP benchmark for INDEP

$$
\begin{aligned}
\mathrm{LP}^{\text{fluid}} = \max_{x} \quad & \sum_{i,j} r_{i,j} x_{i,j} \\
s.t. \quad & \sum_{i} x_{i,j} \leq k_i \\
& \sum_{j} x_{i,j} \leq \mathbb{E}[D_j] \\
& x_{i,j} \geq 0
\end{aligned}
$$

$\Rightarrow$

$$
\begin{aligned}
\mathrm{LP}^{\text{trunc}} = \max_{x} \quad & \sum_{i,j} r_{i,j} x_{i,j} \\
s.t. \quad & \sum_{j} x_{i,j} \leq k_i \quad \forall i \\
& \sum_{i \in S} x_{i,j} \leq \mathbb{E}\left[\min\left\{\sum_{i \in S} k_i, D_j\right\}\right] \quad \forall j, S \subseteq [n] \\
& x_{i,j} \geq 0
\end{aligned}
$$

# LP benchmark for INDEP

$$\mathrm{LP}^{\mathrm{fluid}} = \max_x \sum_{i,j} r_{i,j} x_{i,j}$$
$$s.t. \quad \sum_i x_{i,j} \le k_i$$
$$\sum_j x_{i,j} \le \mathbb{E}[D_j]$$
$$x_{i,j} \ge 0$$

$$\mathrm{LP}^{\mathrm{trunc}} = \max_x \sum_{i,j} r_{i,j} x_{i,j}$$
$$s.t. \quad \sum_j x_{i,j} \le k_i \quad \forall i$$
$$\sum_{i \in S} x_{i,j} \le \mathbb{E}\left[\min\left\{\sum_{i\in S} k_i, D_j\right\}\right] \quad \forall j, S \subseteq [n]$$
$$x_{i,j} \ge 0$$

S: subset of resources;
Hall's marriage condition
and taking expectations

# LP benchmark for INDEP

$$\mathrm{LP}^{\mathrm{fluid}} = \max_x \sum_{i,j} r_{i,j} x_{i,j}$$
$$s.t. \quad \sum_i x_{i,j} \le k_i$$
$$\sum_j x_{i,j} \le \mathbb{E}[D_j]$$
$$x_{i,j} \ge 0$$

$$\mathrm{LP}^{\mathrm{trunc}} = \max_x \sum_{i,j} r_{i,j} x_{i,j}$$
$$s.t. \quad \sum_j x_{i,j} \le k_i \quad \forall i$$
$$\sum_{i \in S} x_{i,j} \le \mathbb{E}\left[\min\left\{\sum_{i \in S} k_i, D_j\right\}\right] \quad \forall j, S \subseteq [n]$$
$$x_{i,j} \ge 0$$

S: subset of resources;
Hall's marriage condition
and taking expectations

# LP benchmark for INDEP

$$\mathrm{LP}^{\mathrm{fluid}} = \max_x \sum_{i,j} r_{i,j} x_{i,j}$$
$$s.t. \quad \sum_i x_{i,j} \le k_i$$
$$\sum_j x_{i,j} \le \mathbb{E}[D_j]$$
$$x_{i,j} \ge 0$$

$$\mathrm{LP}^{\mathrm{trunc}} = \max_x \sum_{i,j} r_{i,j} x_{i,j}$$
$$s.t. \quad \sum_j x_{i,j} \le k_i \quad \forall i$$
$$\sum_{i \in S} x_{i,j} \le \mathbb{E}\left[\min\left\{\sum_{i \in S} k_i, D_j\right\}\right] \quad \forall j, S \subseteq [n]$$
$$x_{i,j} \ge 0$$

S: subset of resources; Hall's marriage condition and taking expectations

**Proposition [A., Ma '22]:** Valid benchmark $\mathrm{LP}^{\mathrm{trunc}} \ge \mathrm{OFF}$

**Proposition [A., Ma '22]:** $\mathrm{LP}^{\mathrm{trunc}}$ is solvable in polynomial time (polymatroid constraints).

# Main results – INDEP

**Theorem [A., Ma, Zhang '23]:** For INDEP, there is no matching policy better than 1/2-competitive even under large inventory and uniform arrivals.

# Main results – INDEP

**Theorem [A., Ma, Zhang '23]:** For INDEP, there is no matching policy better than 1/2-competitive even under large inventory and uniform arrivals.

**Theorem [A., Ma '23]:** For INDEP, there exists a 1/2-competitive matching policy (against our LP) that is computed in polynomial time, even under adversarial arrivals.

# Main results – INDEP

**Theorem [A., Ma, Zhang '23]:** For INDEP, there is no matching policy better than 1/2-competitive even under large inventory and uniform arrivals.

**Theorem [A., Ma '23]:** For INDEP, there exists a 1/2-competitive matching policy (against our LP) that is computed in polynomial time, even under adversarial arrivals.

