

# Open Problem: Shifting Experts on Easy Data

**Manfred K. Warmuth\***

*UC Santa Cruz, USA*

MANFRED@CSE.UCSC.EDU

**Wouter M. Koolen**

*Queensland University of Technology, Brisbane, Australia*

WOUTER.KOOLEN@QUT.EDU.AU

## Abstract

A number of online algorithms have been developed that have small additional loss (regret) compared to the best “shifting expert”. In this model, there is a set of experts and the comparator is the best partition of the trial sequence into a small number of segments, where the expert of smallest loss is chosen in each segment. The regret is typically defined for worst-case data / loss sequences.

There has been a recent surge of interest in online algorithms that combine good worst-case guarantees with much improved performance on easy data. A practically relevant class of easy data is the case when the loss of each expert is iid and the best and second best experts have a gap between their mean loss. In the full information setting, the FlipFlop algorithm by [De Rooij et al. \(2014\)](#) combines the best of the iid optimal Follow-The-Leader (FL) and the worst-case-safe Hedge algorithms, whereas in the bandit information case SAO by [Bubeck and Slivkins \(2012\)](#) competes with the iid optimal UCB and the worst-case-safe EXP3.

We ask the same question for the shifting expert problem. First, we ask what are the simple and efficient algorithms for the shifting experts problem when the loss sequence in each segment is iid with respect to a fixed but unknown distribution. Second, we ask how to efficiently unite the performance of such algorithms on easy data with worst-case robustness.

A particular intriguing open problem is the case when the comparator shifts within a small subset of experts from a large set under the assumption that the losses in each segment are iid.

**Introduction** Much of the work on online learning algorithms has been done in the realm of learning well compared to a set of experts. In each trial, the learner (probabilistically) chooses an expert from the set, and at the end of the trial, the expert chosen by the learner as well as all other experts incur a loss. The goal is to find learning algorithms that have small *regret*, which is defined as the total loss of the algorithm minus the loss of the best strategy in a comparator class. In the base case, the comparator class is simply the set of experts and many algorithms have been developed for this case: The Weighted Majority and Hedge algorithms use weights that decay exponentially with the loss of each expert leading to multiplicative updates ([Littlestone and Warmuth, 1994](#); [Freund and Schapire, 1997](#)), and the Follow the Perturbed Leader algorithm perturbs the losses by additive noise and then chooses the expert of minimum perturbed loss ([Kalai and Vempala, 2005](#); [Hutter and Poland, 2005](#)). All these algorithms are designed to achieve optimal regret when the

---

\* The author was supported by NSF grant IIS-0917397

losses of the experts are chosen by an adversary and the regret depends on the loss of the best comparator.

The key to achieving small worst-case expected regret is that the algorithms make a probabilistic choice of the expert in each trial. Deterministic algorithms can be fooled with an alternating loss pattern. This focus on the worst-case lead to some elegant algorithms including an efficient minimax algorithm based on random walks (Abernethy and Warmuth, 2010). However the worst-case assumption might be overly pessimistic. If the losses are iid, then the simple deterministic Follow the Leader (FL) algorithm is already minimax optimal (Kotłowski, 2013), and thus for most natural easy data, FL already works well and the fancy algorithms that guard against the worst-case are not needed.

**Shifting** The problem becomes more challenging when we model data that changes with time. Now the comparator consists of a partition of the trials into segments and a choice of a best expert in each segment. Again the goal is to achieve small worst-case regret bounds with this richer class of comparators and an interesting set of new algorithm has been developed for this task (Herbster and Warmuth, 1998, 2001; Bousquet and Warmuth, 2002; Koolen et al., 2012), most notably multiplicative update plus a mixing update.

Much of the literature on shifting experts deals with Vovk’s mixable loss functions (Vovk, 1998). Here the difference between iid and worst-case losses is perhaps insignificant, as the algorithms manage to balance the regret between all outcome sequences. However, in the linear/absolute/dot loss case this equalization property is lost, and a significant regret reduction can be achieved on some sequences. Worst-case bounds for the shifting experts problem with the dot loss have been obtained in Cesa-Bianchi et al. (2012).

**Problem: Efficient low regret algorithms and bounds for iid shifting experts**

We propose to focus on the shifting expert setting when the losses in each segment are iid with a gap, i.e. the mean loss of the best and the mean loss of the next best expert differ by at least some fixed constant. There are many good simple algorithms for this setting and a first goal is to delineate between them in terms of their regret bounds. When a new section starts, then the algorithm must be able to recover the new best expert. Each of the below algorithm has a mechanism for doing that.

- FL on the best current partition over  $n$  experts with  $k$  segments. Using dynamic programming, it takes  $O(nk)$  time to find the best such partition. In the simplest case,  $k$  is known to the algorithm.
- FL on a shifting window of length  $w$  ending with the last trial. This requires  $O(nw)$  storage between trials.
- FL on the total losses, where after each trial the total loss  $L_i$  of each expert  $i$  is reset to  $\min(L_i, w + \min_j L_j)$ , ie the total losses are kept at most  $w$  apart. This requires  $O(n)$  storage.
- FL on the total losses of each expert where the losses in the  $q$ -th last trial are multiplied by an exponentially decaying factor  $\beta^q$ . This requires  $n$  totals to be maintained.
- Use the weight updates designed for the worst case (i.e. a multiplicative update plus a mixing update, which typically requires  $2n$  weights). However instead of drawing

an expert from the current weight vector, the learner predicts with the expert of maximum weight.

