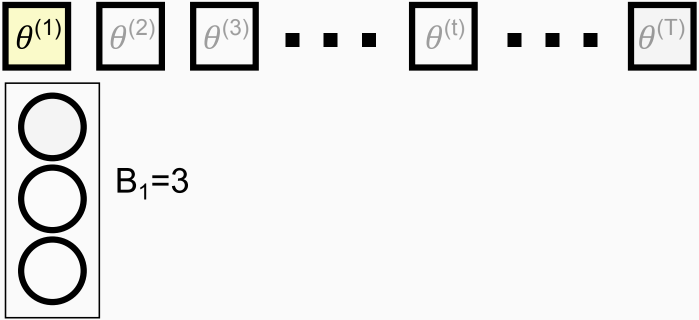
Allocating Resources, in the Future

Sid Banerjee
School of ORIE

May 3, 2018

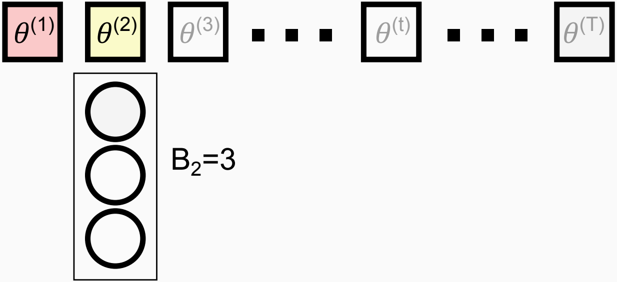
Simons Workshop on Mathematical and Computational Challenges in Real-Time Decision Making

online resource allocation: basic model



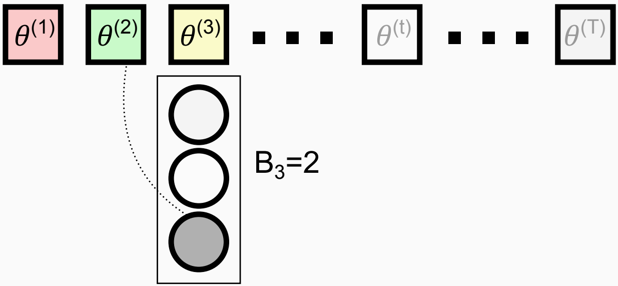
- single resource, initial capacity B; T agents arrive sequentially
- agent t has type $\theta^{(t)}$ = reward earned if agent is allocated

online resource allocation: basic model



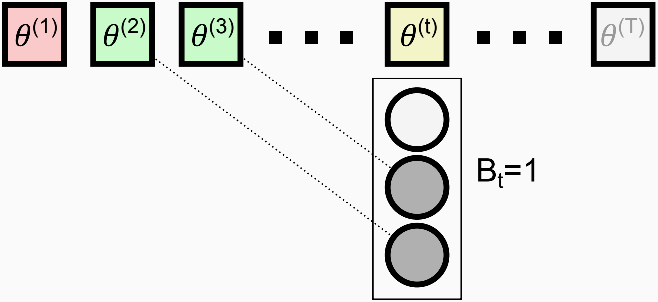
- single resource, initial capacity B; T agents arrive sequentially
- agent t has type $\theta^{(t)}$ = reward earned if agent is allocated
- principle makes irrevocable decisions

online resource allocation: basic model



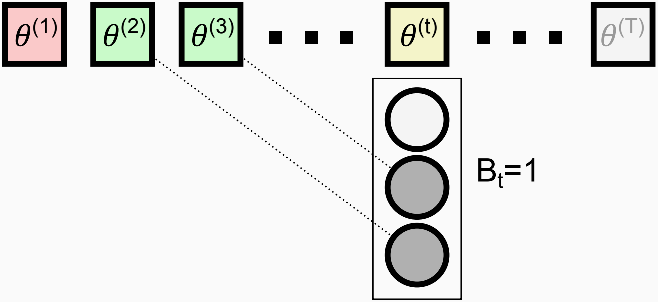
- single resource, initial capacity B; T agents arrive sequentially
- agent t has type $\theta^{(t)}$ = reward earned if agent is allocated
- principle makes irrevocable decisions; resource is non-replenishable

