Online Learning in Repeated Auctions

Jonathan Weed
Massachusetts Institute of Technology

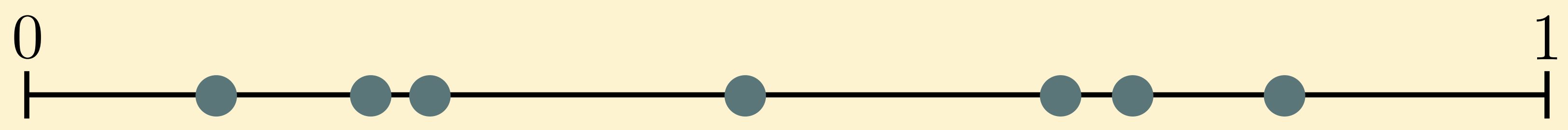
Workshop on Learning
National University Singapore
November 20, 2015

Joint work with

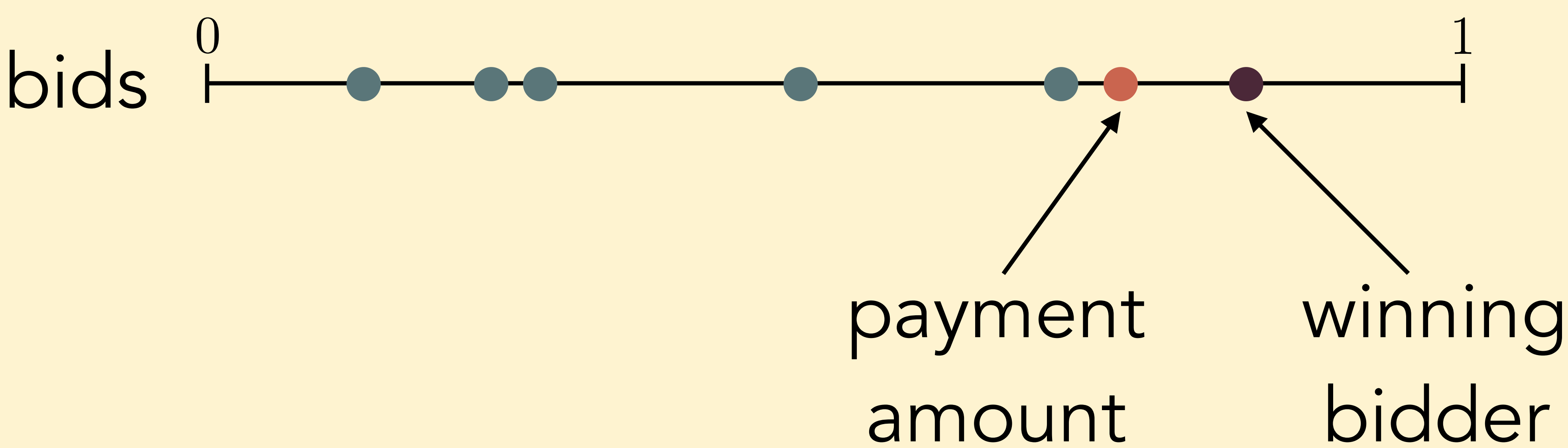
Vianney Perchet
Université Paris Diderot

