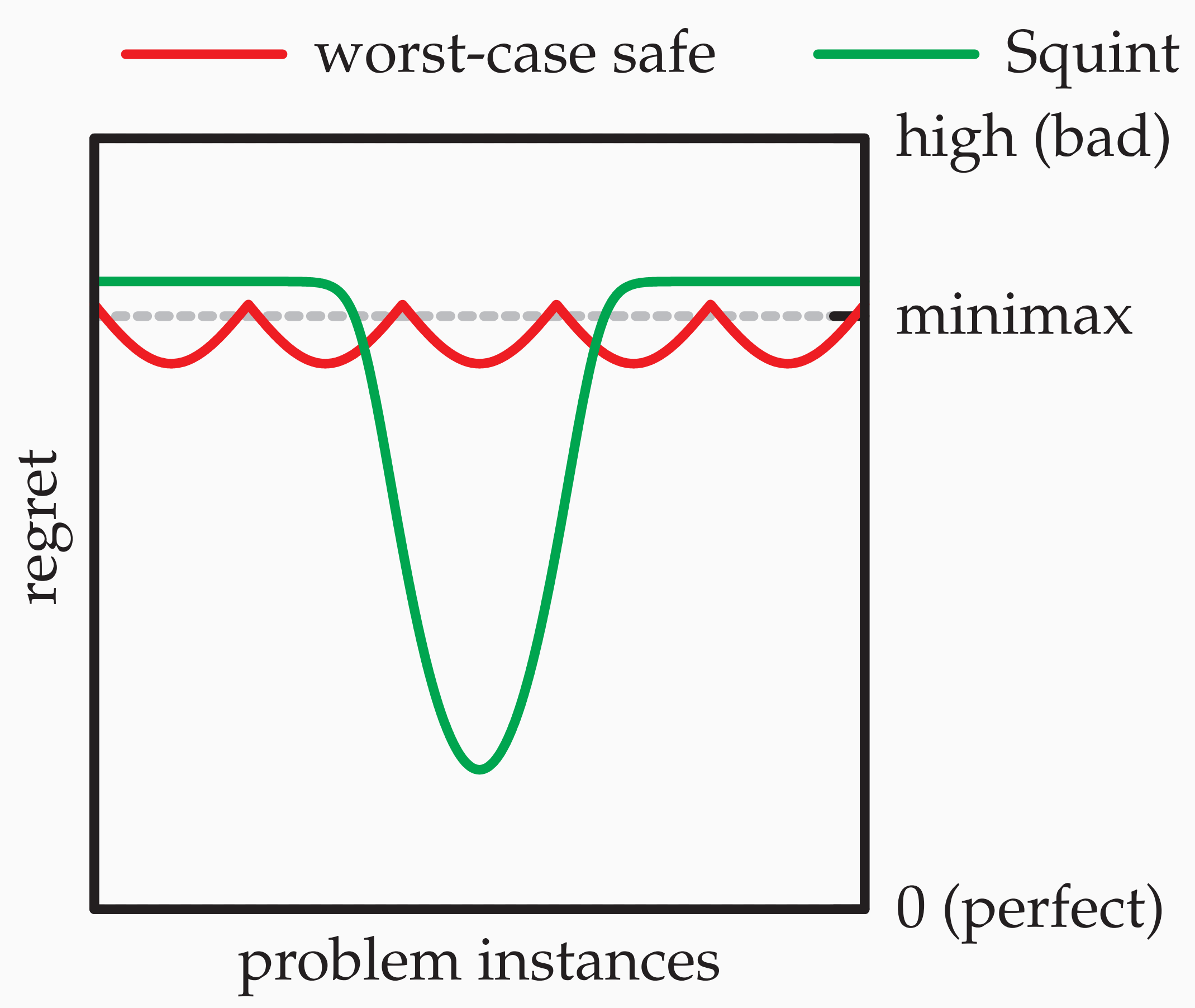
Second-order Quantile Methods for Experts and Combinatorial Games