online resource allocation: basic model



- single resource, initial capacity B; T agents arrive sequentially
- agent t has type $\theta^{(t)}$ = reward earned if agent is allocated
- principle makes irrevocable decisions; resource is non-replenishable
- assumptions on agent types $\{\theta_t\}$
  - finite set of values $\{v_i\}_{i=1}^n$ (e.g. $\theta^{(t)} = v_i$ with prob p_i i.i.d.)
  - in general: arrivals can be time varying, correlated

online resource allocation: basic model

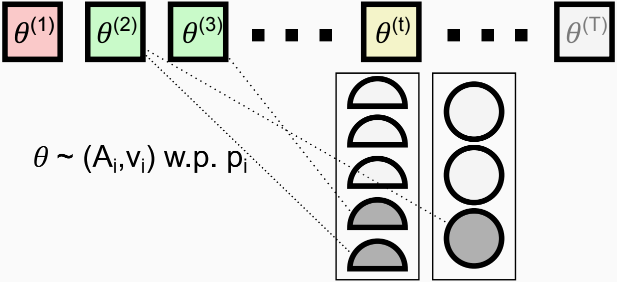


- single resource, initial capacity B; T agents arrive sequentially
- agent t has type $\theta^{(t)}$ = reward earned if agent is allocated
- principle makes irrevocable decisions; resource is non-replenishable
- assumptions on agent types $\{\theta_t\}$
  - finite set of values $\{v_i\}_{i=1}^n$ (e.g. $\theta^{(t)} = v_i$ with prob p_i i.i.d.)
  - in general: arrivals can be time varying, correlated

online resource allocation problem

allocate resources to maximize sum of rewards

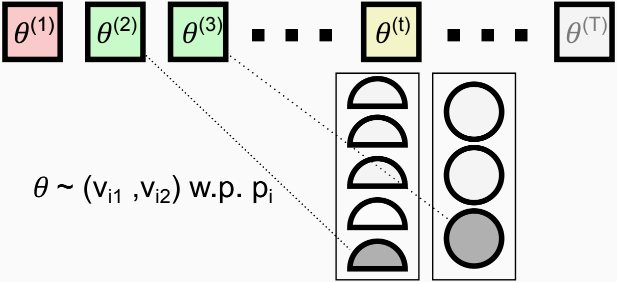
online resource allocation: first generalization



- d resources, initial capacities $(B^1, B^2, \ldots, B^d)$
- T agents; each has type $\theta_i = (A_i, v_i)$
  - $A_i \in \{0,1\}^d$: resource requirement, v_i : value
  - agent has type θ_i with prob p_i

also known as: network revenue management; single-minded buyer

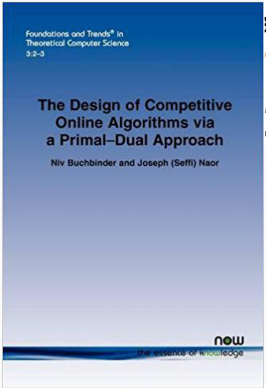
online resource allocation: second generalization

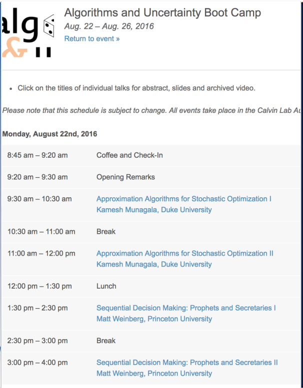


- d resources, initial capacities $(B^1, B^2, \ldots, B^d)$
- T agents arrive sequentially
- each has type $\theta = (v_{i1}, v_{i2}, \ldots, v_{id})$, wants single resource

also known as: online weighted matching; unit-demand buyer

online allocation across fields





Algorithms and Uncertainty Boot Camp

Aug. 22 – Aug. 26, 2016

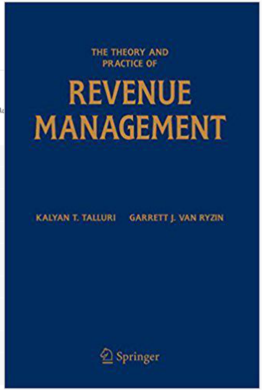
Return to event »

- Click on the titles of individual talks for abstract, slides and archived video.