*Proof idea:*
Reduction to single-offline node prophet inequality via lossless rounding

# The central rounding lemma

$$\mathrm{LP}^{\mathrm{fluid}} = \max_x \sum_t \sum_{i,j} r_{i,j} x^t_{i,j}$$
$$s.t. \quad \sum_i \sum_t x^t_{i,j} \le k_i$$
$$\sum_j x^t_{i,j} \le \theta_{t,j}$$

$$\mathrm{LP}^{\mathrm{trunc}} = \max_x \sum_{i,j} r_{i,j} x_{i,j}$$
$$s.t. \quad \sum_j x_{i,j} \le k_i \quad \forall i$$
$$\sum_{i \in S} x_{i,j} \le \mathbb{E}\left[\min\left\{\sum_{i \in S} k_i, D_j\right\}\right] \quad \forall j, S \subseteq [n]$$
$$x_{i,j} \ge 0$$

# The central rounding lemma

$$\mathrm{LP}^{\mathrm{fluid}} = \max_x \sum_t \sum_{i,j} r_{i,j} x^t_{i,j}$$
$$s.t. \quad \sum_i \sum_t x^t_{i,j} \leq k_i$$
$$\sum_j x^t_{i,j} \leq \theta_{t,j}$$

$$\mathrm{LP}^{\mathrm{trunc}} = \max_x \sum_{i,j} r_{i,j} x_{i,j}$$
$$s.t. \quad \sum_j x_{i,j} \leq k_i \quad \forall i$$
$$\sum_{i \in S} x_{i,j} \leq \mathbb{E}\left[\min\left\{\sum_{i\in S} k_i, D_j\right\}\right] \quad \forall j, S \subseteq [n]$$
$$x_{i,j} \geq 0$$

**Lemma [A., Ma '22]:** There exists a lossless rounding for $\mathrm{LP}^{\mathrm{trunc}}$ for each type j

$$\Pr_{\pi\sim\lambda_j}\left[\sum_\ell \mathbb{I}[\pi(\ell) = i] \cdot \Pr\left[D_j \geq \ell\right]\right] = x^*_{i,j}$$

# The central rounding lemma

$\mathrm{LP}^{\mathrm{fluid}} = \max_{x}$ … $\mathrm{LP}^{\mathrm{trunc}}$ …

$s.t.$ … $\forall i$

… $\sum_{i \in S} k_i, D_j \}\Big]$ $\quad \forall j, S \subseteq [n]$

$x_{i,j} \geq 0$

For any fixed type demand distribution, the LP fractional solutions represents all achievable match rates to the *m* resources

**Lemma [A., Ma '22]:** There exists a lossless rounding for $\mathrm{LP}^{\mathrm{trunc}}$ for each type j:

$$\Pr_{\pi \sim \lambda_j}\left[\sum_{\ell} \mathbb{I}[\pi(\ell) = i] \cdot \Pr\left[D_j \geq \ell\right]\right] = x^*_{i,j}$$

# Rounding lemma

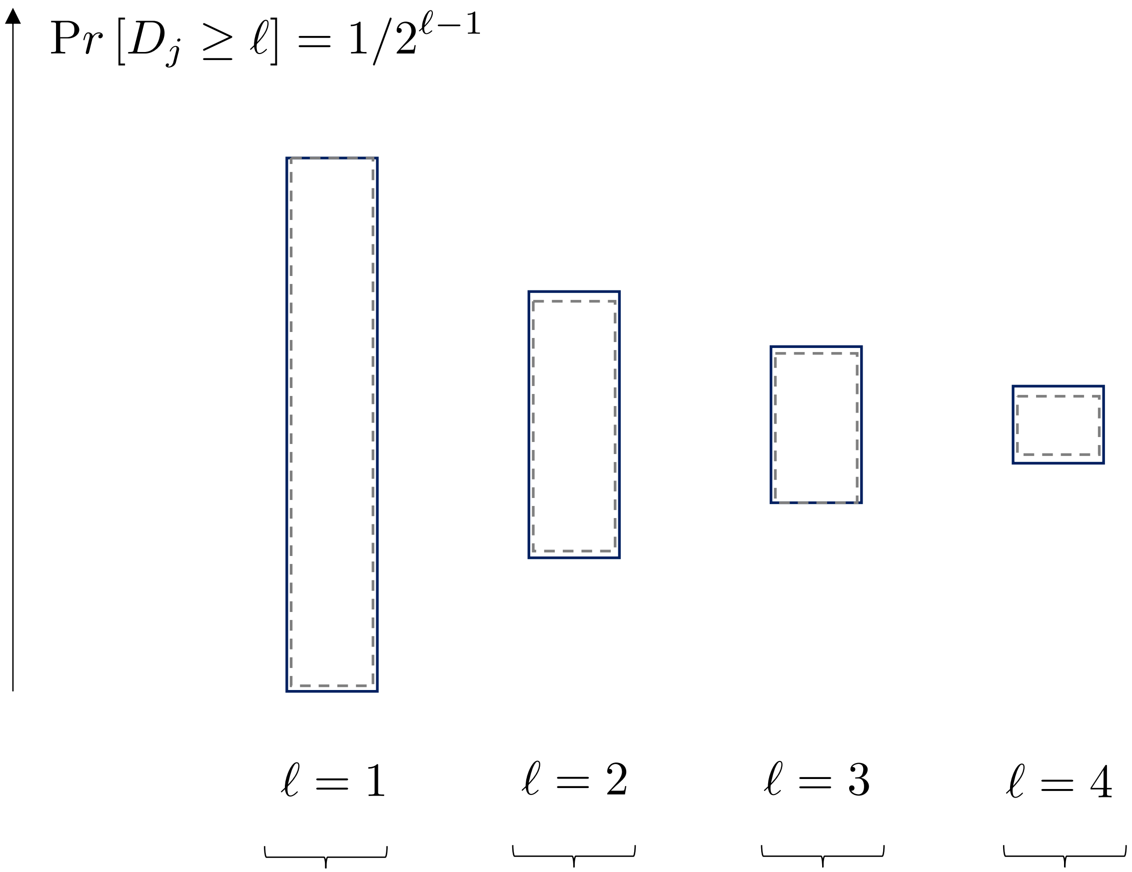
- E.g., nonparametric demand $\Pr\left[D_j \geq \ell\right] = 1/2^{\ell-1}$ with $\ell = 1, \ldots, 4$
- Feasible fractional matching: $x_j^* = \left(\frac{1}{8}, \frac{3}{8}, \frac{7}{8} + \epsilon\right)$
- Binding constraints:

$$\frac{7}{8} + \epsilon \leq \mathrm{E}[\min\{D_j, 1\}] = 1$$

$$\frac{11}{8} + \epsilon \leq \mathrm{E}[\min\{D_j, 2\}] = \frac{3}{2}$$

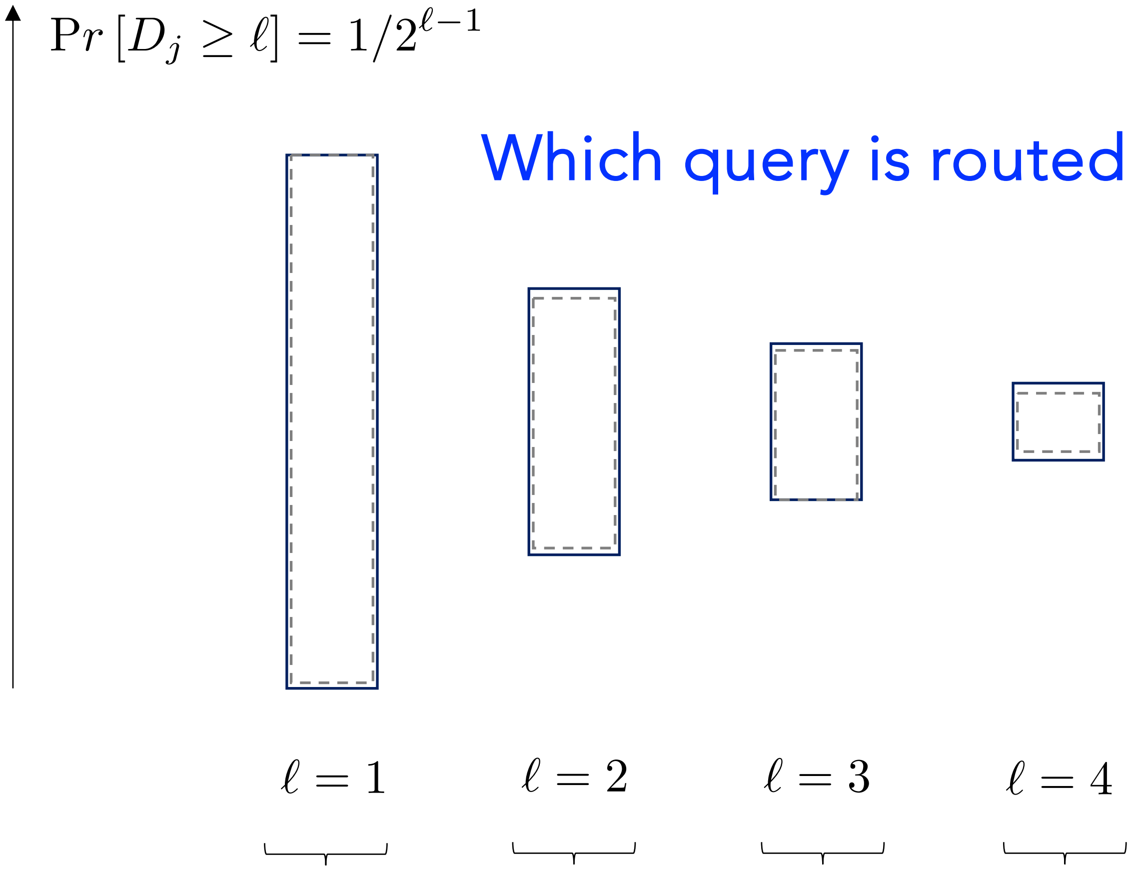
$$\frac{13}{8} + \epsilon \leq \mathrm{E}[\min\{D_j, 3\}] = \frac{7}{4}$$

