No-Regret Incentive-Compatible Online Learning under Exact Truthfulness with Non-Myopic Experts

Junpei Komiyama, Nishant Mehta, Ali Mortazavi

NYU

University of Victoria

A forecasting game

You (Mechanism) want to sequentially forecast various global events

Mechanism is clueless about the events but has access to N experts

Each expert:

- has fixed beliefs about the events
- believes all events are independent and does not update its beliefs over time

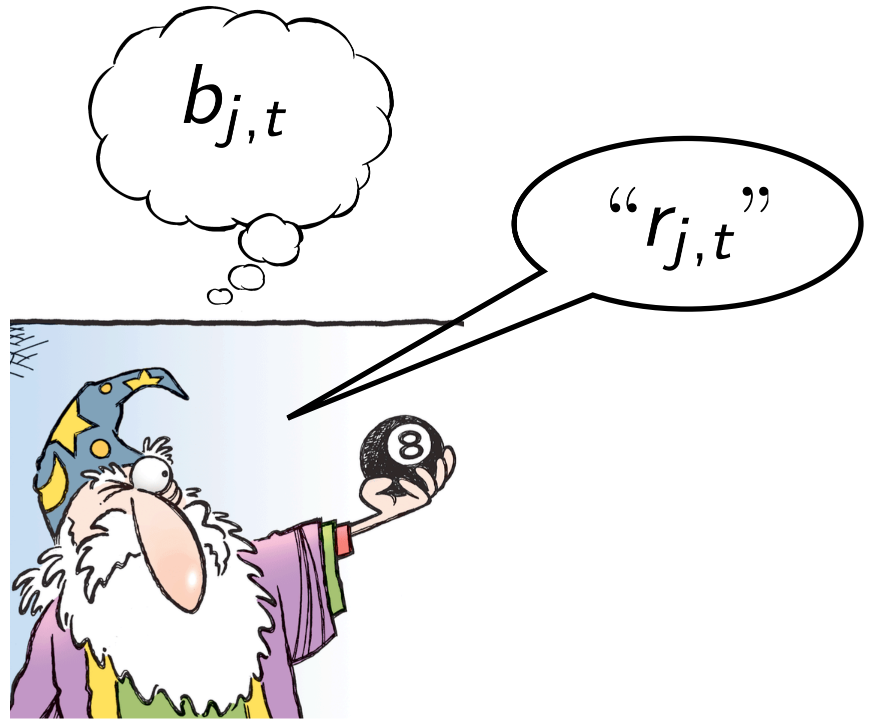
For $t = 1, 2, \ldots, T$:

Each expert $j \in [N]$ selects a forecast (report) $r_{j,t} \in [0, 1]$ and reveals its report

Mechanism draws expert $I_t \sim p_t$, pays it a fixed sum of money  , and selects report $r_{I_t,t}$

Nature reveals outcome $y_t \in \{0, 1\}$ and Mechanism suffers loss $\ell(r_{I_t,t}, y_t)$

Mechanism updates by setting p_{t+1}



expert j

Forecaster's goal: Obtain low *belief regret* ⟷ regret against best expert in hindsight when considering experts' ***beliefs***

$$\sum_{t=1}^{T} \ell(r_{I_t,t}, y_t) - \min_{j \in [N]} \sum_{t=1}^{T} \ell(b_{j,t}, y_t)$$

Belief Regret and Truthfulness

Belief Regret: $\sum_{t=1}^{T} \ell(r_{I_t,t}, y_t) - \min_{j \in [N]} \sum_{t=1}^{T} \ell(b_{j,t}, y_t)$

Exactly truthful mechanisms

- The best thing for an expert to do is to report her belief
- Expert has no reason to spend resources on finding the best way to misreport (lie), so more resources are available to obtain better beliefs. This is good.
- Classic regret becomes identical to belief regret, so it suffices to bound classical regret