Please note that this schedule is subject to change. All events take place in the Calvin Lab Au

Monday, August 22nd, 2016

8:45 am – 9:20 am	Coffee and Check-In
9:20 am – 9:30 am	Opening Remarks
9:30 am – 10:30 am	Approximation Algorithms for Stochastic Optimization I Kamesh Munagala, Duke University
10:30 am – 11:00 am	Break
11:00 am – 12:00 pm	Approximation Algorithms for Stochastic Optimization II Kamesh Munagala, Duke University
12:00 pm – 1:30 pm	Lunch
1:30 pm – 2:30 pm	Sequential Decision Making: Prophets and Secretaries I Matt Weinberg, Princeton University
2:30 pm – 3:00 pm	Break
3:00 pm – 4:00 pm	Sequential Decision Making: Prophets and Secretaries II Matt Weinberg, Princeton University



- related problems studied in Markov decision processes, online algorithms, prophet inequalities, revenue management, etc.
- informational variants:
  distributional knowledge $\prec$ bandit settings $\prec$ adversarial inputs

the ‘deep’ learning revolution

vast improvements in machine learning for data-driven prediction

axiomatizing the zeitgeist

the deep learning revolution

vast improvements in machine learning for data-driven prediction

- axiom: have access to black-box predictive algorithms

axiomatizing the zeitgeist

the deep learning revolution

vast improvements in machine learning for data-driven prediction

- axiom: have access to black-box predictive algorithms

core question of this talk

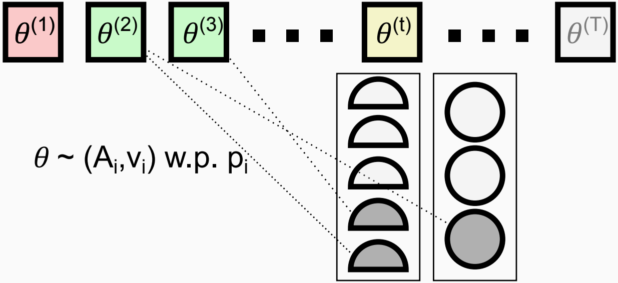
how does having such an oracle affect online resource allocation?

- TL;DR - new online allocation policies with strong regret bounds
- re-examining old questions leads to surprising new insights

bridging online allocation and predictive models

The Bayesian Prophet: A Low-Regret Framework for Online Decision Making
Alberto Vera & S.B. (2018)
https://ssrn.com/abstract_id=3158062

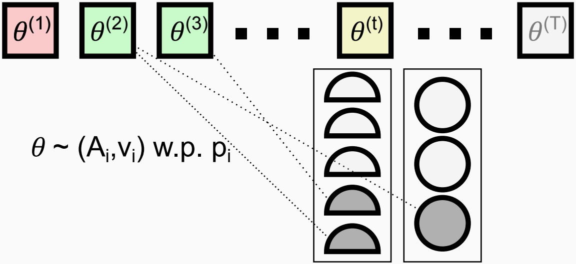
focus of talk: allocation with single-minded agents



- d resources, initial capacities $(B^1, B^2, \ldots, B^d)$
- T agents arrive sequentially; each has type $\theta = (A, v)$
  - $A =$ resource requirement, $v =$ value
  - agent has type θ_i with prob p_i, i.i.d.

