Tight Bounds for Online Learning with Adversarially Corrupted Experts

Anonymous authors
Paper under double-blind review

Abstract

We study learning from expert advice with corrupted feedback, where the learner observes only *noisy* losses but is evaluated on the underlying *true* losses. Unlike prior work on stochastic perturbations, we consider a fully adversarial model in which perturbations are arbitrary but bounded by a per-expert budget C. We establish the first minimax-optimal risk of order $\tilde{\Theta}(C \log K)$ for the corrupted-feedback setting, even when the corruption budget C is *unknown*. Our achievability result builds on a novel potential-based analysis of the Exponentially Weighted Average (EWA) algorithm together with an epoch-based scheduling scheme inspired by the doubling trick. By contrast, we prove that the FTRL-style self-confident EWA (Auer et al., 2002a), and more generally a broad class of FTRL-style adaptive schedules, suffers risk at least $\Omega(C^\beta)$ for any $1 < \beta < \frac{3}{2}$, highlighting a fundamental separation between risk evaluations on noisy and true losses. Finally, we extend our guarantees beyond the realizable case by establishing sublinear regret bounds. Together, these results establish a complete minimax theory for the corrupted-expert model, capturing both the optimal rates and the algorithmic principles.

1 Introduction

Learning from expert advice (Cesa-Bianchi & Lugosi, 2006; Shalev-Shwartz, 2007; Rakhlin et al., 2010; Wu et al., 2026) is one of the most studied paradigms in online learning, forming the basis of algorithms for sequential prediction, game theory, and adaptive decision-making. In its classical form, the learner competes with the best expert in hindsight under adversarially chosen losses, and tight guaranties are well understood. Yet this clean picture changes dramatically when the learner's observations are *corrupted*. In many realistic environments—from adversarial platforms and cybersecurity to noisy crowdsourcing—the feedback revealed to the learner may be an intentionally distorted version of the true losses, while performance is ultimately evaluated with respect to those hidden truths.

To further motivate our setting, consider online binary classification under *Massart noise* (Ben-David et al., 2009). Let $\mathcal{H} = \{h_1, \ldots, h_K\} \subset \{0,1\}^{\mathcal{X}}$ be a hypothesis class of size K, and let Nature fix an unknown target $h_{k^*} \in \mathcal{H}$. At each round $t = 1, \ldots, T$: (i) a feature $\mathbf{x}_t$ is revealed, (ii) the learner predicts $\hat{y}_t \in \{0,1\}$, and (iii) a *noisy* label $\tilde{y}_t = \mathsf{Bernoulli}(\gamma_t) \oplus h_{k^*}(\mathbf{x}_t)$ is observed, where $\gamma_t \le \gamma < \frac{1}{2}$ for some *unknown* γ_t (but the upper bound γ is known). The learner aims to minimize the cumulative loss with respect to the *true* labels:

$$\tilde{r}_T(\mathcal{H}, \gamma) := \sup_{\mathbf{x}^T, h_{k^*}, \gamma^T} \mathbb{E}\left[\sum_{t=1}^{T} \mathbf{1}\{h_{k^*}(\mathbf{x}_t) \neq \hat{y}_t\}\right].$$

It was shown in Ben-David et al. (2009) that

$$\tilde{r}_T(\mathcal{H}, \gamma) \;\le\; \frac{\log|\mathcal{H}|}{1 - 2\sqrt{\gamma(1-\gamma)}}. \tag{1}$$

Note that a key nontrivial aspect of the result in (1) is that the noise parameters γ_t are *unknown*, which prevents the learner from directly estimating the actual noise levels. This phenomenon was further generalized

by Wu et al. (2024) to more general stochastic noise mechanisms. However, a major limitation of Ben-David et al. (2009) and Wu et al. (2024) is that the noise upper bound γ remains assumed to be *known* to the learner, an assumption that is often unrealistic in practice.

To address these limitations and to capture scenarios where stochastic assumptions completely break down, Papagiannis et al. (2025) recently introduced a general framework (namely the so-called "nasty-noise" model) in which corruption can occur *arbitrarily*, without adhering to any stochastic mechanism, but subject only to a bounded corruption budget. Formally, at each time step t, Nature selects an arbitrary *true* loss vector $\ell_t \in [0,1]^K$ (assigning losses to K experts) but reveals only a perturbed *noisy* version $\tilde{\ell}_t \in [0,1]^K$. The choice of $(\ell_t, \tilde{\ell}_t)$ is *completely* arbitrary, except that: (i) for every expert, the discrepancy between its cumulative true and noisy losses is upper bounded by a parameter C (which may grow with the horizon T); and (ii) there exists at least one expert that incurs zero cumulative true loss (i.e., the well-specified case).

It was shown in Papagiannis et al. (2025) that if the corruption parameter C is *known*, then the minimax cumulative risk grows as $\Theta(C \log K)$, via an elimination-based algorithm. When C is *unknown*, however, Papagiannis et al. (2025) proved that the standard Exponentially Weighted Average (EWA) algorithm (Cesa-Bianchi & Lugosi, 2006) with a constant learning rate attains cumulative risk of order $O(C^2 + C \log K)$. This result was further employed in (Papagiannis et al., 2025, Theorem 3) to derive a constant-risk bound for the Massart-noise model *without* assuming knowledge of the noise upper bound γ.[1] In addition, they showed an $\Omega(C^2/\log K)$ lower bound for EWA with constant learning rates. These findings leave open the following fundamental questions:

> *Can one achieve the minimax risk of order $O(C \log K)$ without knowing the parameter C? If so, can this optimal rate be obtained by the EWA algorithm with an adaptive schedule of learning rates based on the observed (noisy) losses?*

Summary of results. In this paper, we provide a complete resolution to the open questions above, which were also explicitly posed in Papagiannis et al. (2025). Specifically, we show (in Theorem 1) that for any learning rate η, the risk (on true losses) of the EWA algorithm is bounded by $O\left(\frac{\log K}{\eta} + \eta C(C+1) + C \log K\right)$, which is $O(\log K/\eta + \eta C^2 + C \log K)$ in the nontrivial regime $C \geq 1$. By choosing $\eta = 1/C$ when $C \geq 1$, this recovers the $O(C \log K)$ bound of Papagiannis et al. (2025) for the case of *known* C. Crucially, however, our result holds for *any* learning rate η, rather than only for given constants. The proof relies on a novel algebraic variant of the "elliptic potential lemma" (Cesa-Bianchi & Lugosi, 2006), which allows us to tightly control the discrepancies between risks on noisy and true losses. This technique is of independent interest.