Philippe Rigollet
Massachusetts Institute of Technology

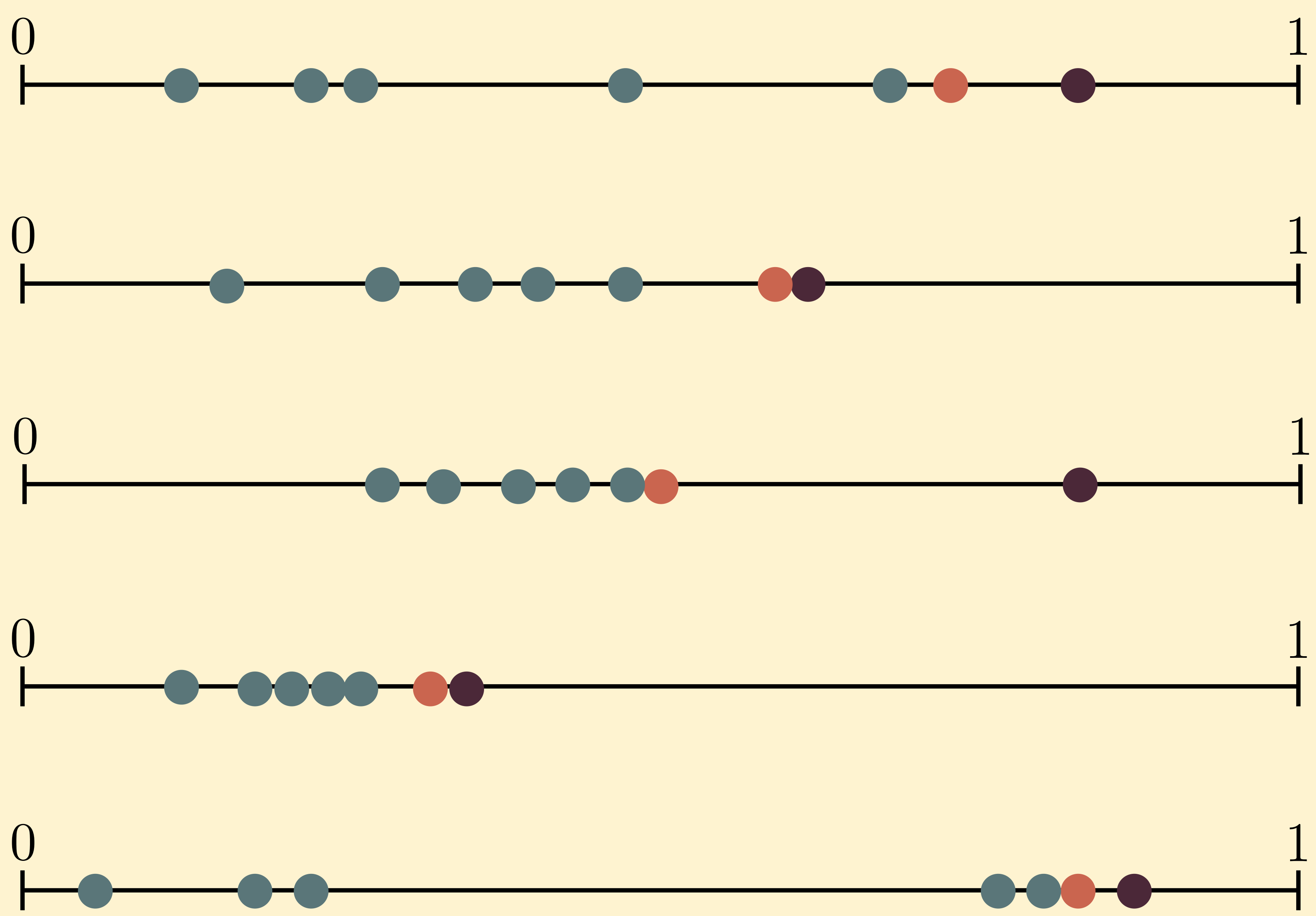
Second Price Auctions

bids 

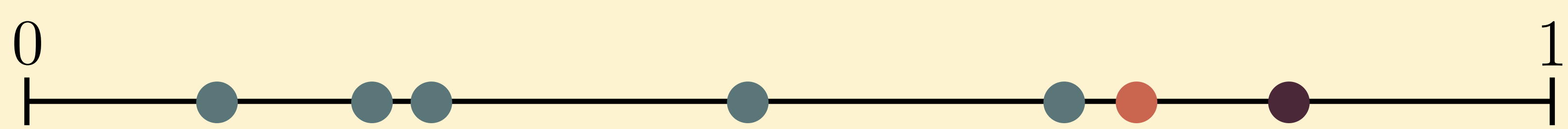
Second Price Auctions



Second Price Auctions



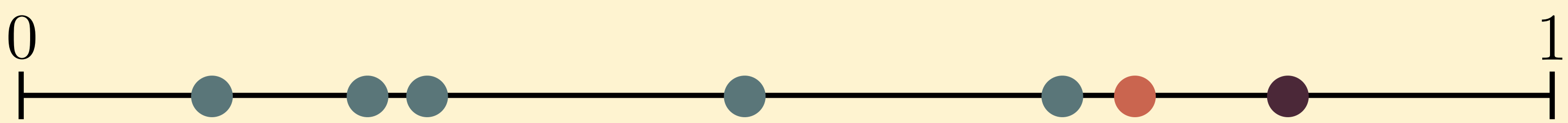
Second Price Auctions



Seller
Profit-maximizing
(with reserve price)
[Myerson ’81]

Bidder
Truthful
(should report true value)
[Vickrey ’61]

Second Price Auctions



Seller

Profit-maximizing

(with reserve price)

[Myerson '81]

How should seller learn reserve price?

[lots of recent work]

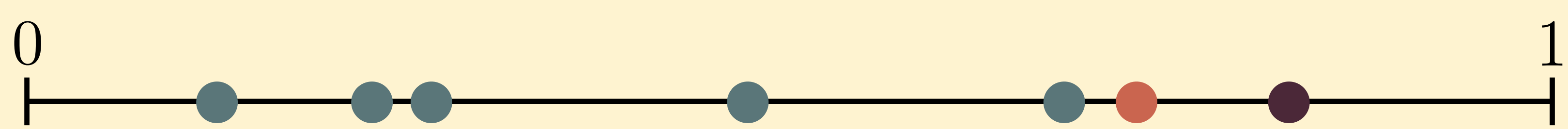
Bidder

Truthful

(should report true value)

[Vickrey '61]

Second Price Auctions

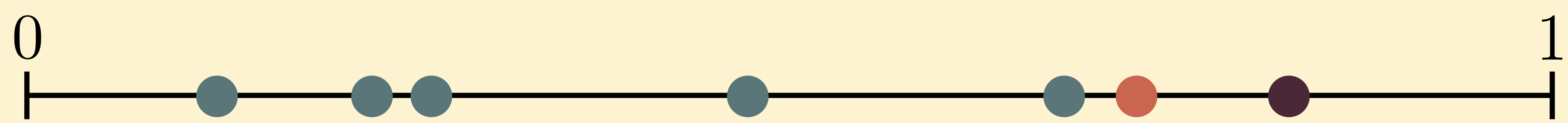


Seller
Profit-maximizing
(with reserve price)
[Myerson '81]

How should seller
learn reserve price?
[lots of recent work]

Bidder
Truthful
(should report true value)
[Vickrey '61]

Second Price Auctions



Seller

Profit-maximizing

(with reserve price)

[Myerson ’81]

How should seller learn reserve price?