online allocation problem

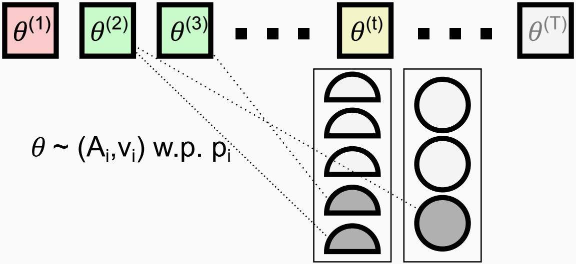
allocate resources to maximize sum of rewards



optimal policy

can be computed via dynamic programming

- requires exact distributional knowledge
- ‘curse of dimensionality’: $|\text{state-space}| = T \times B_1 \times \ldots \times B_d$
- does not quantify cost of uncertainty



optimal policy

can be computed via dynamic programming

- requires exact distributional knowledge
- 'curse of dimensionality': $|\text{state-space}| = T \times B_1 \times \ldots \times B_d$
- does not quantify cost of uncertainty

'prophet' benchmark

V^{off}: OFFLINE optimal policy; has full knowledge of $\{\theta_1, \theta_2, \ldots, \theta_T\}$

performance measure: regret

prophet benchmark: V^{off}

- `OFFLINE` knows entire type sequence $\{\theta_t | t = 1 \dots T\}$
- for the network revenue management setting, V^{off} given by

$$\begin{aligned} max. \sum_{i=1}^{n} x_i v_i \\ s.t. \sum_{i=1}^{n} A_i x_i \leq B \\ 0 \leq x_i \leq N_i[1:T] \end{aligned}$$

  – $N_i[1:T] \sim \#$ of arrivals of type $\theta_i = (A_i, v_i)$ over $\{1, 2, \dots, T\}$

regret

$$\mathbb{E}[Regret] = \mathbb{E}[V^{off} - V^{alg}]$$

online allocation with prediction oracle

given black-box predictive oracle about performance of `OFFLINE`
(specifically, for any t, B, have statistical info about $V^{off}[t, T]$)

online allocation with prediction oracle

given black-box predictive oracle about performance of `OFFLINE` (specifically, for any t, B, have statistical info about $V^{off}[t, T]$)

- let $\pi_t = \mathbb{P}\big[V^{off}[t, T]$ decreases if `OFFLINE` accepts t^{th} arrival$\big]$

online allocation with prediction oracle

given black-box predictive oracle about performance of `OFFLINE` (specifically, for any t, B, have statistical info about $V^{off}[t,T]$)

- let $\pi_t = \mathbb{P}\left[V^{off}[t,T] \text{ decreases if } \texttt{OFFLINE} \text{ accepts } t^{th} \text{ arrival}\right]$

Bayes selector

accept t^{th} arrival iff $\pi_t > 0.5$

online allocation with prediction oracle

given black-box predictive oracle about performance of OFFLINE (specifically, for any t, B, have statistical info about $V^{off}[t, T]$)

- let $\pi_t = \mathbb{P}\big[V^{off}[t, T]$ decreases if OFFLINE accepts t^{th} arrival$\big]$

Bayes selector

accept t^{th} arrival iff $\pi_t > 0.5$

theorem [Vera & B, 2018]

(under mild tail bounds on $N_i[t : T]$)

Bayes selector has $\mathbb{E}[Regret]$ independent of $T, B_1, B_2, \ldots, B_d$

online allocation with prediction oracle

given black-box predictive oracle about performance of OFFLINE (specifically, for any t, B, have statistical info about $V^{off}[t, T]$)

- let $\pi_t = \mathbb{P}\big[V^{off}[t, T]$ decreases if OFFLINE accepts t^{th} arrival$\big]$

Bayes selector

accept t^{th} arrival iff $\pi_t > 0.5$

theorem [Vera & B, 2018]

(under mild tail bounds on $N_i[t : T]$)

Bayes selector has $\mathbb{E}[Regret]$ independent of $T, B_1, B_2, \ldots, B_d$