The resulting regret bounds for these algorithms should be  $\log n$  per segment under the assumption that there is fixed gap between the mean loss of the best and the mean loss of the second best in each segment.

**Problem: Long-term memory for iid shifting experts** One of the most interesting cases of expert algorithms is the long term memory setting where the best expert shifts within a small subset. In the worst case setting, this was solved with a multiplicative update followed by certain mixing updates (Bousquet and Warmuth, 2002; Koolen et al., 2012). What are the optimal algorithms for the long-term memory setting when the segments are assumed to be iid? Intuitively the algorithm has to locally follow the leader but help with recovering to the new local leader when the (hidden) segment changes. It also has to recover faster towards experts that have done well in the past (long term memory).

If  $n$  is the total number of experts and  $s$  the size of the relevant subset, then the question is whether there is an algorithm that achieves expected regret  $\log n$  per member in the subset and  $\log s$  per shift within the subset. Note that this is a version of Yoav Freund’s (Freund, 2000) original open problem which initiated the long-term memory work of Bousquet and Warmuth (2002) that focused on the worst-case.

**Problem: IID and worst-case-optimal shifting experts** Ideally we want a single algorithm that can handle both the malignant worst-case and the benevolent iid case. For the base problem, this has been done in a whole line of research by fancy tunings of the Hedge algorithm (see Van Erven et al., 2011; De Rooij et al., 2014) and most recently with the FL algorithm applied to the dropout perturbed losses (Van Erven et al., 2014). In the bandit feedback model, the best of both worlds was achieved by Bubeck and Slivkins (2012). However it is not at all clear how these algorithm can be adapted to the shifting expert setting, let alone the long-term memory scenario.

**Wrap-up** The importance of this open problem stems directly from the practical abundance of non-stationary data, and the desire to learn from it using online algorithms. We hope to have sketched an important family of problems. The details of the iid shifting experts problem can be worked out in many ways. One particular choice is whether we want to compare on each segment with the loss of the best expert on the sampled data or in expectation. Another is whether the segmentation is itself generated randomly, or chosen adversarially. In the simplest scenario, the algorithm knows the parameters of the problem like the gap, the number of trials, the number of segments in the comparator partition, and the number of distinct best experts. The ultimate online algorithm should of course adapt to these practically unknown parameters on the fly.

## References

- Jacob Abernethy and Manfred K. Warmuth. Repeated games against budgeted adversaries. In *NIPS*, pages 1–9, 2010.
- Olivier Bousquet and Manfred K. Warmuth. Tracking a small set of experts by mixing past posteriors. *Journal of Machine Learning Research*, 3:363–396, 2002.

- Sébastien Bubeck and Aleksandrs Slivkins. The best of both worlds: Stochastic and adversarial bandits. *Journal of Machine Learning Research - Proceedings Track*, 23:42.1–42.23, 2012.
- Nicolò Cesa-Bianchi, Pierre Gaillard, Gábor Lugosi, and Gilles Stoltz. Mirror Descent meets Fixed Share (and feels no regret). In *NIPS Proceedings*, pages 989–997, 2012.
- Steven de Rooij, Tim van Erven, Peter D. Grünwald, and Wouter M. Koolen. Follow the leader if you can, Hedge if you must. *Journal of Machine Learning Research*, 2014. To appear.
- Yoav Freund. Private communication, 2000.
- Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55:119–139, 1997.
- M. Herbster and M. K. Warmuth. Tracking the best expert. *Machine Learning*, 32:151–178, 1998.
- Mark Herbster and Manfred K. Warmuth. Tracking the best linear predictor. *Journal of Machine Learning Research*, 1:281–309, 2001.
- Marcus Hutter and Jan Poland. Adaptive online prediction by following the perturbed leader. *Journal of Machine Learning Research*, 6:639–660, 2005.
- Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005. ISSN 0022-0000.
- Wouter M. Koolen, Dmitri Adamskiy, and Manfred K. Warmuth. Putting Bayes to sleep. In *Advances in Neural Information Processing Systems (NIPS) 25*, pages 135–143, December 2012.
- Wojciech Kotłowski. Follow the leader bound in the iid case. Private communication, September 2013.
- Nick Littlestone and Manfred K. Warmuth. The Weighted Majority algorithm. *Information and Computation*, 108(2):212–261, 1994.
- Tim van Erven, Steven de Rooij, Wouter M. Koolen, and Peter Grünwald. Adaptive Hedge. In *Advances in Neural Information Processing Systems (NIPS) 24*, pages 1656–1664, December 2011.
- Tim van Erven, Wojciech Kotłowski, and Manfred K. Warmuth. Follow the leader with dropout perturbations. In *Proceedings of the 27th Annual Conference on Learning Theory*, 2014.
- Volodya Vovk. A game of prediction with expert advice. *Journal of Computer and System Sciences*, 56(2):153–173, 1998.