Note that Theorem 1 still requires tuning the parameter η, which is infeasible when C is unknown. To address this, we propose a novel epoch-based algorithm (see Algorithm 2 and Theorem 3) that employs a "guess-and-double" strategy for handling the unknown corruption budget. This achieves a risk upper bound of order $O((C+1)\log(e+C)\log K)$ in the case of *unknown* C, matching the general $\Omega(C \log K)$ lower bound up to only a logarithmic factor for $C \geq 1$, and substantially improving upon the $O(C^2 + C \log K)$ upper bound from Papagiannis et al. (2025) for large C. We also revisit the one-sided relaxation used in (Papagiannis et al., 2025, Corollary 1). Even after strengthening that condition to hold at every prefix, Proposition 1 gives only an $O((C+1)\log^2 K)$-type EWA guarantee, while Remark 1 shows that the extra logarithmic factor is unavoidable already when $C = 1$. Thus the one-sided claim in (Papagiannis et al., 2025, Corollary 1) cannot hold without an additional factor of $\log K$.

Note that the "guess-and-double" strategy resembles the classical *doubling trick* in the online-learning literature, which *restarts* the algorithm whenever the current guess is violated. A natural question is whether a more *smoothed* adaptive schedule—such as the FTRL-style self-confident EWA of Auer et al. (2002a)—could attain the optimal $C \log K$ risk without restarts. Theorem 2 shows that the self-confident rule, and more generally FTRL-style schedules whose learning rates remain above a natural threshold, incur risk at least $\Omega(C^\beta)$ for any $1 < \beta < \frac{3}{2}$. The result does not apply to every FTRL schedule; rather, it identifies a broad and important class of smooth adaptive rules that *cannot* match the guarantee of our restart-based algo-

[1] In fact, any risk bound for the "nasty-noise" model *without* known C can be turned into such a guarantee via a black-box reduction.

rithm. Finally, we extend our results to scenarios where the true losses are *misspecified*, i.e., the cumulative true loss of the best expert is nonzero. We introduce a notion of *regret* with respect to the true losses,[2] and establish in Theorem 4 a sublinear regret bound $O\Big(\sqrt{(L^*_{\min} + C + C(C + L^*_{\min}))\ln K} + C\log K\Big)$, where $L^*_{\min}$ denotes the minimal cumulative true loss of the best expert. For $C \geq 1$, this simplifies to $O(\sqrt{C(C + L^*_{\min})\ln K} + C\log K)$. In Theorem 5 we show that Algorithm 2 also applies when C and $L^*_{\min}$ are unknown.

Highlights. In summary, our results:

(i) settle the minimax rate for learning with adversarially corrupted experts by achieving $\tilde{O}(C\log K)$ risk *without* prior knowledge of C;

(ii) reveal a fundamental separation between risks on noisy and true losses by showing the suboptimality of natural FTRL-style adaptive EWA schedules, including the classic self-confident rule;

(iii) extend the guarantees beyond the well-specified setting to obtain sublinear regret when the best expert incurs nonzero loss. Collectively, these contributions provide a complete minimax theory for corrupted-feedback expert advice.

In short, we present unifying new algorithmic techniques with structural insights of broader interest to adversarial online learning.

1.1 Other related works

Our framework is mostly comparable to Amir et al. (2020), which considers an adversarial corruption model with a budget on the cumulative *supremum* corruption across experts at each time step. This assumption is more stringent, as it directly yields a trivial upper bound on the discrepancy between noisy and true losses. In contrast, our model imposes a *per-expert* corruption budget, which is more flexible and motivated by the semi-stochastic noise models studied in Ben-David et al. (2009); Wu et al. (2024) (see also Bshouty et al. (2002)). In fact, in our model the contribution of corruption to the risk can even dominate the base risk attainable under noisy losses, which is the main nontriviality of our setting. Beyond corruption models, our setting also bears resemblance to *prediction with limited feedback*, as treated in (Cesa-Bianchi & Lugosi, 2006, Section 6) (see also Kakade & Kalai (2005); Wu et al. (2023). However, our "nasty expert" differs in that the noisy losses are selected *arbitrarily* (subject to certain constraints), rather than through a *time-independent* feedback function as in Cesa-Bianchi & Lugosi (2006). Prediction with *stochastic noise* was first investigated in Weissman et al. (2001), which studies sequential prediction with binary outcomes passing through a Binary Symmetric Channel (BSC). This setting was recently generalized in Bhatt & Kostina (2024; 2025) (see also references therein) under the assumption that the true losses are estimable. The regret analysis of *observable* losses using the Exponentially Weighted Average (EWA) algorithm has been widely studied; see Auer et al. (2002b); Cesa-Bianchi & Lugosi (2006); Gaillard et al. (2014); Mourtada & Gaïffas (2019).

2 Problem formulation and preliminaries

Let K be the number of experts. We consider the following online learning framework, which operates in T rounds. At each time step $t = 1, 2, \ldots, T$, the following events occur sequentially: (i) Nature selects the *true* loss vector $\ell_t \in [0,1]^K$ but keeps it secret from the learner; (ii) the learner selects a distribution $\hat{p}_t \in \Delta([K])$ over experts, samples $\hat{k}_t \sim \hat{p}_t$, predicts $\hat{k}_t$, and incurs the loss $\ell_t[\hat{k}_t]$; (iii) Nature selects a *noisy* loss vector $\tilde{\ell}_t \in [0,1]^K$ and reveals it to the learner.