# Rounding lemma




$$x_j^* = \left(\frac{1}{8}, \frac{3}{8}, \frac{7}{8} + \epsilon\right)$$

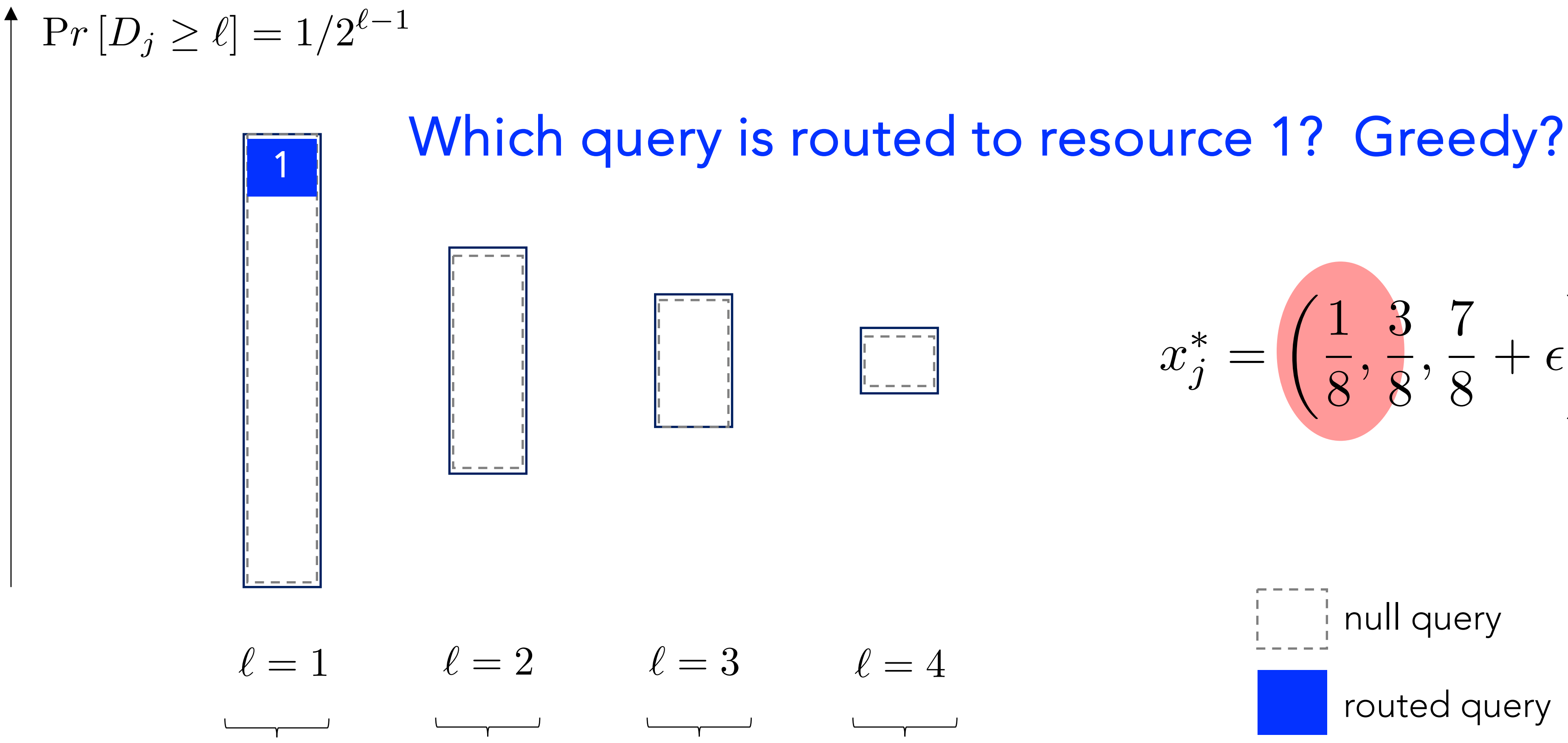
# Rounding lemma




Which query is routed to resource 1?

$$x_j^* = \left(\frac{1}{8}, \frac{3}{8}, \frac{7}{8} + \epsilon\right)$$

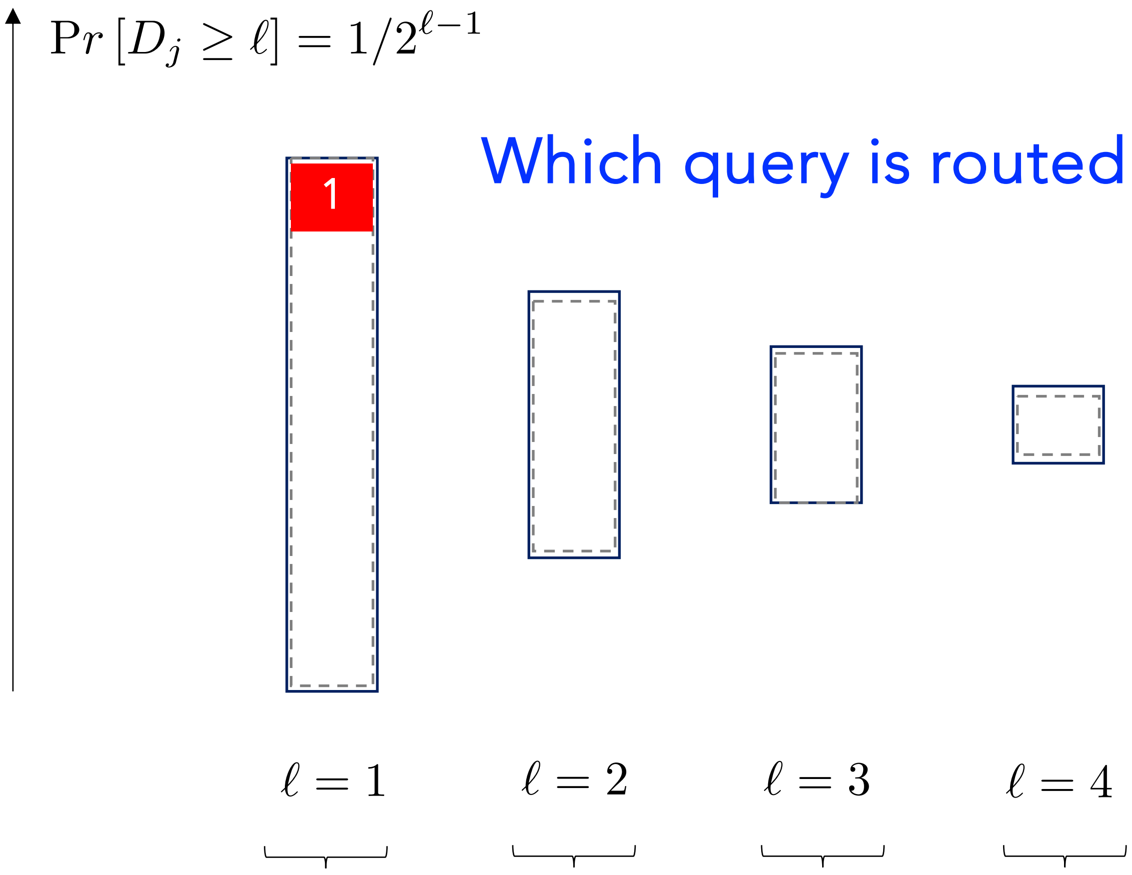
# Rounding lemma




Which query is routed to resource 1? Greedy?