Wouter M. Koolen Tim van Erven

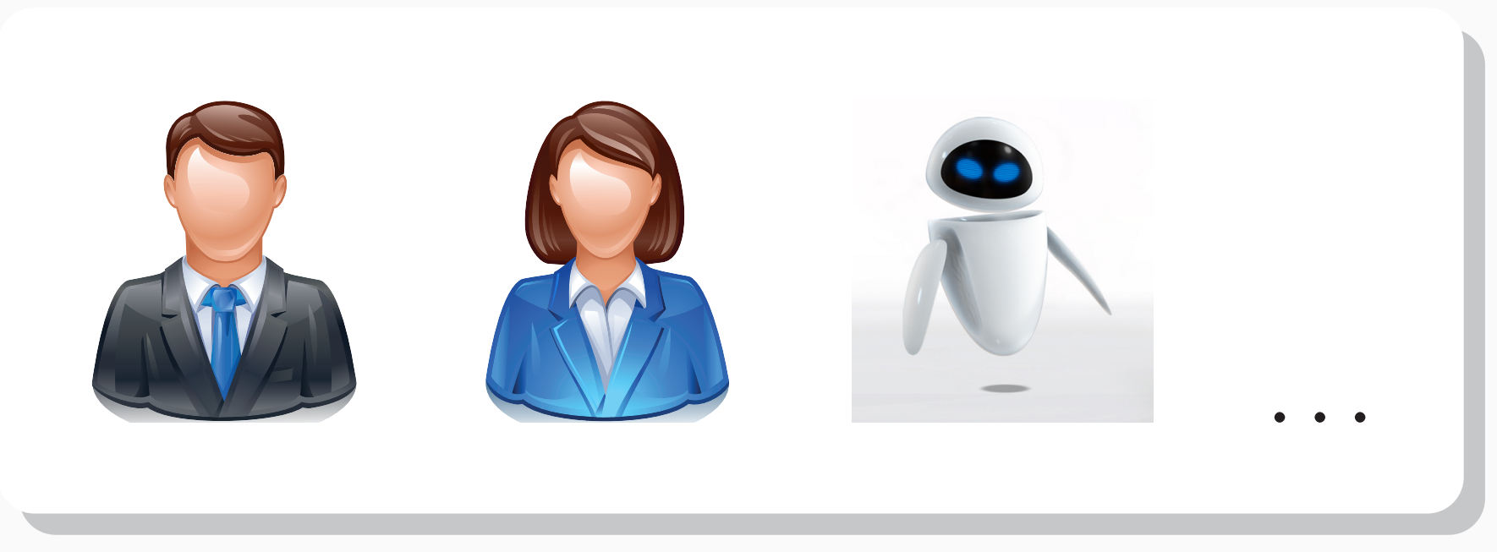
Our new algorithm **Squint**

- adapts to the difficulty of the learning problem by **learning the learning rate**,
- thereby integrating both the popular **second-order** and **quantile** adaptivities,
- at the **run time** of standard Hedge.



Hedge setting

K experts



In round $t = 1, 2, \ldots$

- Learner plays a probability distribution $\boldsymbol{w}_t = (w_t^1, \ldots, w_t^K)$ on experts
- Adversary reveals the expert loss vector $\boldsymbol{\ell}_t = (\ell_t^1, \ldots, \ell_t^K) \in [0,1]^K$

- Learner incurs loss $\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t$

The goal is to have small **regret**

$$R_T^k := \underbrace{\sum_{t=1}^T \boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t}_{\text{Learner}} - \underbrace{\sum_{t=1}^T \ell_t^k}_{\text{Expert } k}$$

with respect to every expert k at every time T.