[2]Note that, both our notions of *risk* and *regret* are defined with respect to the true losses; the noisy losses serve only as observations and do not appear in the performance measure.

Let $C \leq T$ be an arbitrary parameter that controls the noise level, which may grow with respect to T. The goal of the learner is to design a prediction strategy $\hat{p}^T$ that minimizes:

$$\mathsf{risk}_{T,K,C} := \sup_{\ell^T, \tilde{\ell}^T} \mathbb{E}\left[\sum_{t=1}^{T} \ell_t[\hat{k}_t]\right] = \sup_{\ell^T, \tilde{\ell}^T} \sum_{t=1}^{T} \langle \hat{p}_t, \ell_t \rangle, \tag{2}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product, and $\ell^T := \{\ell_1, \ldots, \ell_T\}$, $\tilde{\ell}^T := \{\tilde{\ell}_1, \ldots, \tilde{\ell}_T\}$ are selected *arbitrarily* subject to the following constraints:

(A1) There *exists* $k^* \in [K]$ such that

$$\sum_{t=1}^{T} \ell_t[k^*] = 0; \tag{3}$$

(A2) For *all* $k \in [K]$, we have

$$\sum_{t=1}^{T} |\ell_t[k] - \tilde{\ell}_t[k]| \leq C. \tag{4}$$

Here, the first constraint ensures that there exists an expert incurring zero cumulative *true* loss (i.e., realizable), while the second constraint ensures that for any expert, the discrepancy between the true and noisy cumulative losses is upper bounded by C. Crucially, we assume that the parameter C is *unknown* to the learner. In Section 3.4, we will further relax the assumption (A1) by considering a *regret* formulation. However, we should emphasize that even with the realizable assumption, the problem remains fundamentally challenging and leads to unexpected results.

Algorithm 1 EWA algorithm with Noisy Losses

Input: Learning rate $\eta > 0$, number of experts K
Initialize weights $w_k^1 \leftarrow 1$ for all $k \in [K]$
for $t = 1, 2, \ldots, T$ **do**
  Compute probability distribution over experts:

$$\hat{p}_t[k] \leftarrow \frac{w_k^t}{\sum_{j=1}^{K} w_j^t} \quad \text{for } k \in [K]$$

  Sample $\hat{k}_t$ from the distribution $\hat{p}_t$
  Predict using expert $\hat{k}_t$
  Observe the noisy loss vector $\tilde{\ell}_t \in [0,1]^K$
  Update weights for all experts: $w_k^{t+1} \leftarrow w_k^t \cdot \exp(-\eta \tilde{\ell}_t[k]) \quad \text{for } k \in [K]$
end for

Exponentially Weighted Average (EWA). Arguably the most established approach to learning from expert advice is the Exponentially Weighted Average (EWA) algorithm (Cesa-Bianchi & Lugosi, 2006), given in Algorithm 1 for the case of observable noisy losses. The following risk bound on *noisy* losses is classical (Auer et al., 2002a).

Lemma 1 (Risk of EWA on Noisy Losses)**.** *Let $\tilde{\ell}_1, \ldots, \tilde{\ell}_T \in \mathbb{R}^K$ be arbitrary noisy loss vectors, and let $\hat{p}_1, \ldots, \hat{p}_T$ be the distributions produced by Algorithm 1. Then*

$$\sum_{t=1}^{T} \langle \hat{p}_t, \tilde{\ell}_t \rangle \;\leq\; O\Big((\eta+1)\tilde{L} + \tfrac{\log K}{\eta}\Big) \leq O\left((\eta+1)C + \frac{\log K}{\eta}\right),$$

where $\tilde{L} := \min_{k \in [K]} \sum_{t=1}^{T} \tilde{\ell}_t[k] \leq C$ denotes the minimal cumulative noisy loss.

In order to convert such a risk bound to the *true* loss, we consider the following decomposition:

$$\sum_{t=1}^{T}\langle \hat{p}_t, \ell_t\rangle = \underbrace{\sum_{t=1}^{T}\langle \hat{p}_t, \tilde{\ell}_t\rangle}_{\text{Algorithmic (noisy) risk}} + \underbrace{\sum_{t=1}^{T}\langle \hat{p}_t, \ell_t - \tilde{\ell}_t\rangle}_{\text{Discrepancy between true and noisy risks}} . \tag{5}$$

Note that, in contrast to the cumulative *supremum* corruption considered in Amir et al. (2020) [3], our condition (4) does *not* guarantee that the discrepancy term is bounded by C. In fact, the discrepancy may vary significantly across algorithms (i.e., the choices of $\hat{p}_t$), and controlling this behavior is the central technical challenge of our analysis.

To gain intuition about the behavior of the risk on *true* losses, we present the following lower bound (see also Papagiannis et al. (2025)) which – for completeness – we prove in Appendix A.

Lemma 2 (Lower Bound on True Risk)**.** *For any prediction rule $\hat{p}_1, \ldots, \hat{p}_T$ (with or without knowledge of C), there exist true and noisy loss sequences $\ell_1, \ldots, \ell_T$ and $\tilde{\ell}_1, \ldots, \tilde{\ell}_T$ satisfying conditions (3) and (4) such that $\sum_{t=1}^{T}\langle \hat{p}_t, \ell_t\rangle \;\geq\; \Omega(C \cdot \log K)$.*

If we take $\eta = \sqrt{\log K/C}$ in Lemma 1, the risk of EWA on *noisy* losses scales as $\sqrt{C \cdot \log K} + C$. This is of lower asymptotic order than the lower bound in Lemma 2, implying that the discrepancy term can, in fact, be the *dominant* contribution to the true risk. Perhaps surprisingly, (Papagiannis et al., 2025, Lemma 4) further shows that when $\eta = 1$ (or any constant), the true risk of EWA is bounded below by $\Omega\Big(\frac{C^2}{\log K}\Big)$, even though the noisy risk is only $O(C + \log K)$.

3 Main results

In this section, we present our main results, which settle the minimax risk for the adversarially corrupted online learning model *without* prior knowledge of the corruption budget C.