Throughout, we will assume the loss function is *strictly proper*

For all beliefs $b \in [0,1]$ and $r \in [0,1]$ with $r \neq b$,

$$\mathrm{E}_{y \sim \mathrm{Bernoulli}(b)} \left[\ell(b, y)\right] < \mathrm{E}_{y \sim \mathrm{Bernoulli}(b)} \left[\ell(r, y)\right]$$

So truthful reporting strictly minimizes the expected loss according to the subjective belief

An Expert's Incentive

Myopic expert

- selects report to maximize probability of being selected in the next round $\Pr(I_{t+1} = i)$

Non-myopic expert

- selects report to maximize sum of probabilities of being selected in all future rounds $\sum_{s=t+1}^{T+1} \Pr(I_s = i)$

  OR

- selects report to maximize probability of being selected at the end of the game $\Pr(I_{T+1} = i)$

What rates are known for expected belief regret?

	Myopic		Non-myopic	
	Exact Truthfulness	Approximate Truthfulness	Exact Truthfulness	Approximate Truthfulness
Full-Information	Weighted-Score Update $\sqrt{T \log N}$ $\Rightarrow$ (Freeman, Pennock, Podimata, Wortman Vaughan, 2021)	$\sqrt{T \log N}$	?	Exponentially Weighted Average Forecaster $\sqrt{T \log N}$ (Frongillo, Gomez, Thilagar, Waggoner, 2021)
Bandit	TS-Prod $\sqrt{TN}$ $\Rightarrow$ (Zimmert and Marinov, 2024)	$\sqrt{TN}$	?	?

What rates are known for expected belief regret?

	Myopic		Non-myopic	
	Exact Truthfulness	Approximate Truthfulness	Exact Truthfulness	Approximate Truthfulness
Full-Information	Weighted-Score Update $\sqrt{T \log N}$ $\Rightarrow$ (Freeman, Pennock, Podimata, Wortman Vaughan, 2021)	$\sqrt{T \log N}$	NEW: $\sqrt{TN} \log T$	Exponentially Weighted Average Forecaster $\sqrt{T \log N}$ (Frongillo, Gomez, Thilagar, Waggoner, 2021)
Bandit	TS-Prod $\sqrt{TN}$ $\Rightarrow$ (Zimmert and Marinov, 2024)	$\sqrt{TN}$	NEW: $T^{2/3} N^{1/3} \log T$	?

Belief Structure

- Each expert maintains a joint distribution over the outcomes and other experts' reports in all rounds

- We assume *Belief Independence* - the joint distribution factorizes over rounds

$$\mathbf{r}_{-i,1}, y_1 \times \mathbf{r}_{-i,2}, y_2 \times \cdots \times \mathbf{r}_{-i,T}, y_T$$

Expert i's belief structure

Note: strong assumption!

Online Incentive Compatibility under Belief Independence

We say an online mechanism is *online incentive compatible under belief independence* if

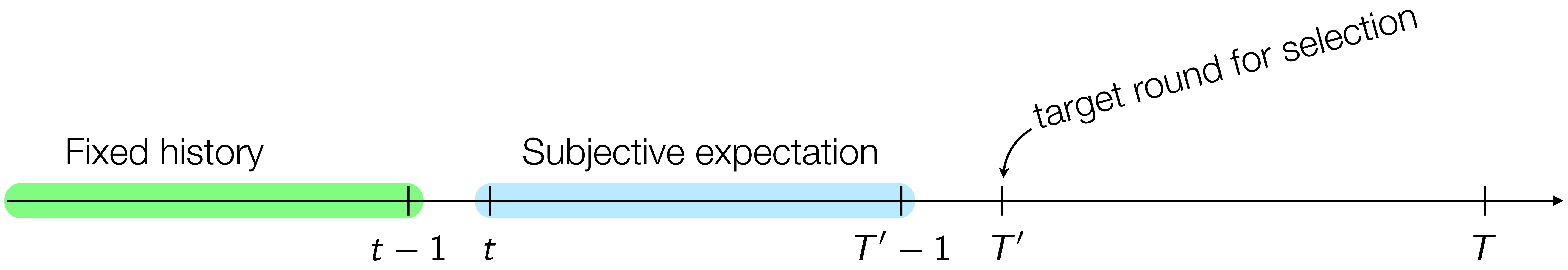
- For all experts i with respective beliefs satisfying belief independence
- For any round t , any history $\mathbf{r}_{1:t-1}, y_{1:t-1}$, any future round (for selection) T'
- For any report sequence $r_{i,t:T'-1}$ that disagrees with $b_{i,t:T'-1}$ in at least one round