$$x_j^* = \left(\frac{1}{8}, \frac{3}{8}, \frac{7}{8} + \epsilon\right)$$

# Rounding lemma



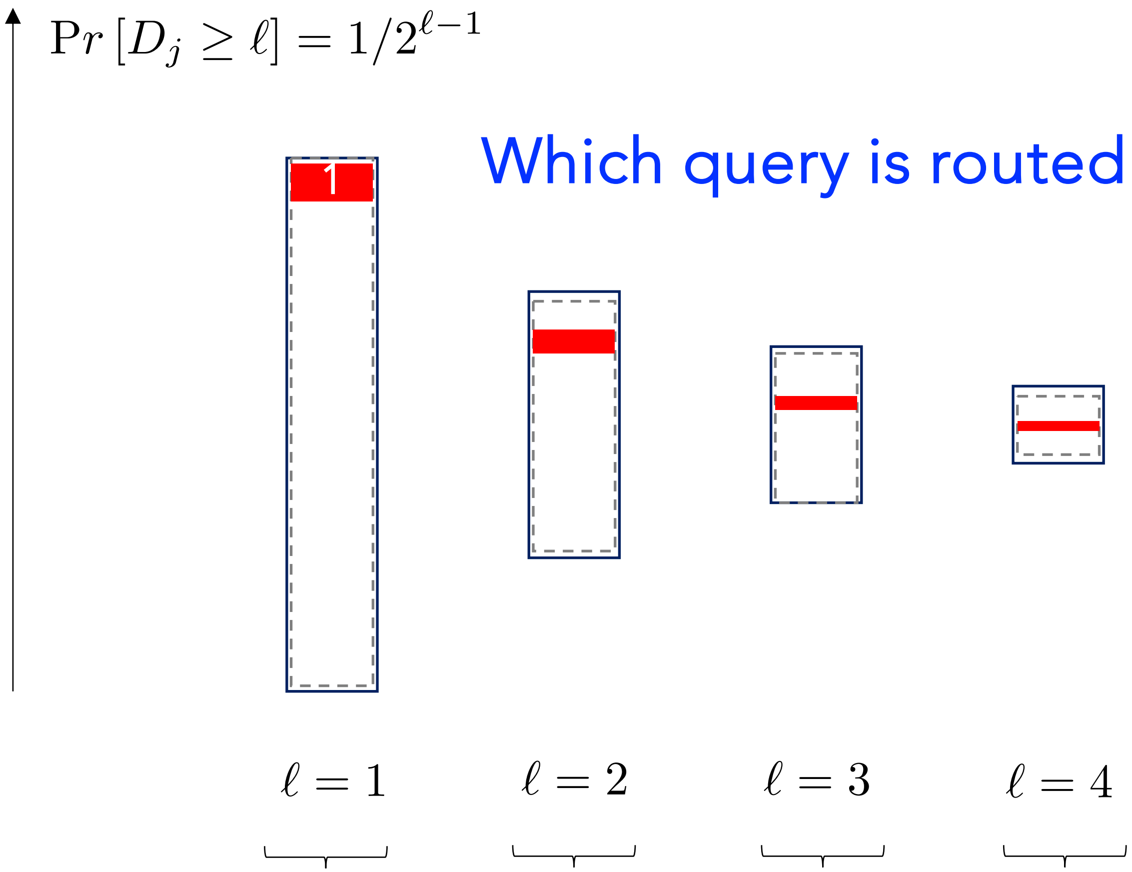

Which query is routed to resource 1? Greedy? ☹

$$x_j^* = \left(\frac{1}{8}, \frac{3}{8}, \frac{7}{8} + \epsilon\right)$$