3.1 A tighter analysis of EWA algorithm

We start with the following key risk bound on true losses for EWA.

Theorem 1. *Let $\ell_1, \ldots, \ell_T \in \mathbb{R}^K$ be arbitrary true loss vectors and $\tilde{\ell}_1, \ldots, \tilde{\ell}_T \in \mathbb{R}^K$ be corresponding noisy loss vectors satisfying conditions (3) and (4). Let $\hat{p}_1, \ldots, \hat{p}_T$ be the distributions produced by EWA Algorithm 1 with any learning rate $\eta > 0$. Then with $\tilde{L} = \min_k \sum_{t=1}^{T} \tilde{\ell}_t[k]$*

$$\sum_{t=1}^{T}\langle \hat{p}_t, \ell_t\rangle \;\leq\; O\Big(\tfrac{\ln K}{\eta} + \eta C(\tilde{L}+1) + C\ln K\Big),$$

Proof. Using the decomposition in (5), the expected true risk can be written as

$$\begin{aligned}
\sum_{t\in[T]} \frac{\sum_{i\in[K]} w_i^t \ell_t[i]}{\sum_{i\in[K]} w_i^t} &= \sum_{t\in[T]} \frac{\sum_{i\in[K]} w_i^t \tilde{\ell}_t[i]}{\sum_{i\in[K]} w_i^t} + \sum_{t\in[T]} \frac{\sum_{i\in[K]} w_i^t (\ell_t[i] - \tilde{\ell}_t[i])}{\sum_{i\in[K]} w_i^t} \\
&\overset{(a)}{\leq} O\left(\eta C + \frac{\ln K}{\eta}\right) + \sum_{t\in[T]} \frac{\sum_{i\in[K]} w_i^t (\ell_t[i] - \tilde{\ell}_t[i])}{\sum_{i\in[K]} w_i^t} \\
&\leq O\left(\eta C + \frac{\ln K}{\eta}\right) + \sum_{k\in[K]} \max_{t\in[T]} \frac{w_k^t}{\sum_{i\in[K]} w_i^t} \sum_{t\in[T]} |\ell_t[k] - \tilde{\ell}_t[k]| \\
&\overset{(b)}{\leq} O\left(\eta C + \frac{\ln K}{\eta}\right) + C \sum_{k\in[K]} \max_{t\in[T]} \frac{w_k^t}{\sum_{i\in[K]} w_i^t},
\end{aligned} \tag{6}$$

[3] They assume $\sum_{t=1}^{T} \|\ell_t - \tilde{\ell}_t\|_\infty \leq C$ for certain budget parameter C.

where (a) follows from Lemma 1 and (b) follows by condition (4).

Let $t_k = \arg\max_{t\in[T]} \frac{w_k^t}{\sum_{i\in[K]} w_i^t}$. Without loss of generality, assume that $t_1 \le t_2 \le \cdots \le t_{K-1}$. In addition, assume that expert K achieves the smallest cumulative noisy loss, i.e., $\sum_{t=1}^T \tilde{\ell}_t[K] = \min_k \sum_{t=1}^T \tilde{\ell}_t[k]$. Then, we have (by excluding the expert K, which contributes at most 1):

$$\begin{aligned}
\sum_{k\in[K-1]} \max_{t\in[T]} \frac{w_k^t}{\sum_{i\in[K]} w_i^t} &= \sum_{k\in[K-1]} \frac{w_k^{t_k}}{\sum_{i\in[K]} w_i^{t_k}} \overset{(a)}{\le} \sum_{k\in[K-1]} \frac{w_k^{t_k}}{\sum_{i\in[K-1]\setminus[k-1]} w_i^{t_i} + w_K^{t_k}} \\
&\overset{(b)}{\le} \sum_{k\in[K-1]} \frac{w_k^{t_k}}{\sum_{i\in[K-1]\setminus[k-1]} w_i^{t_i} + e^{-\eta\tilde{L}}} \\
&\overset{(c)}{\le} \sum_{k\in[K-1]} \ln \frac{\sum_{i\in[K-1]\setminus[k-1]} w_i^{t_i} + e^{-\eta\tilde{L}}}{\sum_{i\in[K-1]\setminus[k]} w_i^{t_i} + e^{-\eta\tilde{L}}} \\
&\le \ln \frac{e^{-\eta\tilde{L}} + \sum_{i\in[K-1]} w_i^{t_i}}{e^{-\eta\tilde{L}}} \le \ln K + \eta\tilde{L},
\end{aligned}$$

where (a) follows from the fact that $w_k^t \ge w_k^{t'}$ for $t < t'$, (b) follows since $w_K^t \ge e^{-\eta\tilde{L}}$ for all $t \in [T]$, and (c) follows since $x \le -\ln(1-x)$ for $x \in (0,1)$. Combining with (6), we obtain

$$\sum_{t\in[T]} \frac{\sum_{i\in[K]} w_i^t \ell_t[i]}{\sum_{i\in[K]} w_i^t} \le O\left(\eta C + \frac{\ln K}{\eta} + C\ln K + \eta C\tilde{L}\right),$$

as needed. □

The key difference between the true-risk bound in Theorem 1 and the classical noisy-risk bound in Lemma 1 is the additional $\eta C(\tilde{L}+1)$ term, which is $\eta C(C+1)$ in the worst case. This quadratic dependence for $C \ge 1$ is inherent for EWA: (Papagiannis et al., 2025, Lemma 4) gives an $\Omega(\eta C^2/\log K)$ lower bound for the true risk. For $\eta = 1$, Theorem 1 also recovers the risk bound of (Papagiannis et al., 2025, Lemma 3), which was previously proved only for constant learning rates. The log-potential argument in the proof of Theorem 1 is the main device that lets us control the discrepancy term; it is reminiscent of the classical elliptic-potential lemma (Cesa-Bianchi & Lugosi, 2006).