Classic Hedge result

The **Hedge** algorithm with **learning rate** η

$$w_{t+1}^k := \frac{e^{-\eta L_t^k}}{\sum_k e^{-\eta L_t^k}} \quad \text{where} \quad L_t^k = \sum_{s=1}^t \ell_s^k,$$

upon proper tuning of η ensures

$$R_T^k \prec \sqrt{T \ln K} \qquad \text{for each expert } k.$$

Tight for adversarial (worst-case) losses.
Underwhelming in practice.

Second-order adaptivity

Cesa-Bianchi, Mansour, and Stoltz 2007, Hazan and Kale 2010, Chiang, Yang, Lee, Mahdavi, Lu, Jin, and Zhu 2012, De Rooij, Van Erven, Grünwald, and Koolen 2014, Gaillard, Stoltz, and Van Erven 2014, Steinhardt and Liang 2014

$$R_T^k \prec \sqrt{V_T^k \ln K} \qquad \text{for each expert } k.$$

for some second-order $V_T^k \leq L_T^k \leq T$

- stochastic case, learning sub-algorithms
- specialized algorithms, hard-coded $\ln K$.

Quantile adaptivity

Hutter and Poland 2005, Chaudhuri, Freund, and Hsu 2009, Chernov and Vovk 2010, Luo and Schapire 2014

Prior π on experts:

$$\min_{k \in \mathcal{K}} R_T^k \prec \sqrt{T(-\ln \pi(\mathcal{K}))} \qquad \text{for each subset } \mathcal{K} \text{ of experts}$$

- over-discretization, company baseline
- specialized algorithms, hard-coded T. "impossible tunings". efficiency.

Squint guarantees both

Squint algorithm with bound

$$R_T^{\mathcal{K}} \prec \sqrt{V_T^{\mathcal{K}}(-\ln \pi(\mathcal{K}) + C_T)} \qquad \text{for each subset } \mathcal{K} \text{ of experts}$$

where $R_T^{\mathcal{K}} = \mathbb{E}_{\pi(k|\mathcal{K})} R_T^k$ and $V_T^{\mathcal{K}} = \mathbb{E}_{\pi(k|\mathcal{K})} V_T^k$ denote the average (under the prior π) among the reference experts $k \in \mathcal{K}$ of the cumulative regret $R_T^k = \sum_{t=1}^T r_t^k$ and the (uncentered) variance of the excess losses $V_T^k = \sum_{t=1}^T (r_t^k)^2$ (where $r_t^k = (\boldsymbol{w}_t - \boldsymbol{e}_k)^\intercal \boldsymbol{\ell}_t$).

pretty two-line proof

Squint potential motivation

Fix prior π on experts $k \in \{1, \ldots, K\}$ and prior γ on learning rates $\eta \in [0, 1/2]$.

Potential function (weighted sum of objectives)

$$\Phi_t := \mathop{\mathbb{E}}_{\pi(k)\gamma(\eta)} \left[e^{\eta R_t^k - \eta^2 V_t^k} \right],$$

and associated weights

$$w_{t+1}^k := \frac{\pi(k) \, \mathbb{E}_{\gamma(\eta)} \left[e^{\eta R_t^k - \eta^2 V_t^k} \eta \right]}{\text{normalisation}}.$$

$1 = \Phi_0 \geq \Phi_1 \geq \Phi_2 \geq \cdots$

$e^{x - x^2} \leq 1 + x$ for $x \geq -1/2$

$$\Phi_{t+1} - \Phi_t = \mathop{\mathbb{E}}_{\pi(k)\gamma(\eta)} \left[e^{\eta R_t^k - \eta^2 V_t^k} \left(e^{\eta r_{t+1}^k - (\eta r_{t+1}^k)^2} - 1 \right) \right] \leq \mathop{\mathbb{E}}_{\pi(k)\gamma(\eta)} \left[e^{\eta R_t^k - \eta^2 V_t^k} \eta (\boldsymbol{w}_{t+1} - \boldsymbol{e}_k)^\intercal \boldsymbol{\ell}_{t+1} \right] = 0$$