[lots of recent work]

Bidder

Truthful

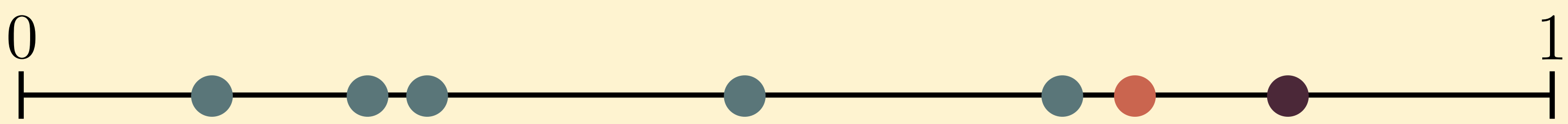
(should report true value)

[Vickrey ’61]

How should bidder act?

[this talk]

How should bidder act?



low information,
but safe

high information,
but costly

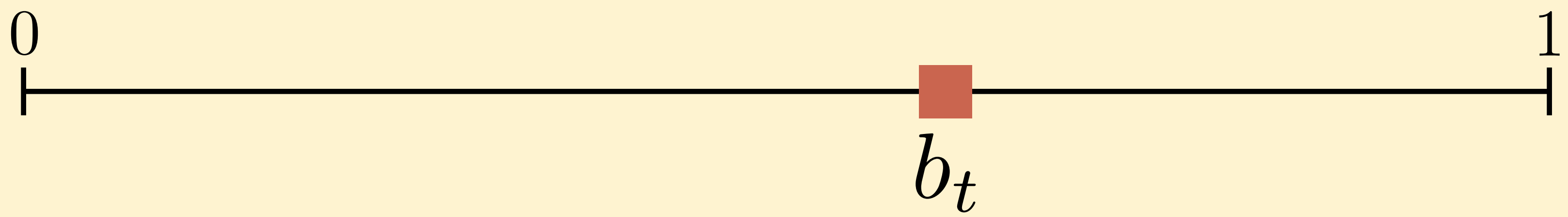
"bandit"-like tradeoff between
exploration and exploitation

Model

round t:

Model

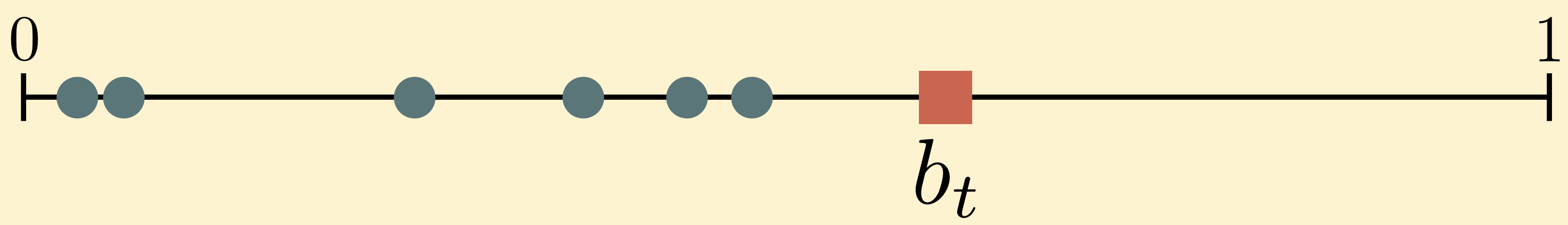
round t:



player bids b_t

Model

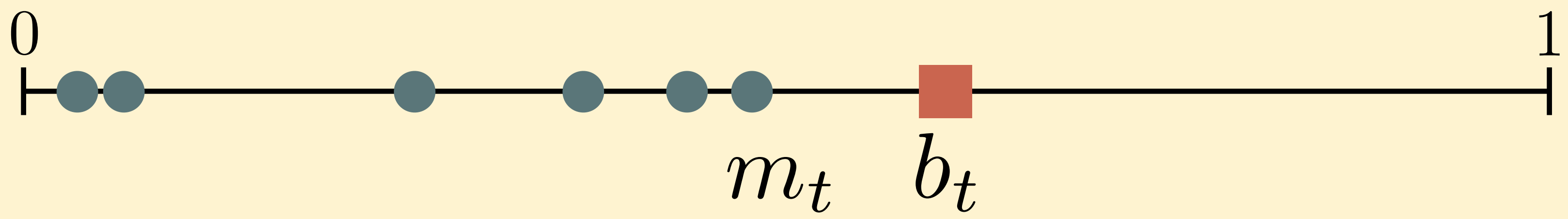
round t:



player bids b_t

Model

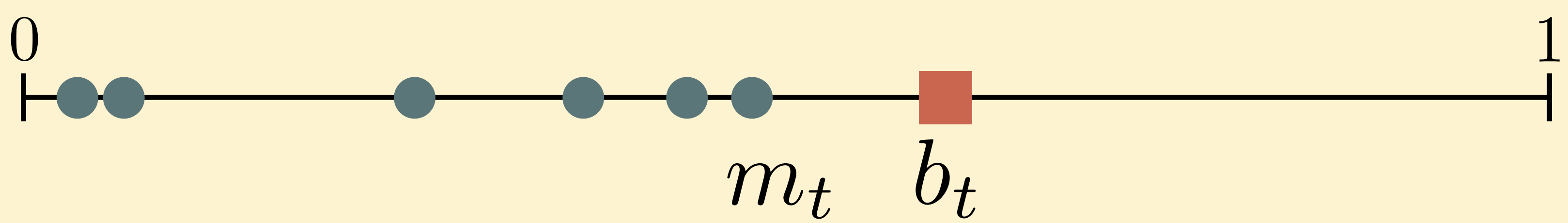
round t:



player bids b_t

Model

round t:



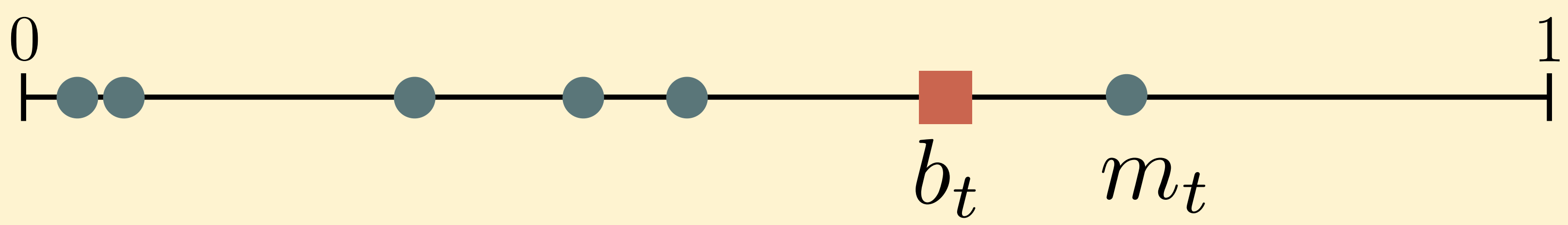
player bids b_t

if $b_t > m_t$ (maximum of adversaries' bids):

bidder wins item, observes $v_t \in [0, 1]$
bidder pays m_t

Model

round t:



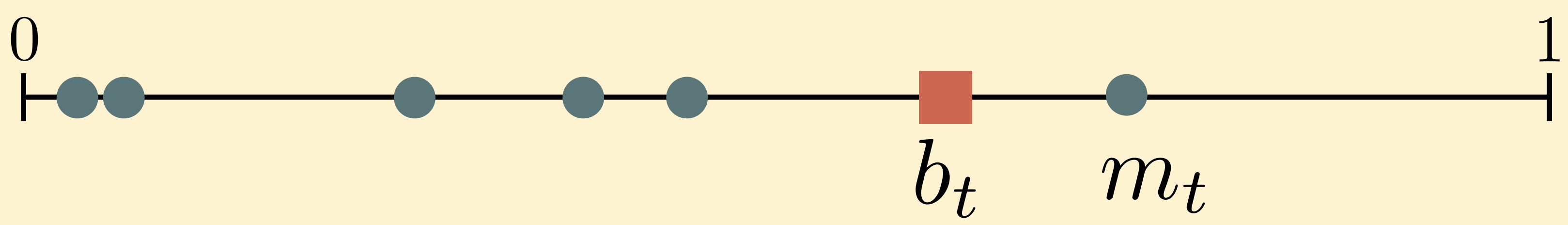
player bids b_t

if $b_t < m_t$:

bidder does not observe v_t

Model

round t:



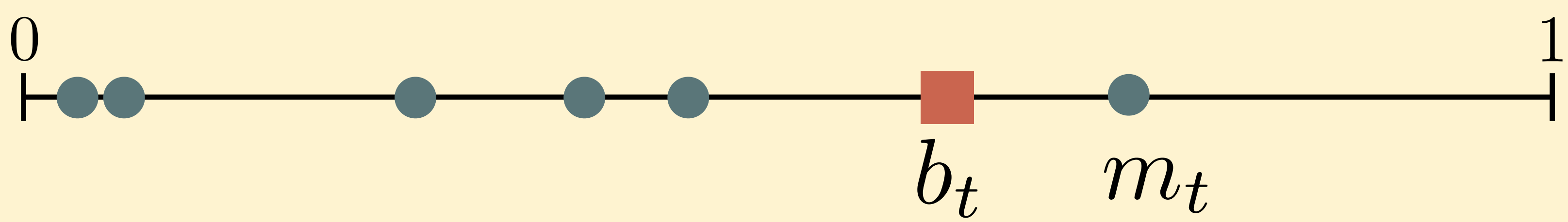
player bids b_t

Total utility:

$$\sum_{t=1}^{T} (v_t - m_t) \mathbb{1}\{b_t > m_t\}$$

Model

round t:



player bids b_t

Total regret:

$$\max_{b\in[0,1]} \sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b > m_t\} - \sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b_t > m_t\}$$

Model

$$\max_{b\in[0,1]} \sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b > m_t\} - \sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b_t > m_t\}$$

Stochastic framework:

v_t i.i.d.

$\mathbb{E}[v_t] = v$ (unknown)

Adversarial framework:

no assumption on v_t

Stochastic framework

Bound *pseudo regret:*

$$\max_{b\in[0,1]} \mathbb{E}\left[\sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b > m_t\}\right] - \mathbb{E}\left[\sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b_t > m_t\}\right]$$

Stochastic framework

Bound *pseudo regret:*

$$\max_{b\in[0,1]} \mathbb{E}\left[\sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b > m_t\}\right] - \mathbb{E}\left[\sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b_t > m_t\}\right]$$

Stochastic framework

Bound *pseudo regret:*

$$\max_{b\in[0,1]} \mathbb{E}\left[\sum_{t=1}^{T}(v-m_t)\mathbb{1}\{b>m_t\}\right]-\mathbb{E}\left[\sum_{t=1}^{T}(v-m_t)\mathbb{1}\{b_t>m_t\}\right]$$