We next discuss a related one-sided condition. It was claimed in (Papagiannis et al., 2025, Corollary 1) that an $O(C^2 + C\log K)$-type EWA bound continues to hold under a signed one-sided corruption condition, as presented in (7) below. The next proposition shows that such conditions only admit an $O((C+1)\log^2 K)$-type bound. In Remark 1, we also shows that the additional factor of $\log K$ is not a proof artifact; thus, the assertion in (Papagiannis et al., 2025, Corollary 1) does not hold as stated.

Proposition 1 (Any-time signed one-sided corruption)**.** *Assume that $\ell_t, \tilde{\ell}_t \in [0,1]^K$ and that the following any-time signed one-sided conditions hold: there is a reference expert k^* such that*

$$\sum_{t=1}^T \tilde{\ell}_t[k^*] \le C, \qquad \sum_{s=1}^t \ell_s[k] - \sum_{s=1}^t \tilde{\ell}_s[k] \le C \quad (k \in [K],\ t \in [T]). \tag{7}$$

Then Algorithm 1 with learning rate $\eta > 0$ guarantees

$$\sum_{t=1}^T \langle \hat{p}_t, \ell_t \rangle \le O\left((\ln(eK) + \eta C)\left(C + 1 + \frac{\ln(eK)}{\eta} \right) \right).$$

In particular, choosing $\eta = 1/(C+1)$ gives $O((C+1)\ln^2(eK))$ risk.

Proof. We bound the true risk directly, rather than decomposing it into noisy risk and discrepancy as in Theorem 1. Let

$$L_t(k) := \sum_{s=1}^t \ell_s[k], \qquad \tilde{L}_t(k) := \sum_{s=1}^t \tilde{\ell}_s[k],$$

with $L_0(k) = \tilde{L}_0(k) = 0$. Since noisy losses are nonnegative, $\tilde{L}_t(k^*) \le C$ for every prefix t. The signed prefix condition gives

$$L_t(k) \le \tilde{L}_t(k) + C, \qquad k \in [K],\ t \in [T].$$

Fix an expert k and define the disjoint intervals $I_t(k) := (L_{t-1}(k), L_t(k)]$. For $u \ge 0$, let

$$P_k(u) := \max\{\hat{p}_t[k] : \tilde{L}_{t-1}(k) \ge u\},$$

where the maximum is defined to be zero if the set is empty. Then

$$\sum_{t=1}^{T} \hat{p}_t[k]\ell_t[k] = \sum_{t=1}^{T} \int_{I_t(k)} \hat{p}_t[k]\, dx.$$

The part of these intervals below level $C+1$ contributes at most $(C+1)P_k(0)$. For the remaining part, take $x \in I_t(k) \cap (C+1, \infty)$ and set $u := x - C - 1$. Since $x \le L_t(k) \le C + \tilde{L}_t(k)$ and $\tilde{\ell}_t[k] \le 1$, we have $u \le \tilde{L}_{t-1}(k)$. Hence $\hat{p}_t[k] \le P_k(u)$. Since the shifted intervals $\{(I_t(k) - C - 1) \cap (0, \infty)\}_t$ are disjoint,

$$\sum_{t=1}^{T} \hat{p}_t[k]\ell_t[k] \le (C+1)P_k(0) + \int_0^\infty P_k(u)\, du.$$

We now sum over experts. The log-potential estimate in the proof of Theorem 1, applied after relabeling $k^* = K$, gives

$$\sum_{k \ne k^*} P_k(0) \le \ln K + \eta C,$$

where the lower bound $\exp(-\eta \tilde{L}_{t-1}(k^*)) \ge \exp(-\eta C)$ replaces the best-noisy-expert term. Since $P_{k^*}(0) \le 1$, we have

$$\sum_{k=1}^{K} P_k(0) \le \ln(eK) + \eta C.$$

Set $M := \ln(eK) + \eta C$. Thus $\sum_k P_k(u) \le M$ for every $u \ge 0$. Also, if $\tilde{L}_{t-1}(k) \ge u$, then

$$\hat{p}_t[k] = \frac{e^{-\eta \tilde{L}_{t-1}(k)}}{\sum_i e^{-\eta \tilde{L}_{t-1}(i)}} \le e^{-\eta(u-C)},$$

again because $\tilde{L}_{t-1}(k^*) \le C$. Therefore

$$\sum_{k=1}^{K} P_k(u) \le \min\{M,\ Ke^{-\eta(u-C)}\}.$$

With $U := C + \ln(eK)/\eta$, integrating the last display gives

$$\int_0^\infty \sum_{k=1}^{K} P_k(u)\, du \le MU + \int_U^\infty Ke^{-\eta(u-C)}\, du \le M\left(C + \frac{\ln(eK)}{\eta}\right) + \frac{1}{e\eta}.$$

The last term is at most $M \ln(eK)/\eta$, because $M \ge 1$ and $\ln(eK) \ge 1$. Combining this with the $(C+1)M$ contribution from the low levels proves the proposition. □

Remark 1 (Any-time one-sided corruption costs an extra logarithm)**.** *The additional* $\log K$ *factor in Proposition 1 cannot be removed under the any-time signed one-sided condition. Indeed, any claim with only one logarithmic factor in* K *can fail under the any-time conditions*

$$\sum_{t=1}^{T} \tilde{\ell}_t[k^*] \le C, \qquad \sum_{s=1}^{t} \ell_s[k] - \sum_{s=1}^{t} \tilde{\ell}_s[k] \le C \quad (k \in [K],\ t \in [T]).$$

To see this, fix a large integer m, take $K = m^2 + 1$, $C = 1$, and let expert 1 have true and noisy loss zero throughout. Partition the other m^2 experts into groups $G_1, \dots, G_m$ of size m, and put $q := \lfloor \log m \rfloor$. First, for q rounds, assign noisy loss one and true loss zero to every expert in $\bigcup_j G_j$. Every such expert then has weight e^{-q} and signed cumulative discrepancy $-q$. Next process $G_1, \dots, G_m$ successively. While processing G_j, give each expert in G_j true loss one and noisy loss zero for q rounds. Afterwards, use additional rounds with true loss zero and noisy loss one on G_j until its total EWA weight is at most K^{-3}. All experts not mentioned on a round receive zero true and noisy loss.