- arrivals can be time-varying, correlated; discounted rewards
- works for general settings (single-minded, unit-demand, etc.)
- can use approx oracle (e.g., from samples)

standard approach: randomized admission control (RAC)

offline optimum V^{off}

$$\begin{aligned} max. \sum_{i=1}^{n} x_i v_i \\ s.t. \sum_{i=1}^{n} A_i x_i \leq B \\ 0 \leq x_i \leq N_i[1:T] \end{aligned}$$

standard approach: randomized admission control (RAC)

offline optimum V^{off}

$$\begin{aligned} max.\ & \sum_{i=1}^{n} x_i v_i \\ s.t. & \sum_{i=1}^{n} A_i x_i \le B \\ & 0 \le x_i \le N_i[1:T] \end{aligned}$$

(upfront) fluid LP V^{fl}

$$\begin{aligned} max.\ & \sum_{i=1}^{n} x_i v_i \\ s.t. & \sum_{i=1}^{n} A_i x_i \le B \\ & 0 \le x_i \le \mathbb{E}[N_i[1:T]] = Tp_i \end{aligned}$$

- $\mathbb{E}[V^{off}] \le V^{fl}$ (via Jensen's, concavity of V^{off} w.r.t. N_i)
- fluid RAC: accept type θ_i with prob $\frac{x_i}{Tp_i}$

standard approach: randomized admission control (RAC)

offline optimum V^{off}

$$\begin{aligned} max. & \sum_{i=1}^{n} x_i v_i \\ s.t. & \sum_{i=1}^{n} A_i x_i \leq B \\ & 0 \leq x_i \leq N_i[1:T] \end{aligned}$$

(upfront) fluid LP V^{fl}

$$\begin{aligned} max. & \sum_{i=1}^{n} x_i v_i \\ s.t. & \sum_{i=1}^{n} A_i x_i \leq B \\ & 0 \leq x_i \leq \mathbb{E}[N_i[1:T]] = T p_i \end{aligned}$$

- $\mathbb{E}[V^{off}] \leq V^{fl}$ (via Jensen's, concavity of V^{off} w.r.t. N_i)
- fluid RAC: accept type θ_i with prob $\frac{x_i}{T p_i}$

proposition

fluid RAC has $\mathbb{E}[Regret] = \Theta(\sqrt{T})$

- **[Gallego & van Ryzin'97], [Maglaras & Meissner'06]**
- N.B. this is a static policy!

RAC with re-solving

offline optimum V^{off}

$$\begin{aligned} max. \sum_{i=1}^{n} x_i v_i \\ s.t. \sum_{i=1}^{n} A_i x_i \le B \\ 0 \le x_i \le N_i \end{aligned}$$

re-solved fluid LP $V^{fl}(t)$:

$$\begin{aligned} max. \sum_{i=1}^{n} x_i[t] v_i \\ s.t. \sum_{i=1}^{n} A_i x_i[t] \le B[t] \\ 0 \le x_i[t] \le \mathbb{E}[N_i[t:T]] = (T-t)p_i \end{aligned}$$

AC with re-solving: at time t, accept type θ_i with prob $\frac{x_i[t]}{(T-t)p_i}$

RAC with re-solving

offline optimum V^{off}

$$\begin{aligned} max. \ & \sum_{i=1}^{n} x_i v_i \\ s.t. & \sum_{i=1}^{n} A_i x_i \leq B \\ & 0 \leq x_i \leq N_i \end{aligned}$$

re-solved fluid LP $V^{fl}(t)$:

$$\begin{aligned} max. \ & \sum_{i=1}^{n} x_i[t] v_i \\ s.t. & \sum_{i=1}^{n} A_i x_i[t] \leq B[t] \\ & 0 \leq x_i[t] \leq \mathbb{E}[N_i[t:T]] = (T-t)p_i \end{aligned}$$

AC with re-solving: at time t, accept type θ_i with prob $\frac{x_i[t]}{(T-t)p_i}$

- regret improves to $o(\sqrt{T})$ **[Reiman & Wang'08]**
- $O(1)$ regret under (dual) non-degeneracy **[Jasin & Kumar'12]**