Stochastic framework

Bound *pseudo regret:*

$$\max_{b\in[0,1]} \mathbb{E}\left[\sum_{t=1}^{T}(v-m_t)\mathbb{1}\{b>m_t\}\right]-\mathbb{E}\left[\sum_{t=1}^{T}(v-m_t)\mathbb{1}\{b_t>m_t\}\right]$$

Stochastic framework

Bound *pseudo regret:*

$$\mathbb{E}\left[\sum_{t=1}^{T}(v - m_t)\mathbb{1}\{v > m_t\}\right] - \mathbb{E}\left[\sum_{t=1}^{T}(v - m_t)\mathbb{1}\{b_t > m_t\}\right]$$

Stochastic framework

Bound *pseudo regret:*

$$\bar{R}_T = \mathbb{E}\left[\sum_{t=1}^{T}(v - m_t)\mathbb{1}\{v > m_t\}\right] - \mathbb{E}\left[\sum_{t=1}^{T}(v - m_t)\mathbb{1}\{b_t > m_t\}\right]$$

Stochastic framework

Bound *pseudo regret:*

$$\bar{R}_T = \mathbb{E}\left[\sum_{t=1}^{T}(v - m_t)\mathbb{1}\{v > m_t\}\right] - \mathbb{E}\left[\sum_{t=1}^{T}(v - m_t)\mathbb{1}\{b_t > m_t\}\right]$$

Observation:

just need to learn v!

UCBid

[Upper Confidence Bid]

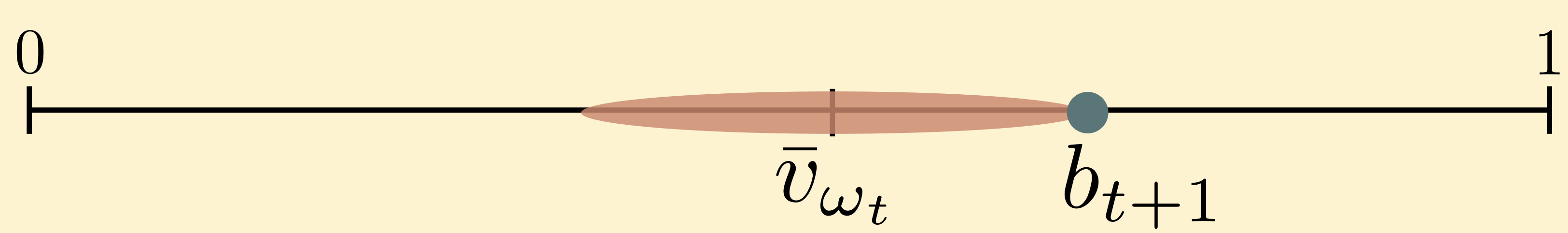
Round 1: bid $b_1 = 1$

Round $t+1$: bid

$$b_{t+1} = \bar{v}_{\omega_t} + \sqrt{\frac{3\log t}{2\omega_t}} \wedge 1$$

empirical average

number of observations



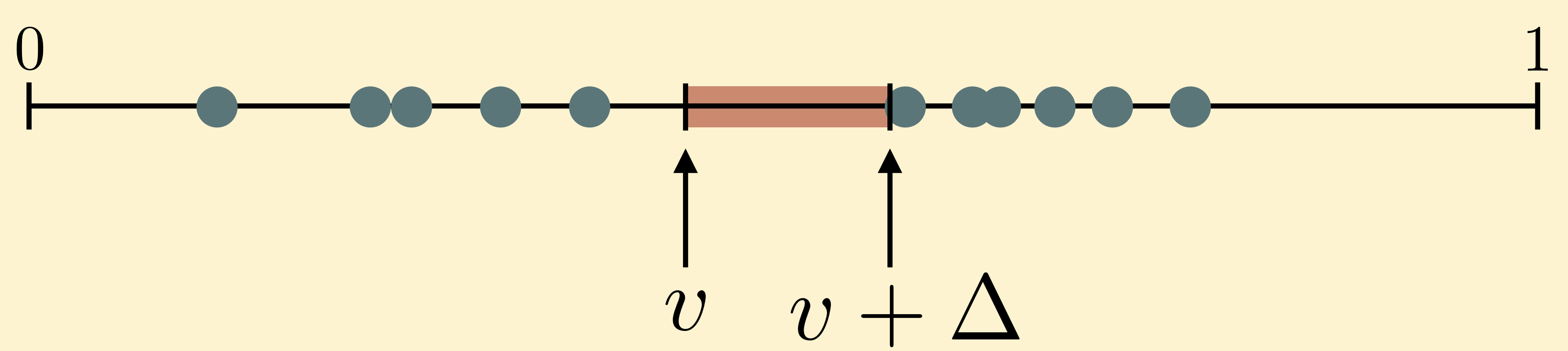
UCBid

[Upper Confidence Bid]

Theorem:

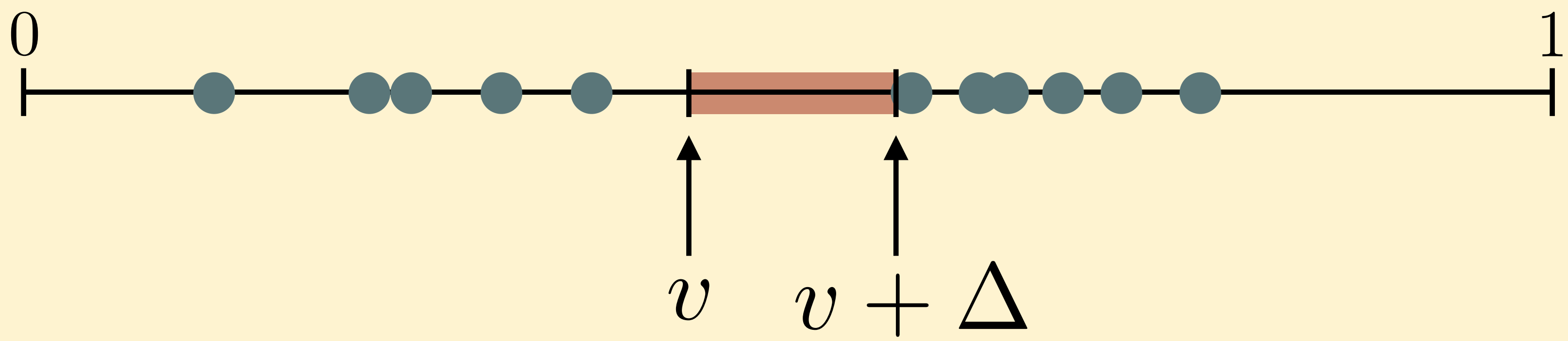
UCBid yields a pseudo regret bound of

$$\bar{R}_T \leq 3 + \frac{12 \log T}{\Delta} \wedge 2\sqrt{6T \log T}$$



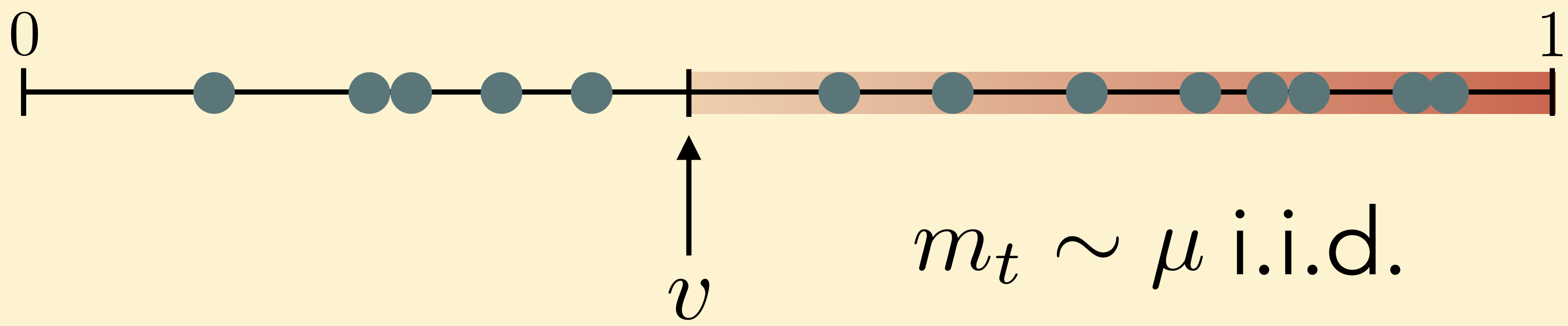
UCBid

[Upper Confidence Bid]



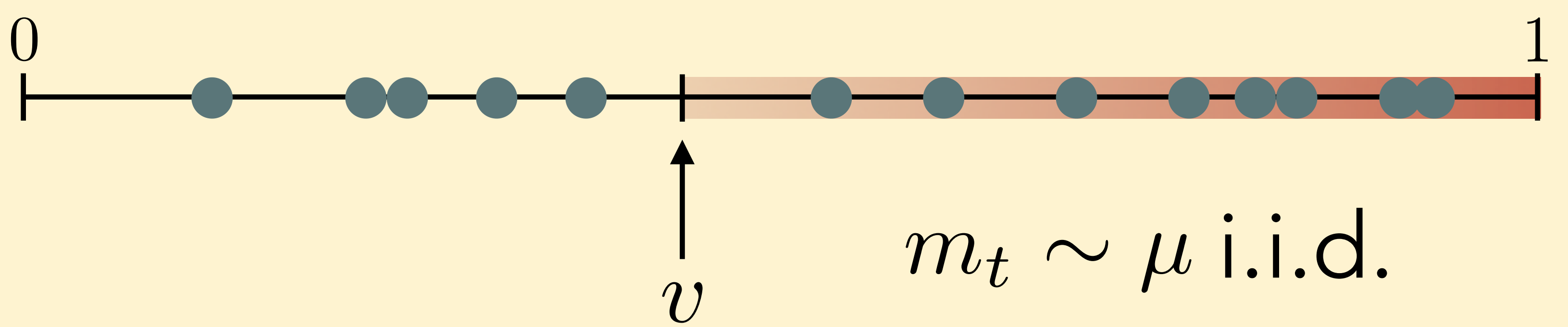
UCBid

[Upper Confidence Bid]



UCBID

[Upper Confidence Bid]



$m_t \sim \mu$ i.i.d.