At all times, every signed discrepancy is at most zero, and hence the displayed any-time conditions hold with $C = 1$. Write $\hat{p}_t[A] := \sum_{k \in A} \hat{p}_t[k]$. During the q costly rounds for group G_j, let $a := me^{-q} \in [1, e)$. The total weight of G_j is a, the unprocessed groups $G_{j+1}, \dots, G_m$ have total weight $(m-j)a$, the already processed groups have total weight at most $(j-1)K^{-3} \le mK^{-3}$, and the reference expert has weight one. Therefore, for all sufficiently large m,

$$\hat{p}_t[G_j] \ge \frac{a}{1 + (m-j+1)a + mK^{-3}} \ge \frac{1}{m-j+3} \ge \frac{1}{2(m-j+2)}.$$

On these q rounds, exactly the experts in G_j have true loss one, so the expected true loss on each such round is $\hat{p}_t[G_j]$. Summing over the costly rounds and then over groups gives

$$\Omega\left(q \sum_{j=1}^{m} \frac{1}{m-j+2} \right) = \Omega\big((\log m)^2\big) = \Omega\big((\log K)^2\big).$$

For $\eta = 1$ and $\tilde{L} = 0$, Theorem 1 would give only $O(\log K)$ when $C = 1$, whereas Proposition 1 gives the matching upper bound $O(\log^2 K)$ up to constants. Consequently, the one-sided relaxation in (Papagiannis et al., 2025, Corollary 1) is not sufficient for the claimed bound without an additional factor of $\log K$; even its any-time strengthening necessarily permits such a factor.

3.2 Fundamental limits of FTRL-style EWA with adaptive learning rates

In this section, we study follow-the-regularized-leader (FTRL)-style EWA variants with learning rates chosen based on past noisy losses. Define

$$\tilde{L}_t(k) := \sum_{s=1}^{t} \tilde{\ell}_s[k], \qquad \tilde{L}_t^* := \min_{k \in [K]} \tilde{L}_t(k).$$

The FTRL-style adaptive update is

$$\hat{p}_t[k] = \frac{\exp\big(-\eta_{t-1} \tilde{L}_{t-1}(k)\big)}{\sum_{u=1}^{K} \exp\big(-\eta_{t-1} \tilde{L}_{t-1}(u)\big)}, \qquad t \ge 2, \tag{8}$$

where η_{t-1} is selected based only on $\tilde{\ell}_1, \dots, \tilde{\ell}_{t-1}$ and $\hat{p}_1$ is arbitrary. We show that a broad class of FTRL schedules, including the self-confident rule of Auer et al. (2002a), is suboptimal when C is unknown.

Theorem 2 (FTRL-style adaptive-rate lower bound)**.** *Let $f_K : [0, \infty) \to (0, \infty)$ be nonincreasing, and suppose that the FTRL-style EWA update in (8) uses learning rates satisfying*

$$\eta_t \ge f_K(\tilde{L}_t^*) \qquad \text{for all } t.$$

Fix $1 < \beta < \frac{3}{2}$. If, for all sufficiently large K,

$$f_K(2K) \ge \frac{2 \log K}{K^{2-\beta}},$$

then, for corruption budget $C = 2K$, there is a choice of noisy and true loss sequences satisfying constraints (3) and (4) such that the FTRL-style algorithm incurs risk at least $\Omega(C^\beta)$.

In particular, the conclusion holds for every $1 < \beta < \frac{3}{2}$ *for the self-confident FTRL schedule*

$$\eta_t = \sqrt{\frac{\log K}{\max\{1, \tilde{L}_t^*\}}}.$$

Proof. **Construction.** Let $D := K^{2-\beta}$ and $M := K^{\beta-1}$, and assume for convenience that both quantities are integers; floors and ceilings only change constants. Let the special experts be $a_i := iD$, $i \in [M]$, and define a bad block of length K for each $2 \le i \le M-1$ by

$$B_i := \{K + (M-i)K + 1, \dots, K + (M-i+1)K\}.$$

Set $T := 2K + MK$. During the first K rounds, use the upper-triangle noisy-loss prefix

$$\tilde{\ell}_t[k] = \mathbf{1}\{t \ge k\}, \qquad t \le K.$$

After this prefix, every non-special coordinate has noisy loss one, whereas the noisy loss of each special coordinate is

$$\tilde{\ell}_t[a_i] = \begin{cases} 0, & K < t \le 2K + (M-i)K, \\ 1, & 2K + (M-i)K < t \le T. \end{cases}$$

The true losses are defined by

$$\ell_t[k] = \begin{cases} 0, & k = a_1, \\ 1, & k = a_i,\ t \in B_i,\ 2 \le i \le M-1, \\ \tilde{\ell}_t[k], & \text{otherwise.} \end{cases}$$

Properties of the construction. For $2 \le i \le M-1$, put $A_i := \{a_1, \dots, a_{i-1}\}$. We claim that the constructed sequences have the following properties:

(P1) Conditions (3) and (4) hold with corruption budget $C = 2K$.

(P2) For every $t \in B_i$, the experts in A_i are precisely the experts with zero true loss.

(P3) For every $t \in B_i$ and $a' \in A_i$,

$$\tilde{L}_{t-1}(a') - \tilde{L}_{t-1}(a_i) \ge D. \tag{9}$$

(P4) For every t in a bad block, $\tilde{L}_{t-1}^* \le 2K$.

We now verify these properties. First, a_1 has zero true loss throughout the construction. Its noisy loss is one on $K - D + 1$ prefix rounds and on at most K later rounds, so

$$\sum_{t=1}^{T} \left|\ell_t[a_1] - \tilde{\ell}_t[a_1]\right| \le (K - D + 1) + K \le 2K.$$

Every special expert a_i with $2 \le i \le M-1$ differs from its noisy loss only on its own block B_i, of length K; all remaining experts differ on no rounds. This proves (P1).