RAC with re-solving

offline optimum V^{off}

$$\begin{aligned} max. \sum_{i=1}^{n} x_i v_i \\ s.t. \sum_{i=1}^{n} A_i x_i \leq B \\ 0 \leq x_i \leq N_i \end{aligned}$$

re-solved fluid LP $V^{fl}(t)$:

$$\begin{aligned} max. \sum_{i=1}^{n} x_i[t] v_i \\ s.t. \sum_{i=1}^{n} A_i x_i[t] \leq B[t] \\ 0 \leq x_i[t] \leq \mathbb{E}[N_i[t:T]] = (T-t)p_i \end{aligned}$$

AC with re-solving: at time t, accept type θ_i with prob $\frac{x_i[t]}{(T-t)p_i}$

- regret improves to $o(\sqrt{T})$ **[Reiman & Wang'08]**
- $O(1)$ regret under (dual) non-degeneracy **[Jasin & Kumar'12]**
- most results use V^{fl} as benchmark (including 'prophet inequality')

proposition [Vera & B'18]

for degenerate instances, $V^{fl} - \mathbb{E}[V^{off}] = \Omega(\sqrt{T})$

Bayes selector for i.i.d arrivals

Bayes selector

$\pi_t = \mathbb{P}\left[V^{off}[t, T] \text{ decreases if } \texttt{OFFLINE} \text{ accepts } t^{th} \text{ arrival}\right]$

– accept t^{th} arrival iff $\pi_t > 0.5$

Bayes selector for i.i.d arrivals

Bayes selector

$\pi_t = \mathbb{P}\left[V^{off}[t,T] \text{ decreases if } \texttt{OFFLINE} \text{ accepts } t^{th} \text{ arrival}\right]$

– accept t^{th} arrival iff $\pi_t > 0.5$

re-solved fluid LP

$$
\begin{aligned}
max.\ & \sum_{i=1}^{n} x_i[t] v_i \\
s.t.\ & Ax[t] \leq B[t], \\
& 0 \leq x_i[t] \leq \mathbb{E}[N_i[t:T]]
\end{aligned}
$$

Bayes selector for i.i.d arrivals

Bayes selector

$\pi_t = \mathbb{P}\left[V^{off}[t, T] \text{ decreases if } \texttt{OFFLINE} \text{ accepts } t^{th} \text{ arrival}\right]$

– accept t^{th} arrival iff $\pi_t > 0.5$

re-solved fluid LP

$$\begin{aligned} max. \ & \sum_{i=1}^{n} x_i[t] v_i \\ s.t. \ & Ax[t] \leq B[t], \\ & 0 \leq x_i[t] \leq \mathbb{E}[N_i[t : T]] \end{aligned}$$

the re-solved LP gives an approximate admission oracle

fluid Bayes selector

accept type θ_i iff $\frac{x_i[t]}{\mathbb{E}[N_i[t:T]]} > 0.5$

Bayes selector for i.i.d arrivals

Bayes selector

$\pi_t = \mathbb{P}\left[V^{off}[t, T] \text{ decreases if } \texttt{OFFLINE} \text{ accepts } t^{th} \text{ arrival}\right]$

– accept t^{th} arrival iff $\pi_t > 0.5$

re-solved fluid LP

$$\begin{aligned} max. \; & \sum_{i=1}^{n} x_i[t] v_i \\ s.t. \; & Ax[t] \leq B[t], \\ & 0 \leq x_i[t] \leq \mathbb{E}[N_i[t:T]] \end{aligned}$$

a

the re-solved LP gives an approximate admission oracle

fluid Bayes selector

accept type θ_i iff $\frac{x_i[t]}{\mathbb{E}[N_i[t:T]]} > 0.5$

proposition [Vera & B, 2018]

fluid Bayes selector has $\mathbb{E}[Regret] \leq 2v_{\max} \sum_{i=1}^{n} p_i^{-1}$

Bayes selector for i.i.d arrivals

Bayes selector

$\pi_t = \mathbb{P}\big[V^{off}[t, T]$ decreases if OFFLINE accepts t^{th} arrival$\big]$