null query

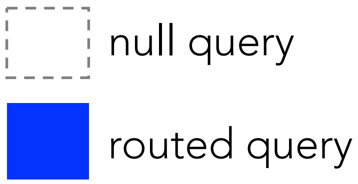
routed query

# Rounding lemma

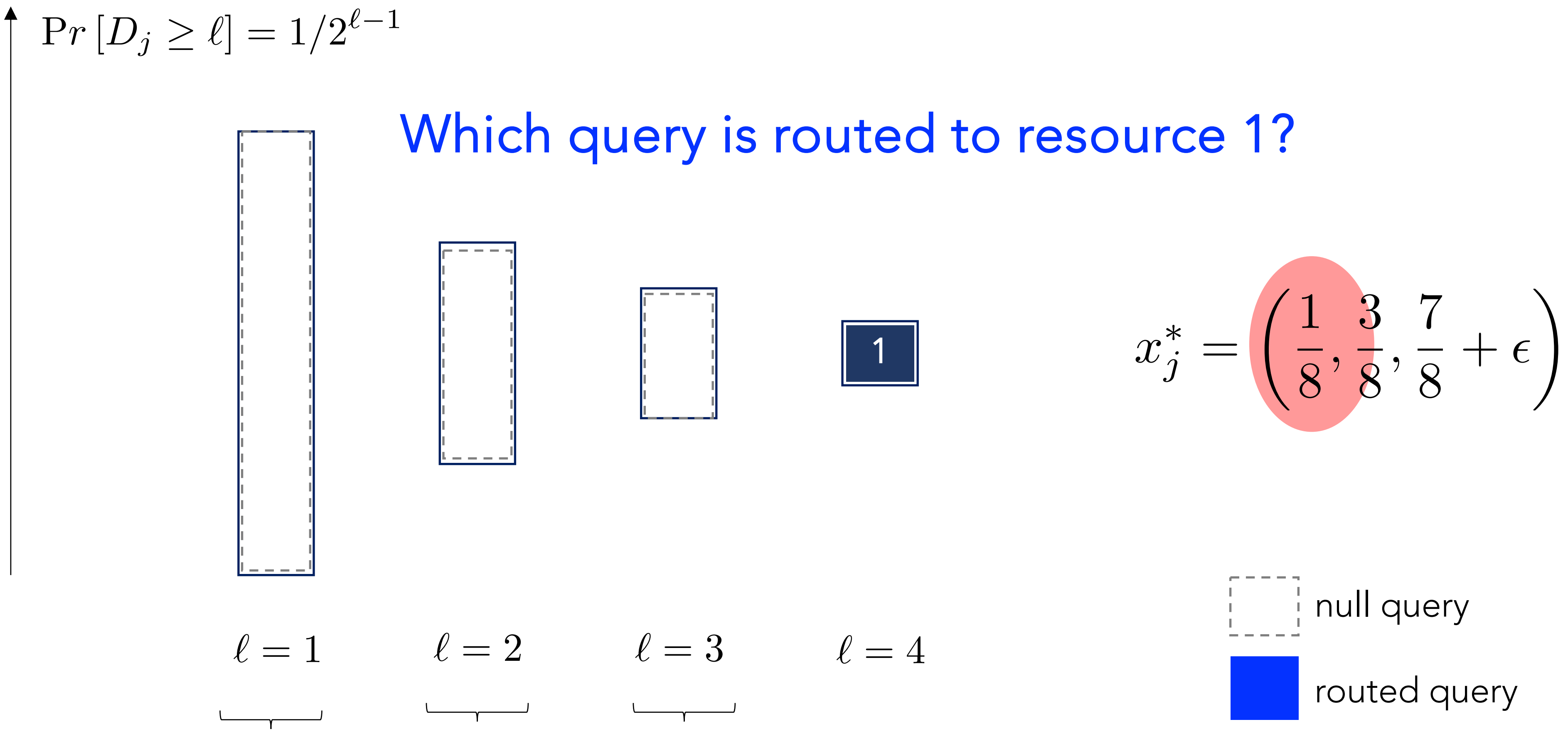



Which query is routed to resource 1? Proportional? ☹

$$x_j^* = \left(\frac{1}{8}, \frac{3}{8}, \frac{7}{8} + \epsilon\right)$$