Next fix $2 \le i \le M-1$. On B_i, the coordinates $a_1, \dots, a_i$ have noisy loss zero, whereas each special coordinate a_j with $j > i$ and each non-special coordinate has noisy loss one. The construction changes a_i from noisy loss zero to true loss one on B_i, while $a_1, \dots, a_{i-1}$ retain true loss zero. Hence (P2) holds.

For $a' = a_j \in A_i$, the prefix satisfies

$$\tilde{L}_K(a') - \tilde{L}_K(a_i) = (K - a' + 1) - (K - a_i + 1) = a_i - a' \ge D.$$

Both a' and a_i have zero noisy loss from round $K+1$ through every round of B_i, so this gap persists until time $t-1$, proving (P3). Finally, since a_1 has cumulative noisy loss at most $2K$ throughout the construction,

$$\tilde{L}^*_{t-1} \le \tilde{L}_{t-1}(a_1) \le 2K,$$

which proves (P4).

Deriving the FTRL lower bound. Fix $i \in \{2, \dots, M-1\}$ and $t \in B_i$. By (P4), monotonicity of f_K, and the theorem's rate condition,

$$\eta_{t-1} \ge f_K(\tilde{L}^*_{t-1}) \ge f_K(2K) \ge \frac{2\log K}{D}.$$

Define the unnormalized FTRL weight

$$W_t(k) := \exp\bigl(-\eta_{t-1}\tilde{L}_{t-1}(k)\bigr).$$

Combining the preceding lower bound on η_{t-1} with (P3), for every $a' \in A_i$,

$$\frac{W_t(a')}{W_t(a_i)} = \exp\bigl(-\eta_{t-1}\left(\tilde{L}_{t-1}(a') - \tilde{L}_{t-1}(a_i)\right)\bigr) \le \exp(-2\log K) = K^{-2}.$$

It follows, after summing this bound over the truly good experts, that

$$W_t(A_i) := \sum_{a' \in A_i} W_t(a') \le |A_i| K^{-2} W_t(a_i).$$

The reference expert a_i is not in A_i, so its weight contributes an additional term to the normalizing denominator. Consequently,

$$\begin{aligned}\hat{p}_t[A_i] &= \frac{W_t(A_i)}{\sum_{u=1}^K W_t(u)} \le \frac{W_t(A_i)}{W_t(A_i) + W_t(a_i)} \\ &\le \frac{|A_i|K^{-2}W_t(a_i)}{|A_i|K^{-2}W_t(a_i) + W_t(a_i)} = \frac{|A_i|K^{-2}}{1 + |A_i|K^{-2}} \le \frac{K^{-1}}{1+K^{-1}} \le \frac{1}{2}.\end{aligned}$$

Here we used $|A_i| = i-1 \le M-1 \le K$, since $\beta < 3/2 < 2$. By (P2), all true-loss-zero mass on B_i is contained in A_i, and hence $\langle \hat{p}_t, \ell_t \rangle \ge 1/2$. Summing over the K rounds in each of the $M-2$ bad blocks gives

$$\sum_{t=1}^T \langle \hat{p}_t, \ell_t \rangle \ge \frac{K}{2}(M-2) = \Omega(K^\beta) = \Omega(C^\beta),$$

where the last equality uses (P1), namely $C = 2K$.

Self-confident schedule. For this schedule, (P4) gives, on each bad block,

$$\eta_{t-1} = \sqrt{\frac{\log K}{\max\{1, \tilde{L}^*_{t-1}\}}} \ge \sqrt{\frac{\log K}{2K}}.$$

Therefore,

$$\eta_{t-1} D \ge \frac{1}{\sqrt{2}} K^{3/2-\beta}\sqrt{\log K} \ge 2\log K$$

for all sufficiently large K, since $\beta < 3/2$. Thus the stated self-confident rule satisfies the required rate condition. □

Corollary 1 (Polynomially decaying FTRL rates)**.** *Suppose the FTRL-style update in (8) uses rates satisfying*

$$\eta_t \ge \frac{c}{(1+\tilde{L}^*_t)^\gamma}$$

for constants $c > 0$ and $\gamma < 1$. Then, for every $1 < \beta < \min\{3/2, 2-\gamma\}$ and all sufficiently large K, there are loss sequences with corruption budget $C = 2K$ satisfying constraints (3) and (4) for which the algorithm incurs risk at least $\Omega(C^\beta)$. In particular, the self-confident rates (Auer et al., 2002a) hold with $\gamma = 1/2$.

Proof. Let $\gamma_+ := \max\{\gamma, 0\}$ and set $f_K(x) := c/(1+x)^{\gamma_+}$. This function is nonincreasing, and the assumed rate lower bound implies $\eta_t \geq f_K(\tilde{L}_t^*)$. Moreover,

$$\frac{f_K(2K)K^{2-\beta}}{2\log K} = \frac{cK^{2-\beta}}{2(1+2K)^{\gamma_+}\log K} \longrightarrow \infty,$$

because $2-\beta-\gamma_+ > 0$ under the stated choice of β. Therefore $f_K(2K) \geq 2\log K/K^{2-\beta}$ for all sufficiently large K, and the claim follows from Theorem 2. □

Note that Corollary 1 *does not* rule out all adaptive FTRL-style schedules. In particular, its threshold condition need not hold for *harmonic* or *dyadic* rates of order $1/(1+\tilde{L}_t^*)$. To clarify this distinction, Appendix C provides an empirical comparison of these schedules on the hard instance used in Theorem 2.

3.3 Achieving minimax optimal risk for unknown C

In Theorem 2, we showed that the FTRL-style self-confident EWA, and more generally schedules satisfying its rate-threshold condition, cannot achieve $\tilde{O}(C\log K)$ risk when C is *unknown*. In contrast, Theorem 1 yields an $O(C\log K)$ risk for $C \geq 1$ if we set $\eta = 1/C$, matching the general lower bound in Lemma 2; however, this choice requires the learner to know C in advance. To address these issues, we introduce a "guess-and-double" strategy in Algorithm 2 to estimate C. The main idea is to monitor the minimal cumulative noisy loss observed so far, which always underestimates C, and use this estimate to tune the learning rate. Once the observed noisy losses exceed our current estimate, we double the estimate and restart the EWA; this restart mechanism is the key distinction that bypasses the lower bound in Theorem 2.