– accept t^{th} arrival iff $\pi_t > 0.5$

re-solved fluid LP

$$\begin{aligned} max. \ & \sum_{i=1}^{n} x_i[t] v_i \\ s.t. \ & Ax[t] \leq B[t], \\ & 0 \leq x_i[t] \leq \mathbb{E}[N_i[t:T]] \end{aligned}$$

a

the re-solved LP gives an approximate admission oracle

fluid Bayes selector

accept type θ_i iff $\frac{x_i[t]}{\mathbb{E}[N_i[t:T]]} > 0.5$

proposition [Vera & B, 2018]

fluid Bayes selector has $\mathbb{E}[Regret] \leq 2 v_{\max} \sum_{i=1}^{n} p_i^{-1}$

– proposed for multi-secretary by **[Gurvich & Arlotto, 2017]**
– NRM via partial resolving **[Bumpensanti & Wang, 2018]**

proof outline

the proof comprises two parts

1. compensated coupling: regret bound for Bayes selector for generic online decision problem
2. bound compensation for online packing problems via LP sensitivity, measure concentration

the compensated coupling: make OFFLINE follow ONLINE

for any time t, budget $B[t]$

- let $V^{off}(t, B[t]) \triangleq$ OFFLINE starting from current state

the compensated coupling: make OFFLINE follow ONLINE

for any time t, budget $B[t]$

- let $V^{off}(t, B[t]) \triangleq$ OFFLINE starting from current state
- for any action a, disagreement set $Q_t(a) \triangleq$ set of sample-paths ω where a is sub-optimal (given $B[t]$)

the compensated coupling: make OFFLINE follow ONLINE

for any time t, budget $B[t]$

- let $V^{off}(t, B[t]) \triangleq$ OFFLINE starting from current state
- for any action a, disagreement set $Q_t(a) \triangleq$ set of sample-paths ω where a is sub-optimal (given $B[t]$)
- can compensate OFFLINE to follow same action a as ONLINE

$$V^{off}(t, B[t]) \leq R_t^{alg} + v_{\max} \mathbf{1}_{\omega \in Q_t(a)} + V^{off}(t+1, B[t+1])$$

the compensated coupling: make OFFLINE follow ONLINE

for any time t, budget $B[t]$

- let $V^{off}(t, B[t]) \triangleq$ OFFLINE starting from current state
- for any action a, disagreement set $Q_t(a) \triangleq$ set of sample-paths ω where a is sub-optimal (given $B[t]$)
- can compensate OFFLINE to follow same action a as ONLINE

$$V^{off}(t, B[t]) \leq R_t^{alg} + v_{\max} \mathbf{1}_{\omega \in Q_t(a)} + V^{off}(t+1, B[t+1])$$

- iterating, we get

$$\mathbb{E}[V^{off}] \leq \mathbb{E}[V^{alg}] + v_{\max} \sum_{t=1}^{T} \mathbb{P}[Q_t(a_t)]$$

note: Bayes selector picks $a_t = \min_a \mathbb{P}[Q_t(a_t)]$

compensated coupling for single resource allocation

for any time t, budget $B[t]$

- if Bayes selector rejects type θ_i, assume `OFFLINE` front-loads θ_i
  - error only if `OFFLINE` rejects all future θ_i
- if Bayes selector accepts type θ_i, assume `OFFLINE` back-loads θ_i
  - error only if `OFFLINE` accepts all future θ_i

compensated coupling for single resource allocation

for any time t, budget $B[t]$

- if Bayes selector rejects type θ_i, assume `OFFLINE` front-loads θ_i
  - error only if `OFFLINE` rejects all future θ_i
- if Bayes selector accepts type θ_i, assume `OFFLINE` back-loads θ_i
  - error only if `OFFLINE` accepts all future θ_i
- claim: smaller of the two events has probability $e^{-c(T-t)}$

summary

online allocation via the Bayes selector

- new online allocation policy with horizon-independent regret
- way to use black-box predictive algorithms
- generic regret bounds for any online decision problem

Thanks!