it holds that

$$\Pr^{(i)}(I_{T'} = i \mid \mathbf{r}_{1:t-1}, y_{1:t-1}) \quad > \quad \Pr^{(i)}(I_{T'} = i \mid \mathbf{r}_{1:t-1}, y_{1:t-1})$$

expert i selects beliefs $b_{i,t:T'-1}$
in rounds t through T'

expert i selects reports $r_{i,t:T'-1}$
in rounds t through T'

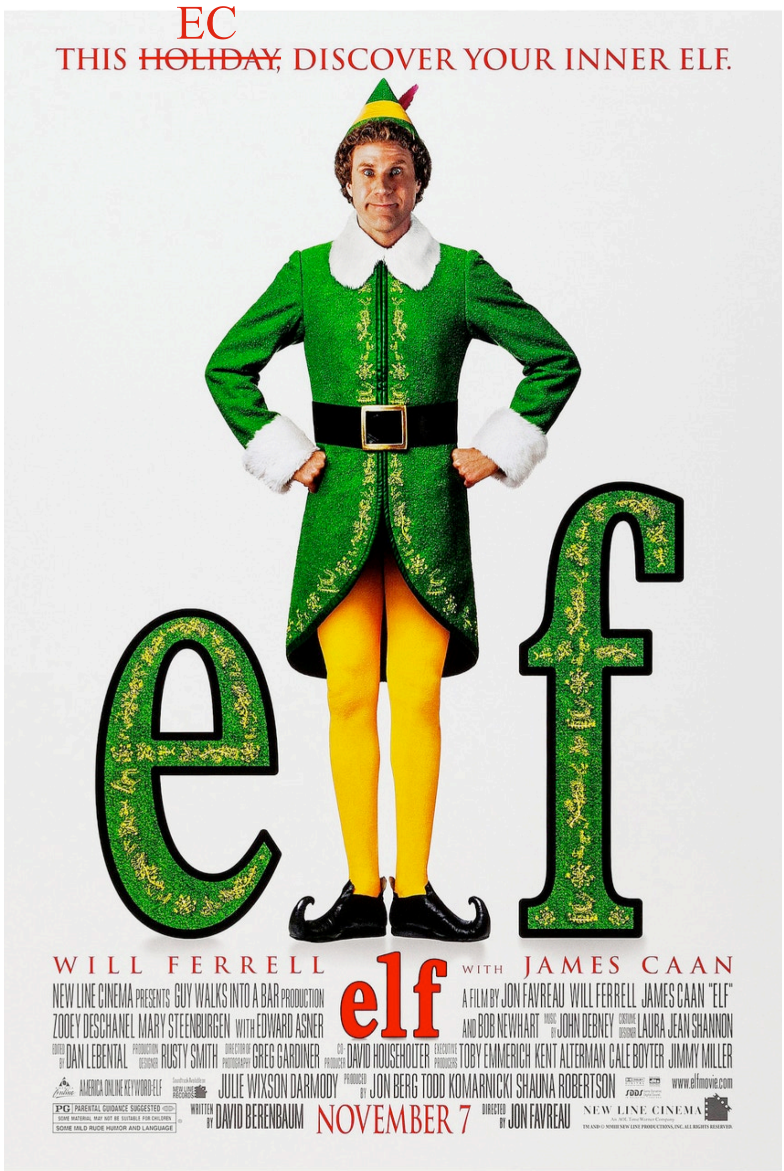


Elven magic

One-shot game (Witkowski, Freeman, Wortman Vaughan, Pennock, Krause, 2018, 2023)

