Impossible Tuning Made Possible:
A New Expert Algorithm and Its Applications

Liyu Chen Haipeng Luo **Chen-Yu Wei**
University of Southern California

Learning from Experts

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

Learner chooses a distribution over experts $p_t \in \Delta_N$

Adversary chooses a loss vector $\ell_t \in \mathbb{R}^N$

Learner suffers loss $\langle p_t, \ell_t \rangle$ and receives ℓ_t

$$\mathsf{Regret}_T(i) = \sum_{t=1}^{T} \langle p_t, \ell_t \rangle - \sum_{t=1}^{T} \ell_{t,i}$$

Known Bounds

For the standard setting where $\ell_t \in [0,1]^N$

- Minimax regret bound: $\text{Regret}_T(i) = \Theta(\sqrt{T \ln N})$
- First-order regret bound: $\text{Regret}_T(i) = \mathcal{O}\left(\sqrt{\sum_{t=1}^T \ell_{t,i} \ln N}\right)$
- Second-order regret bound (Cesa-Bianchi et al. 2007):

$$\text{Regret}_T(i) = \mathcal{O}\left(\frac{\ln N}{\eta} + \eta \sum_{t=1}^T \ell_{t,i}^2\right)$$

Is $\text{Regret}_T(i) = \mathcal{O}\left(\sqrt{\sum_{t=1}^T \ell_{t,i}^2 \ln N}\right)$ possible for all i simultaneously?

→ the "**impossible tuning**" open problem (Gaillard et al. 2014)

Learning from Experts and Hints

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

Learner receives a hint m_t

Learner chooses a distribution over experts $p_t \in \Delta_N$

Adversary chooses a loss vector $\ell_t \in \mathbb{R}^N$

Learner suffers loss $\langle p_t, \ell_t \rangle$ and receives ℓ_t

(Liang and Steinhardt, 2014): $\text{Regret}_T(i) = \mathcal{O}\left(\frac{\ln N}{\eta} + \eta \sum_{t=1}^{T} (\ell_{t,i} - m_{t,i})^2\right)$

Is $\text{Regret}_T(i) = \mathcal{O}\left(\sqrt{\sum_{t=1}^{T} (\ell_{t,i} - m_{t,i})^2 \ln N}\right)$ possible for all i simultaneously?

Our Contribution

- Answered the two long-standing open problems **affirmatively.**
- Proposed a general **learning-the-learning-rate** framework that leads several new data-dependent bounds for **learning from experts** and **online linear optimization.**

How to Solve the Impossible Tuning Issue

Standard exponential weight algorithm:

$$p_{t+1,i} \propto p_{t,i} \exp\left(-\eta \ell_{t,i}\right)$$

Equivalent form in Online Mirror Descent:

$$p_{t+1} = \text{argmin}_{p \in \Delta_N} \left\{ \langle p, \ell_t \rangle + \frac{1}{\eta} \text{KL}(p, p_t) \right\}$$

Dependent on all experts' losses

$$\Rightarrow \quad \text{Regret}_T(i) \leq \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T} \langle p_t, \ell_t^2 \rangle$$

How to Solve the Impossible Tuning Issue

Second-order term correction trick (Liang and Steinhardt, 2014):

$$p_{t+1,i} \propto p_{t,i} \exp\left(-\eta(\ell_{t,i} - 2\eta\ell_{t,i}^2)\right)$$

$$\Rightarrow \sum_{t=1}^{T} \langle p_t, \ell_t - 2\eta\ell_t^2 \rangle - \sum_{t=1}^{T} (\ell_{t,i} - 2\eta\ell_{t,i}^2) \le \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T} \langle p_t, (\ell_t - 2\eta\ell_t^2)^2 \rangle$$

(The regret bound of exponential weight with $\ell_t - 2\eta\ell_t^2$ as loss)

$$\le \frac{\ln N}{\eta} + 2\eta \sum_{t=1}^{T} \langle p_t, \ell_t^2 \rangle$$

rearrange

$$\Rightarrow \sum_{t=1}^{T} \langle p_t, \ell_t \rangle - \sum_{t=1}^{T} \ell_{t,i} \le \frac{\ln N}{\eta} + 2\eta \sum_{t=1}^{T} \ell_{t,i}^2$$

← Only depend on the best expert's loss

How to Solve the Impossible Tuning Issue

$$\text{Regret}(i) \leq \frac{\ln N}{\eta} + 2\eta \sum_{t=1}^{T} \ell_{t,i}^2$$

How to get $\text{Regret}(i) \leq \mathcal{O}\left(\sqrt{\sum_{t=1}^{T} \ell_{t,i}^2 \ln N}\right)$ for all i *simultaneously*?