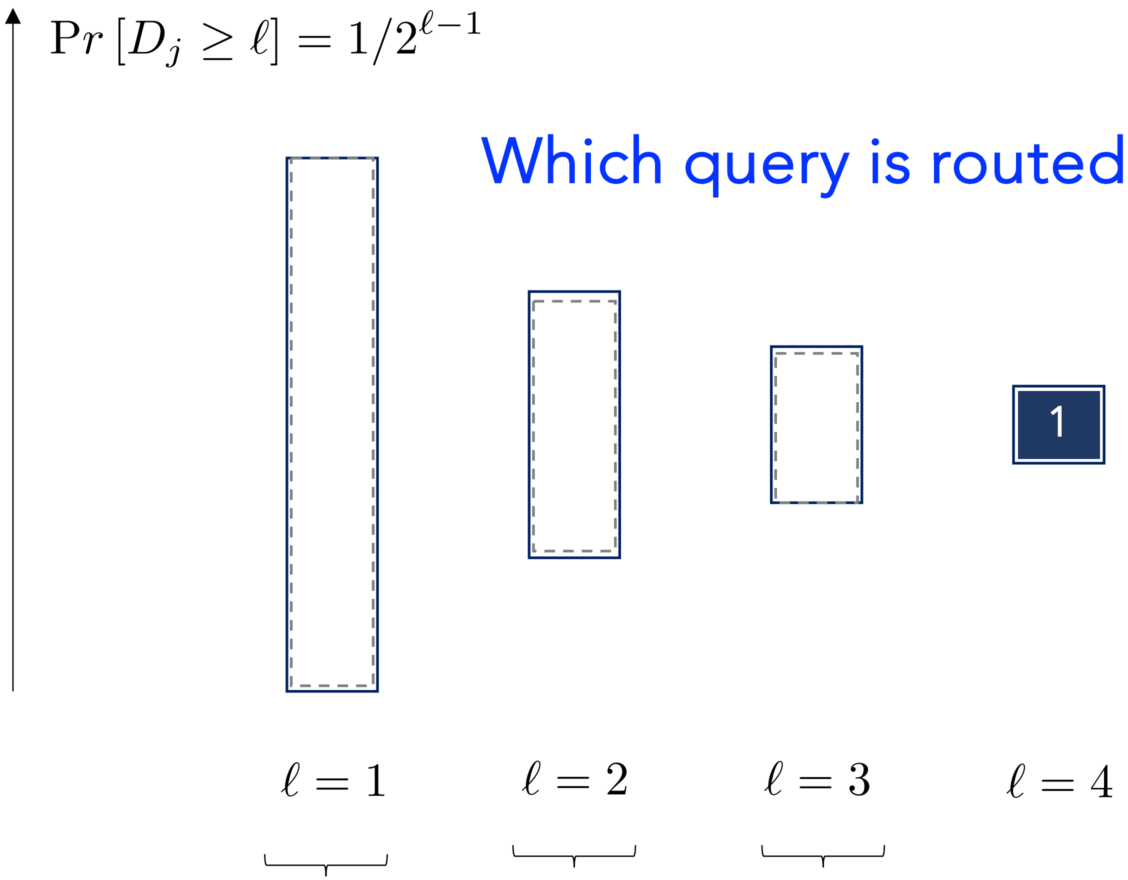


# Rounding lemma

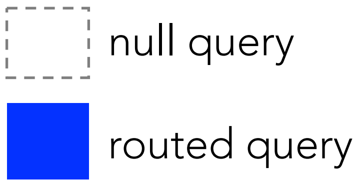


# Rounding lemma

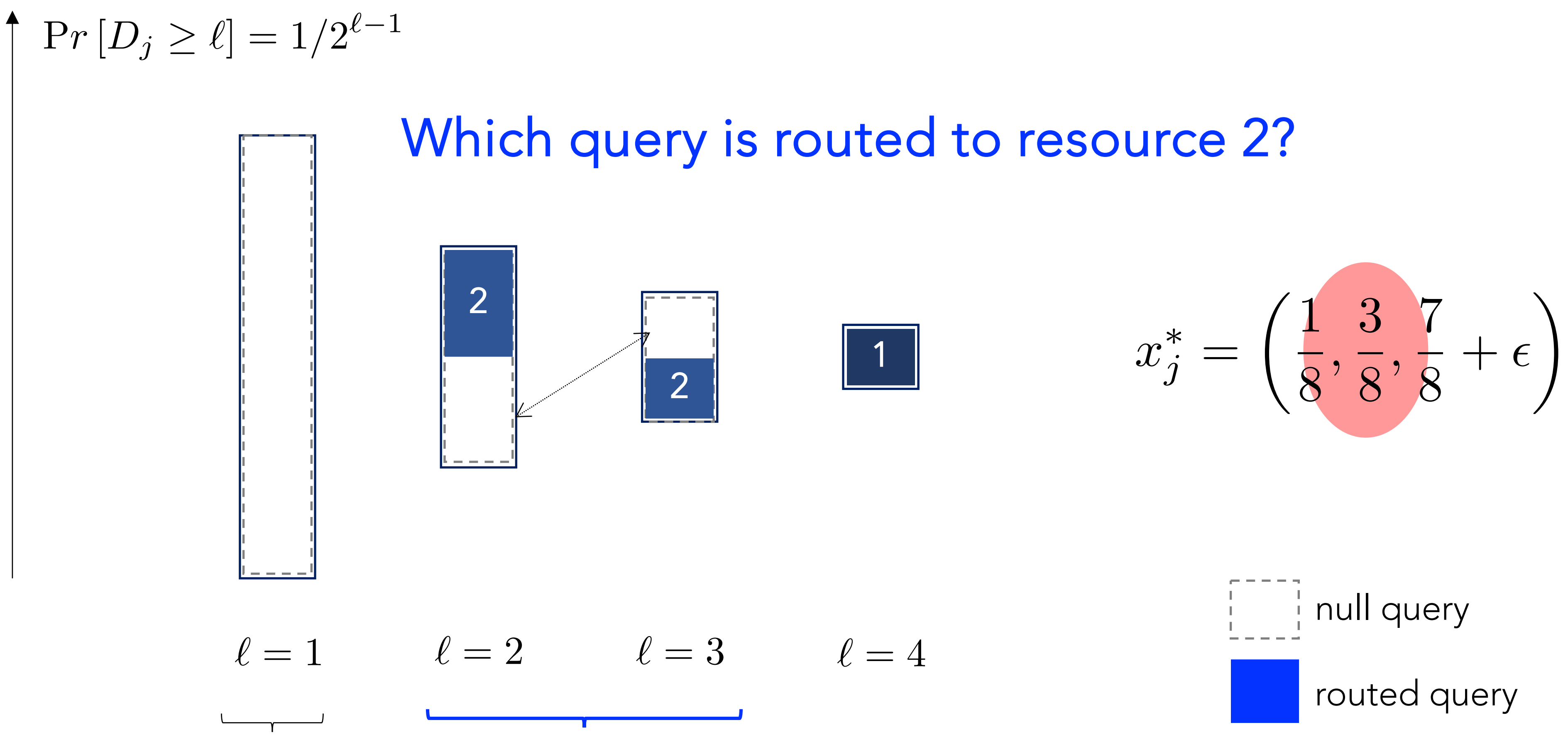
**Which query is routed to resource 2?**