- Suppose all experts simultaneously provide forecasts for all events (one round game)
- Select one expert and give it a single cash prize 
- Goal: have experts report truthfully while maximizing probability of selecting most accurate expert (expert whose internal beliefs give the lowest cumulative loss according to a proper loss)

Event-Lotteries Forecaster (ELF)



- For each event, run a lottery, assigning a “happy point” to a single expert

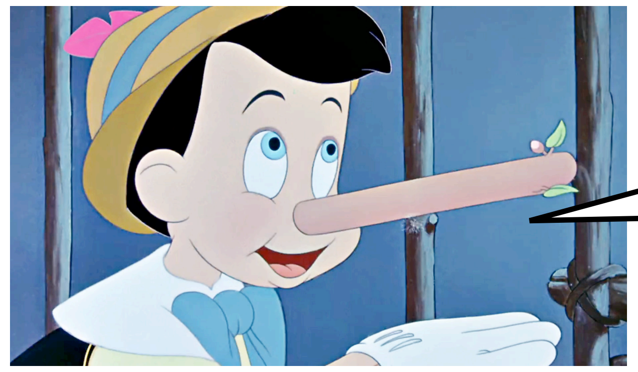
$$\Pr\left(W_{j,t}=1\right)=\frac{1}{N}\left(1-\ell(r_{j,t},y_t)+\frac{1}{N-1}\sum_{k\neq j}\ell(r_{k,t},y_t)\right)$$

- Give the prize to the expert with the most happy points

$$\text{Winner}=\arg\max_{j\in[N]}\sum_{t=1}^{T}W_{j,t}$$

Our Algorithm: FPL-ELF

Main Idea of FPL-ELF



For $t = 1, 2, \ldots, T$:

To determine I_t, run ELF using losses from rounds 1 through $t - 1$ and use forecast of selected expert

with a small but important modification to ELF

This approach ^ was called "ELF-X" by Freeman, Pennock, Podimata, Wortman Vaughan (2020)

What we actually do: swap out ELF with "Simple ELF"

- Simple ELF is ELF but with a per-event lottery that is simpler and has other important modifications

Modifying the lottery

Original ELF: probability of getting a happy point

$$\frac{1}{N}\left(1 - \ell(r_{j,t}, y_t) + \frac{1}{N-1}\sum_{k\neq j}\ell(r_{k,t}, y_t)\right)$$

Properties:

- Variance of $W_{j,t}$ ranges from 0 to $\frac{2}{N}$
- exactly one expert gets a point
- points are good (happy points)

Simple ELF: probability of getting a sad point

$$\frac{1}{N}\left(\frac{1}{2} + \frac{1}{4}\ell(r_{j,t}, y_t)\right)$$

Properties:

- Variance of $W_{j,t}$ is always $\Theta\left(\frac{1}{N}\right)$
- at most one expert gets a point
- points are bad (sad points)

FPL-ELF

For $t = 1, 2, \ldots, T$:

Select $I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} W_{j,s}$

Observe losses for round t

For $s = 1, 2, \ldots, t$:

Draw candidate $C_s \sim \text{Uniform}([N])$

Set $W_{j,s} = 0$ for $j \neq C_s$

Draw $W_{j,s} \sim \text{Bernoulli}\left(\frac{1}{2} + \frac{1}{4}\ell_{j,s}\right)$ for $j = C_s$

FPL-ELF is online incentive compatible under belief independence

FPL-ELF

For $t = 1, 2, \ldots, T$:

Select $I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} W_{j,s}$

Observe losses for round t

For $s = 1, 2, \ldots, t$:

Draw candidate $C_s \sim \text{Uniform}([N])$

Set $W_{j,s} = 0$ for $j \neq C_s$

Draw $W_{j,s} \sim \text{Bernoulli}\left(\frac{1}{2} + \frac{1}{4}\ell_{j,s}\right)$ for $j = C_s$

FPL-ELF is online incentive compatible under belief independence

Connection to Follow the Perturbed Leader

Selection rule is equivalent to

$$I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} \underbrace{(4N \cdot W_{j,s} - 2)}$$

perturbed loss $\tilde{\ell}_{j,s}$

$$\mathbb{E}\left[\tilde{\ell}_{j,s}\right] = \ell_{j,s}$$

Selection rule can be expressed as

$$I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} \ell_{j,s} + \sum_{s=1}^{t-1} X_{j,s}$$

for perturbations $X_{j,s}$

FPL-SELF (Stabilized/Static ELF)

For $t = 1, 2, \ldots, T$:

Select $I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} W_{j,s}$

Observe losses for round t

~~For $s = 1, 2, \ldots, t$:~~

for regret analysis, not necessary to redraw **anything**

Draw candidate $C_t \sim \text{Uniform}([N])$

Set $W_{j,t} = 0$ for $j \neq C_t$

Draw $W_{j,t} \sim \text{Bernoulli}\left(\frac{1}{2} + \frac{1}{4}\ell_{j,t}\right)$ for $j = C_t$

FPL-ELF is online incentive compatible under belief independence

Connection to Follow the Perturbed Leader

Selection rule is equivalent to

$$I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} (4N \cdot W_{j,s} - 2)$$

perturbed loss $\tilde{\ell}_{j,s}$

$$\mathbb{E}\left[\tilde{\ell}_{j,s}\right] = \ell_{j,s}$$

Selection rule can be expressed as

$$I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} \ell_{j,s} + \sum_{s=1}^{t-1} X_{j,s}$$

for perturbations $X_{j,s}$

Analyzing the regret

FPL-ELF bears similarity to Prediction by Random-Walk Perturbation (Devroye, Lugosi, Neu, 2013)



Draw each perturbation $X_{j,s}$ uniformly from $\{-\frac{1}{2}, \frac{1}{2}\}$

In round t, select expert $I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} \ell_{j,s} + \sum_{s=1}^{t-1} X_{j,s}$

Key Lemma: **Expected Regret** $\leq 2 \sum_{t=1}^{T} \Pr(I_{t+1} \neq I_t) + \mathrm{E}\left[\max_{j \in [N]} \sum_{t=1}^{T} X_{j,t}\right]$

Analyzing the regret

FPL-ELF bears similarity to Prediction by Random-Walk Perturbation (Devroye, Lugosi, Neu, 2013)

Draw each perturbation $X_{j,s}$ uniformly from $\{-\frac{1}{2}, \frac{1}{2}\}$

In round t, select expert $I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} \ell_{j,s} + \sum_{s=1}^{t-1} X_{j,s}$

Key Lemma: $\textbf{Expected Regret} \leq 2 \sum_{t=1}^{T} \Pr(I_{t+1} \neq I_t) + \mathrm{E}\left[\max_{j \in [N]} \sum_{t=1}^{T} X_{j,t}\right]$

- To control the probability of switching,  devised the "Lead Pack"
  - $\mathsf{LeadPack}_t$ (Lead Pack in round t) is set of experts that potentially can "take the lead" in round $t+1$

    (total perturbed loss within 2 of perturbed leader's total perturbed loss)

Analyzing the regret

FPL-ELF bears similarity to Prediction by Random-Walk Perturbation (Devroye, Lugosi, Neu, 2013)

Draw each perturbation $X_{j,s}$ uniformly from $\{-\frac{1}{2}, \frac{1}{2}\}$

In round t, select expert $I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} \ell_{j,s} + \sum_{s=1}^{t-1} X_{j,s}$