One attempt by Gaillard et al., (2014)**:** individual adaptive learning rate

$$p_{t+1,i} \propto p_{t,i} \exp\left(-\eta_{t,i}(\ell_{t,i} - 2\eta_{t,i}\ell_{t,i}^2)\right)$$

(doesn't quite work in the analysis)

How to Solve the Impossible Tuning Issue

Our Solution: change the regularizer in the Online Mirror Descent form

$$p_{t+1} = \text{argmin}_{p \in \Delta_N} \left\{ \sum_{j=1}^{N} p_{t,j} \left(\ell_{t,j} - \underbrace{10\eta_{t,j}\ell_{t,j}^2} \right) + \underbrace{\sum_{j=1}^{N} \frac{p_j}{\eta_{t,j}} \ln \frac{p_j}{p_{t,j}}} \right\}$$

$$\eta_{t,j} \approx \sqrt{\frac{\ln N}{1 + \sum_{\tau=1}^{t} \ell_{\tau,j}^2}}$$

Second-order correction with a slightly larger coefficient

Weighted KL divergence

Key analysis step: The larger correction term creates a useful negative regret

$$\sum_{t=1}^{T} \langle p_t, \ell_t \rangle - \sum_{t=1}^{T} \ell_{t,i} \le \frac{\ln N}{\eta_{T,i}} + 10 \sum_{t=1}^{T} \eta_{t,i} \ell_{t,i}^2 - 8 \sum_{t=1}^{T} \sum_{j=1}^{N} \eta_{t,j} p_{t,j} \ell_{t,j}^2$$

These two tricks resolve the impossible tuning issue.

Extension:

A Unified General Framework for “learning the learning rate”

There are many online learning problems which cannot be completely adaptive simply because the **learning rate is hard to tune.**

The algorithm we just introduced can become a **meta-algorithm** that learns over multiple base algorithms each with fixed learning rates, and automatically adapts to the best learning rate.

Examples in the expert setting:

	Easy	Difficult	Existing solutions
2^{nd}-order quantile bound	$\frac{\mathrm{KL}(u,\pi)}{\eta} + \eta \sum_t \langle u - p_t, \ell_t - m_t \rangle^2$	$\sqrt{\mathrm{KL}(u,\pi) \sum_t \langle u - p_t, \ell_t - m_t \rangle^2}$	Squint [Koolen & van Erven’15]
2^{nd}-order switching regret	$\frac{S \ln(NT)}{\eta} + \eta \sum_t \langle u_t, (\ell_t - m_t)^2 \rangle$	$\sqrt{S \ln(NT) \sum_t \langle u_t, (\ell_t - m_t)^2 \rangle}$	---

Extension:

A Unified General Framework for “learning the learning rate”

Examples in online linear optimization:

Easy	Difficult	Existing solutions
$\frac{d \ln T}{\eta} + \eta \sum_t \langle u, \ell_t - m_t \rangle^2$	$\sqrt{d \ln T \sum_t \langle u, \ell_t - m_t \rangle^2}$	[Cutkosky and Orabona’18] $(m_t = 0)$
$\frac{\|u\|^2}{\eta} + \eta \sum_t \|\ell_t - m_t\|^2$	$\|u\| \sqrt{\sum_t \|\ell_t - m_t\|^2}$	[Cutkosky’19]
$\frac{u^\top (I + L)^{1/2} u}{\eta} + \eta \mathrm{tr}(L)^{1/2}$	$\sqrt{u^\top (I + L)^{1/2} u \cdot \mathrm{tr}(L)^{1/2}}$	[Cutkosky’ 20] $(m_t = 0)$
	Best of the above three bounds	[Cutkosky’19] only for unconstrained settings

The Meta Algorithm

Base algorithm (η): achieving $\mathcal{O}\left(\frac{\mathsf{KL}(u,\pi)}{\eta} + \eta\sum_t \langle u, \ell_t^2\rangle\right) - 8\eta\sum_t \langle p_t, \ell_t^2\rangle$ against all u

useful negative regret

Meta algorithm: treating $\eta = 1, \frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{T}$ base algorithms as experts $(N = \log T)$

cancel

Regret against the best base algorithm (say η^*):

$$\sqrt{\log(\log T)\sum_{t=1}^T \langle p_t^{\eta^*}, \ell_t\rangle^2} \leq \frac{\log(\log T)}{\eta^*} + \eta^*\sum_{t=1}^T \langle p_t^{\eta^*}, \ell_t\rangle^2$$

Overall regret: $\mathcal{O}\left(\frac{\log(\log T) + \mathsf{KL}(u,\pi)}{\eta^*} + \eta^*\sum_t \langle u, \ell_t^2\rangle\right) \approx \sqrt{\mathsf{KL}(u,\pi)\sum_t \langle u, \ell_t^2\rangle}$