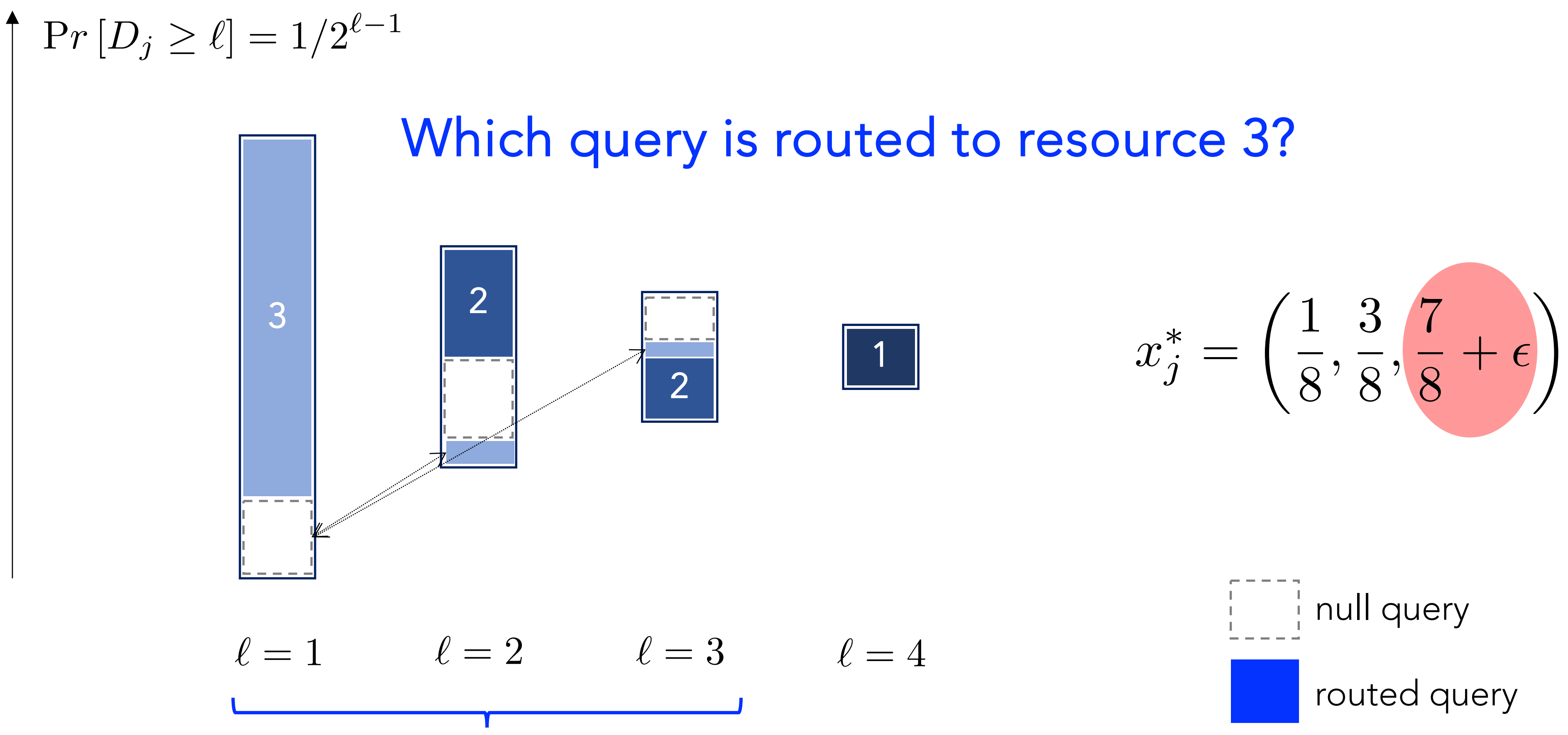

$$x_j^* = \left(\frac{1}{8}, \frac{3}{8}, \frac{7}{8} + \epsilon\right)$$



# Rounding lemma



# Rounding lemma



# Concluding remarks

- Common principles
  - Limitations of fluid relaxation for more rich stochastic matching problems
  - Tighter LP relaxations: more closely approximating the online/offline optimum
  - "Attainability" results: contention resolution or correlated roundings

- Open questions & future directions
  - Breaching (1-1/e)-approximation for dynamic matching
  - Sample complexity of nonparametric stochastic models
  - Other models of correlation: e.g., prediction uncertainty

# Main results – CORREL

**Observation [A., Ma '22]:** For CORREL, no constant-factor competitive ratio is achievable.

**Theorem [A., Ma '22]:** For CORREL, there exists an approximate matching policy that achieving an approximation ratio $\gamma_k^* > (1+\sqrt{k})^{-1}$, where $\gamma_k^*$ is the best-known competitive ratio for k-unit prophet inequality.

*Proof ideas:*

- Conditional LP: valid inequalities conditional on the largest arrival sequence length
- Reduction to online contention resolution scheme [Jiang, Ma, and Zhang, 22]

# Conditional LP

$$\mathrm{LP}^{\mathrm{fluid}} = \max_x \sum_t \sum_{i,j} r_{i,j} x_{i,j}^t$$
$$s.t. \quad \sum_i \sum_t x_{i,j}^t \leq k_i$$
$$\sum_j x_{i,j}^t \leq \theta_{t,j}$$

$\Rightarrow$

$$\mathrm{LP}^{\mathrm{cond}} = \max_x \sum_t \sum_{i,j} r_{i,j} x_{i,j}^t$$
$$s.t. \quad \sum_i \sum_t \frac{1}{\Pr[D \geq t]} \cdot x_{i,j}^t \leq k_i$$
$$\sum_j \frac{1}{\Pr[D \geq t]} \cdot x_{i,j}^t \leq \theta_{t,j}$$

Intuition: "the tightest constraints are given by the largest possible demand"