Key Lemma: $\textbf{Expected Regret} \leq 2 \sum_{t=1}^{T} \Pr(I_{t+1} \neq I_t) + \mathrm{E}\left[\max_{j \in [N]} \sum_{t=1}^{T} X_{j,t}\right] \leq 2 \sum_{t=1}^{T} \Pr(|\mathrm{LeadPack}_t| > 1) + O\left(\sqrt{T \log N}\right)$

- To control the probability of switching,  devised the "Lead Pack"
  - $\mathrm{LeadPack}_t$ (Lead Pack in round t) is set of experts that potentially can "take the lead" in round $t+1$

    (total perturbed loss within 2 of perturbed leader's total perturbed loss)

Analyzing the regret

FPL-ELF bears similarity to Prediction by Random-Walk Perturbation (Devroye, Lugosi, Neu, 2013)

Draw each perturbation $X_{j,s}$ uniformly from $\{-\frac{1}{2}, \frac{1}{2}\}$

In round t, select expert $I_t = \arg\min_{j \in [N]} \sum_{s=1}^{t-1} \ell_{j,s} + \sum_{s=1}^{t-1} X_{j,s}$

Key Lemma: $\mathsf{Expected\ Regret} \le 2 \sum_{t=1}^{T} \Pr(I_{t+1} \ne I_t) + \mathrm{E}\left[\max_{j \in [N]} \sum_{t=1}^{T} X_{j,t}\right] \le 2 \sum_{t=1}^{T} \Pr(|\mathsf{LeadPack}_t| > 1) + O\left(\sqrt{T \log N}\right)$

- To control the probability of switching,  devised the "Lead Pack"
    - $\mathsf{LeadPack}_t$ (Lead Pack in round t) is set of experts that potentially can "take the lead" in round $t + 1$
      (total perturbed loss within 2 of perturbed leader's total perturbed loss)
- Using properties of binomial distribution,  showed $\Pr(|\mathsf{LeadPack}_t| > 1) = O\left(\sqrt{\frac{\log N}{t}}\right)$
- Putting everything together gives $\mathsf{Expected\ Regret} = O\left(\sqrt{T \log N}\right)$

FPL-ELF vs Random-Walk Perturbation

FPL-ELF is like Random-Walk Perturbation, but with two major differences

1) For a given expert and round, the noise random variable depends on the expert's loss in that round

- Cumulative noise $\sum_{s=1}^{t} X_{j,s}$ is centered version of scaled Poisson binomial random variable $4N \cdot \mathtt{PBin}(p_1, \ldots, p_t)$

$$p_s = \frac{1}{N}\left(\frac{1}{2} + \frac{1}{4}\ell_{j,s}\right)$$

- This is the major difficulty! Difficult to analyze 😱

2) The scale of each noise random variable is very large (order N), so perturbed losses can be at scale N

- For   , perturbed losses were constant order; leader change contributes at most 1 to the regret.

  For us, a leader change can contribute N to the regret, so we pay $N \cdot \sum_{t=1}^{N} \Pr\left(I_{t+1} \neq I_t\right)$

Happy lottery vs Sad lottery (Babylonian lottery)

Happy Lottery

- At most one expert gets a happy point in each round
- Selected expert has the most happy points

$$\Pr(I_{t+1} \neq I_t) \leq \Pr(|\mathsf{LeadPack}_t| > 1)$$

Sad Lottery

- At most one expert gets a sad point in each round
- Selected expert has the least sad points
- For there to be a leader change, candidate must be the leader **and** Lead Pack must contain more than one expert

$$\Pr(I_{t+1} \neq I_t) \leq \Pr(C_t = I_t) \cdot \Pr(|\mathsf{LeadPack}_t| > 1)$$
$$= \frac{1}{N} \Pr(|\mathsf{LeadPack}_t| > 1)$$

In terms of what our analysis can give, a sad lottery is much better than a happy lottery!