Margin condition [Tsybakov]:

$$\mu\{(v, v+u]\} \leq C_\mu u^\alpha \quad \forall u > 0$$

UCBid

[Upper Confidence Bid]

$\forall u > 0$
$\mu\{(v, v+u]\} \leq C_\mu u^\alpha$

Theorem:

If $m_t \sim \mu$ i.i.d. and μ satisfies margin condition, then

$$\bar{R}_T \leq \begin{cases} c_1 T^{\frac{1-\alpha}{2}} \log^{\frac{1+\alpha}{2}}(T) & \text{if } \alpha < 1 \\ c_2 \log^2(T) & \text{if } \alpha = 1 \\ c_3 \log(T) & \text{if } \alpha > 1 \end{cases}$$

UCBid

[Upper Confidence Bid]

$\forall u > 0$
$\mu\{(v, v+u]\} \le C_\mu u^\alpha$

Theorem:

If $m_t \sim \mu$ i.i.d. and μ satisfies margin condition, then

$$\bar{R}_T \le \begin{cases} c_1 T^{\frac{1-\alpha}{2}} \log^{\frac{1+\alpha}{2}}(T) & \text{if } \alpha < 1 \\ c_2 \log^2(T) & \text{if } \alpha = 1 \\ c_3 \log(T) & \text{if } \alpha > 1 \end{cases}$$

Lower bound:

$$\bar{R}_T \ge \begin{cases} C_\alpha T^{\frac{1-\alpha}{2}} & \text{if } \alpha < 1 \\ C_\alpha \log T & \text{if } \alpha \ge 1 \end{cases}$$

Adversarial Framework

$$\max_{b\in[0,1]} \sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b > m_t\} - \sum_{t=1}^{T}(v_t - m_t)\mathbb{1}\{b_t > m_t\}$$

No assumptions on v_t and m_t—may even be coupled.

Adversarial Framework

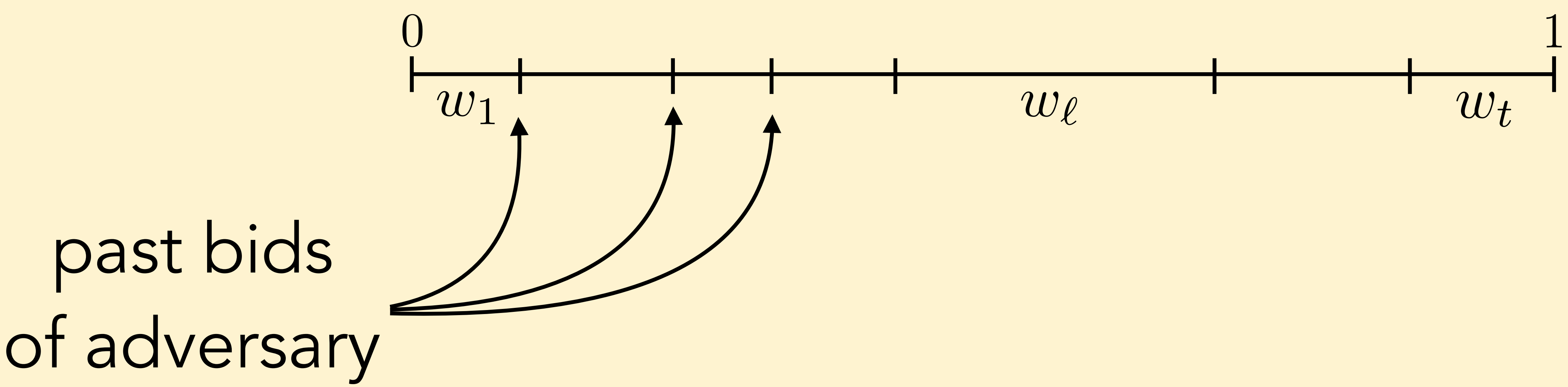
$$\max_{b\in[0,1]} \sum_{t=1}^{T} (v_t - m_t)\mathbb{1}\{b > m_t\} - \sum_{t=1}^{T} (v_t - m_t)\mathbb{1}\{b_t > m_t\}$$

No assumptions on v_t and m_t—may even be coupled.

~~just need to learn v!~~ $\longrightarrow$ mean can be arbitrarily bad

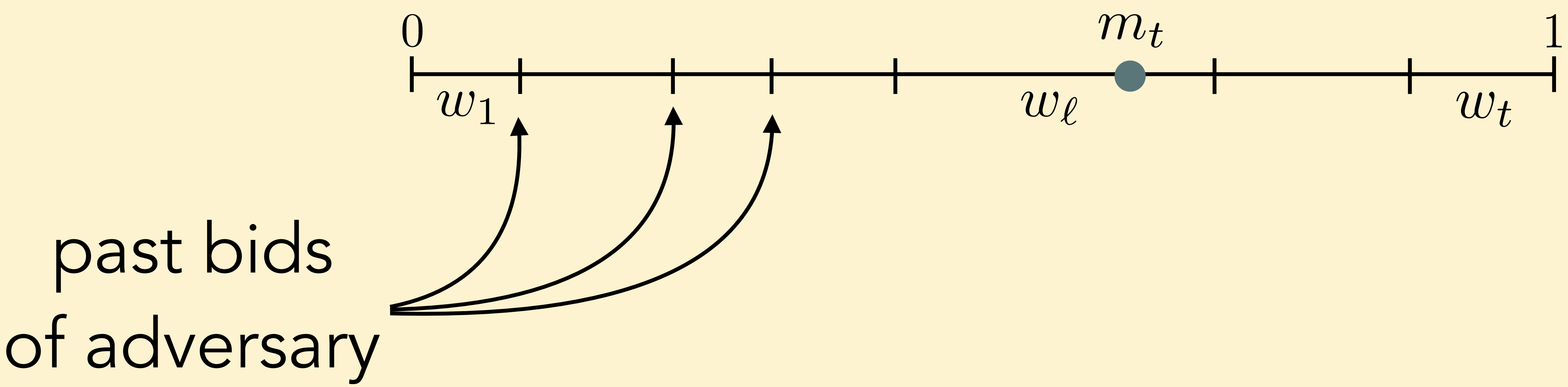
EXPTREE

Idea: Maintain a series of nested partitions of $[0, 1]$.