by the choice of weights $\boldsymbol{w}_{t+1}$

Future work

Loss range adaptivity, bandits, online convex optimization

Regret guarantee

Choose k and fix $\hat{\eta} = \frac{R_T^k}{2V_T^k}$. Now as

$$1 \geq \Phi_T \geq \pi(k)\gamma(\hat{\eta}) e^{\hat{\eta} R_T^k - \hat{\eta}^2 V_T^k} = \pi(k)\gamma(\hat{\eta}) e^{\frac{(R_T^k)^2}{4V_T^k}}$$

we have

$$R_T^k \leq 2\sqrt{V_T^k(-\ln \pi(k) - \ln \gamma(\hat{\eta}))}.$$

For quantile bound take $\sum_{k \in \mathcal{K}}$

Three priors

Idea: have prior $\gamma(\eta)$ put sufficient mass around optimal $\hat{\eta}$

1. Uniform prior (generalizes to conjugate)
   $$\gamma(\eta) = 2$$
   Efficient algorithm, $C_T = \ln V_T^{\mathcal{K}}$.
2. Chernov&Vovk (2010) prior
   $$\gamma(\eta) = \frac{\ln 2}{\eta \ln^2(\eta)}$$
   Not efficient, $C_T = \ln \ln V_T^{\mathcal{K}}$.
3. Improper(!) log-uniform prior
   $$\gamma(\eta) = 1/\eta$$
   Efficient algorithm, $C_T = \ln \ln T$.

Squint with log-uniform prior

Closed-form expression for weights:

$$w_{t+1}^k \propto \pi(k) \int_0^{1/2} e^{\eta R_t^k - \eta^2 V_t^k} \eta \frac{1}{\eta} \, \mathrm{d}\eta$$

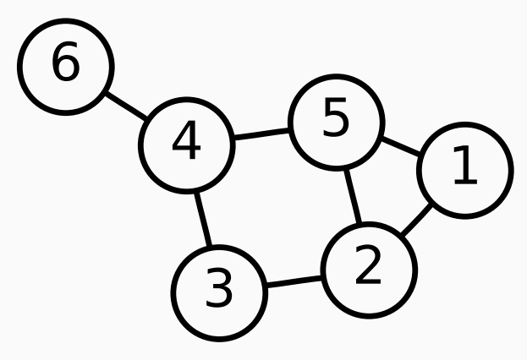
$$\propto \pi(k) e^{\frac{(R_t^k)^2}{4V_t^k}} \frac{\operatorname{erf}\left(\frac{R_t^k}{2\sqrt{V_t^k}}\right) - \operatorname{erf}\left(\frac{R_t^k - V_t^k}{2\sqrt{V_t^k}}\right)}{\sqrt{V_t^k}}$$

constant time per expert per round

Extensions

Combinatorial concept class $\mathcal{C} \subseteq \{0,1\}^K$:

- Shortest path
- Spanning trees
- Permutations



Component Squint guarantees:

$$R_T^{\boldsymbol{u}} \prec \sqrt{V_T^{\boldsymbol{u}}(\operatorname{comp}(\boldsymbol{u}) + KC_T)} \qquad \text{for each } \boldsymbol{u} \in \operatorname{conv}(\mathcal{C}).$$

The reference set of experts $\mathcal{K}$ is subsumed by an "average concept" vector $\boldsymbol{u} \in \operatorname{conv}(\mathcal{C})$, for which our bound relates the coordinate-wise average regret $R_T^{\boldsymbol{u}} = \sum_{t,k} u_k r_t^k$ to the averaged variance $V_T^{\boldsymbol{u}} = \sum_{t,k} u_k (r_t^k)^2$ and the prior entropy $\operatorname{comp}(\boldsymbol{u})$.

No range factor
Drop-in replacement for Component Hedge
Koolen, Warmuth, and Kivinen 2010 with same run-time