(We can kill off a factor of N that was introduced from the scale of the perturbed losses)

Regret bound

The expected regret of FPL-ELF is at most $O\left(\sqrt{TN}\log T\right)$

- log T is likely an artifact of the analysis
- unknown if an exactly truthful mechanism can do better than $O\left(\sqrt{TN}\right)$
- under approximate truthfulness, can do *much* better: $O\left(\sqrt{T\log N}\right)$ (Frongillo, Gomez, Thilagar, Waggoner, 2021)

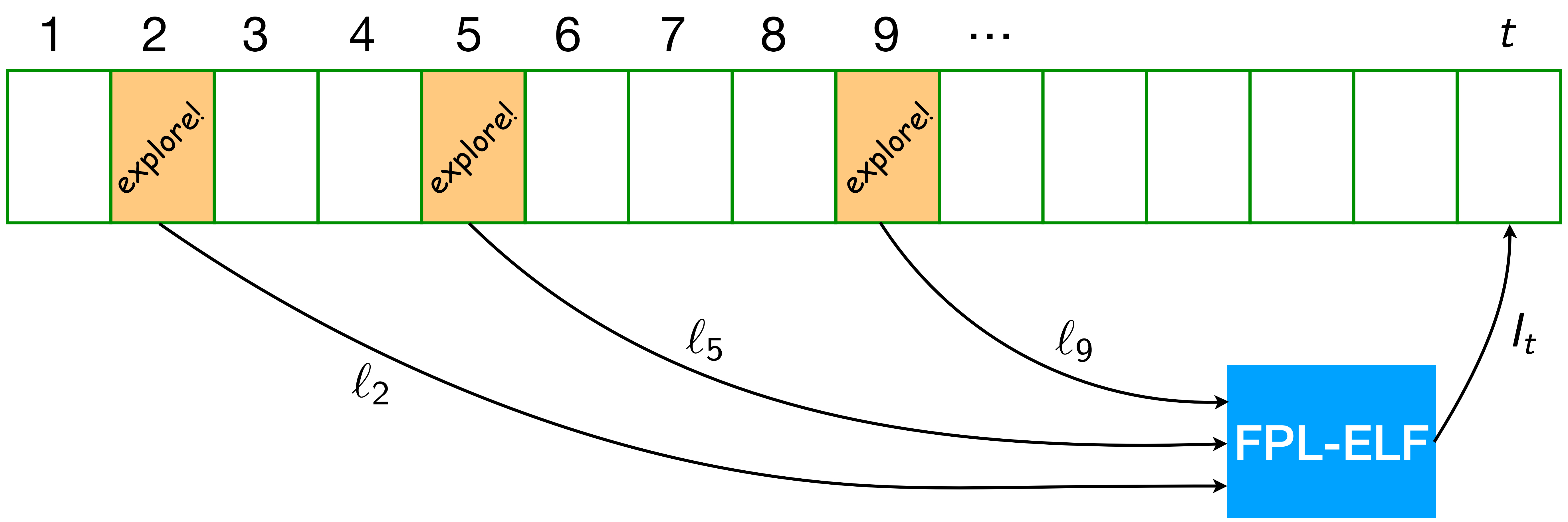
Bandit setting: FPL-ELF-ε

Bandit algorithm - we use an *exploration-separated* version of FPL-ELF

Each round is
- an exploration round with probability ε
- an exploitation round with probability $1 - \varepsilon$

In exploration round, we draw the candidate C_t uniformly from $[N]$ and select this candidate, so $I_t = C_t$

In exploitation round, we use our Simple ELF lottery over the past exploration rounds



The expected regret of FPL-ELF-ε is at most $O\left(T^{2/3}N^{1/3}\log T\right)$

Open Problems

What is the optimal rate for the regret for truthful mechanisms?

In the bandit setting, is exploration separation necessary for exact truthfulness?

In the bandit setting, for truthful mechanisms is it possible to get regret lower than $O\left(T^{2/3}\right)$?

Questions?