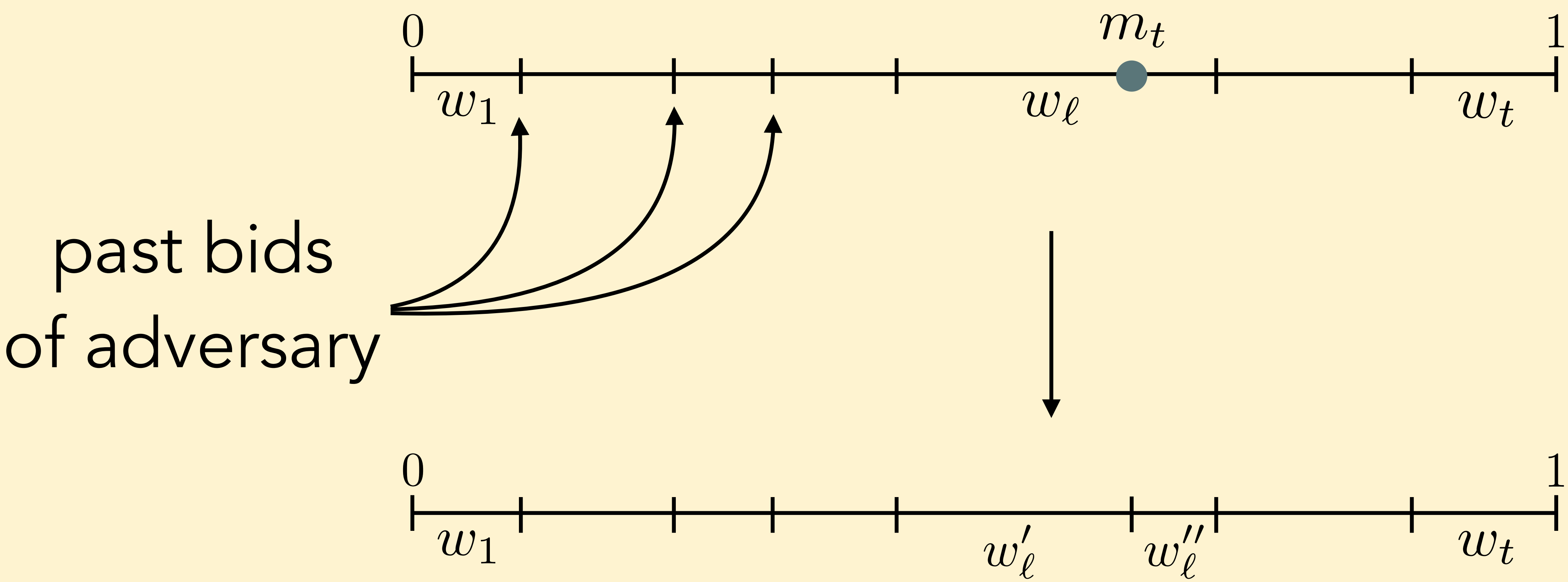
ExpTree

Idea: Maintain a series of nested partitions of $[0,1]$.



ExpTree

Idea: Maintain a series of nested partitions of $[0,1]$.



$w'_\ell + w''_\ell = w_\ell$, proportional to lengths

ExpTree

Idea: Maintain a series of nested partitions of $[0, 1]$.

Play variant of Exp3 on intervals, reassigning weights with each split.

Theorem:

ExpTree yields a pseudo-regret bound of

$$\bar{R}_T \leq 4\sqrt{T \log(1/\Delta^\circ)}$$

where Δ° is width of interval containing best fixed bid.

ExpTree

Theorem:

ExpTree yields a pseudo-regret bound of

$$\bar{R}_T \leq 4\sqrt{T \log(1/\Delta^\circ)}$$

where Δ° is width of interval containing best fixed bid.

ExpTree

Theorem:

ExpTree yields a pseudo-regret bound of

$$\bar{R}_T \leq 4\sqrt{T \log(1/\Delta^\circ)}$$

where Δ° is width of interval containing best fixed bid.

Is dependence on Δ° necessary? **Yes.**

ExpTree

Theorem:

ExpTree yields a pseudo-regret bound of

$$\bar{R}_T \leq 4\sqrt{T \log(1/\Delta^\circ)}$$

where Δ° is width of interval containing best fixed bid.

Is dependence on Δ° necessary? **Yes.**

Lower bound:

$$\bar{R}_T \geq \frac{1}{32}\sqrt{T \lfloor \log_2(1/2\Delta^\circ) \rfloor}$$

Further Questions

- What are the effect of covariates?
- Are better bounds available for well behaved adversaries?

	Upper bound	Lower bound
Stochastic	$O(\log T/\Delta \wedge \sqrt{T\log T})$	
With margin condition	$O(T^{\frac{1-\alpha}{2}}\log^{\frac{1+\alpha}{2}}(T))$	$\Omega(T^{\frac{1-\alpha}{2}})$
Adversarial	$O(\sqrt{T\log(1/\Delta^\circ)})$	$\Omega(\sqrt{T\log(1/\Delta^\circ)})$