In our next main result, we show that Algorithm 2 achieves risk $O((C+1)\log(e+C)\log K)$, which nearly matches the general lower bound in Lemma 2 for $C \geq 1$.

Theorem 3. *Let $\ell_1, \ldots, \ell_T \in \mathbb{R}^K$ be arbitrary true loss vectors and $\tilde{\ell}_1, \ldots, \tilde{\ell}_T \in \mathbb{R}^K$ be corresponding noisy loss vectors satisfying conditions (3) and (4). Then Algorithm 2, which does not require prior knowledge of C, guarantees*

$$\sum_{t=1}^{T} \langle \hat{p}_t, \ell_t \rangle \;\leq\; O((C+1)\log(e+C)\ln K)\,.$$

For $C \geq 1$, this matches the lower bound in Lemma 2 up to the logarithmic factor $\log(e+C)$.

Proof. Since losses are nonnegative, condition (3) implies that the reference expert has true loss zero at every time. Together with condition (4), this implies that its cumulative noisy loss is at most C at every time. Thus L_m doubles at each epoch while remaining at most $C+1$, and there are at most $O(\log(e+C))$ epochs. Let $I_m \subseteq [T]$ denote the set of time steps belonging to epoch m. During each epoch, the algorithm runs as a *refresh* EWA algorithm with learning rate $1/L_m$, and the minimum cumulative noisy loss in the epoch is at most $2L_m+1$. The restriction to an epoch again satisfies conditions (3) and (4) with budget C. Invoking Theorem 1, we obtain

$$\sum_{t\in I_m} \langle \hat{p}_t, \ell_t \rangle \leq O\left(\log K \cdot L_m + \frac{C(2L_m+1)}{L_m} + C\log K\right) \leq O\left(\log K \cdot L_m + C + C\log K\right).$$

Summing over all epochs and using the fact that $\sum_m L_m \leq O(C+1)$, we obtain

$$\sum_{t=1}^{T} \langle \hat{p}_t, \ell_t \rangle \;\leq\; O\left((C+1)\log K + C\log K \cdot \log(e+C)\right) = O\left((C+1)\log(e+C)\log K\right),$$

as claimed. □

Remark 2 (Dyadic FTRL also realizes the doubling idea)**.** *The restart-based implementation in Algorithm 2 is not the only way to realize the dyadic decomposition used in the proof of Theorem 3. The same order of risk can also be obtained by the FTRL-style update (8) with the dyadic learning rate*

$$\eta_t = 2^{-\lceil \log_2(1+\tilde{L}_t^*)\rceil}.$$

Indeed, on each dyadic band of $\tilde{L}_t^*$, this rate is constant and of order the inverse band scale. Repeating the proof of Theorem 3 band by band gives the same $O((C+1)\log(e+C)\log K)$ guarantee; the losses accumulated before a band contribute only a constant multiple of that band's scale. Thus Theorem 2 should not be read as ruling out all FTRL-style adaptive rules. It rules out the lower-envelope class of schedules whose rates remain too large on the hard instance, such as the Auer-type self-confident square-root rate.

Algorithm 2 EWA with Doubling for Unknown C

Input: Number of experts K
Initialize: epoch index $m \leftarrow 1$, threshold $L_1 \leftarrow 1$, learning rate $\eta_1 \leftarrow 1/L_1$, weights $w_k^1 \leftarrow 1$, cumulative noisy losses $\tilde{L}_0(k) \leftarrow 0$ for all $k \in [K]$
for $t = 1, 2, \ldots, T$ **do**
  if $\min_{k\in[K]} \tilde{L}_{t-1}(k) \;\geq\; 2L_m$ **then**
    $m \leftarrow m + 1$
    $L_m \leftarrow \min_{k\in[K]} \tilde{L}_{t-1}(k), \quad \eta_m \leftarrow 1/L_m$
    Reset weights: $w_k^t \leftarrow 1$ for all $k \in [K]$
  end if
  Compute probability distribution over experts:

$$\hat{p}_t[k] \;\leftarrow\; \frac{w_k^t}{\sum_{j=1}^K w_j^t} \qquad \text{for } k \in [K]$$

  Sample $\hat{k}_t$ from the distribution $\hat{p}_t$
  Predict using expert $\hat{k}_t$
  Observe the noisy loss vector $\tilde{\ell}_t \in [0,1]^K$
  Update weights for all experts:

$$w_k^{t+1} \;\leftarrow\; w_k^t \exp\bigl(-\eta_m \tilde{\ell}_t[k]\bigr) \qquad \text{for } k \in [K]$$

  Update cumulative noisy losses:

$$\tilde{L}_t(k) \;\leftarrow\; \tilde{L}_{t-1}(k) + \tilde{\ell}_t[k] \qquad \text{for } k \in [K]$$

end for

The loss model in (Papagiannis et al., 2025, Theorem 2) is exactly the model defined by conditions (3) and (4). Consequently, if its constant-rate EWA predictor is replaced by Algorithm 2, Theorem 3 provides the alternative unknown-C upper bound

$$O((C+1)\log(e+C)\log K)\,.$$

The factor $\log(e+C)$ is the price paid by the current epoch analysis for not knowing the corruption budget in advance. Compared with the $O(C^2 + C\log K)$ bound for constant-rate EWA in (Papagiannis et al., 2025, Theorem 2), this is asymptotically smaller whenever C dominates $\log(e+C)\log K$.

On the one-sided relaxation in Papagiannis et al. (2025). The Massart-noise application in (Papagiannis et al., 2025, Theorem 3) invokes the one-sided relaxation in (Papagiannis et al., 2025, Corollary 1), which controls only the signed cumulative discrepancy. Even if that signed inequality is strengthened to hold at every prefix, Proposition 1 gives only an $O((C+1)\log^2 K)$-type guarantee for EWA, while Remark 1 gives a matching $\Omega(\log^2 K)$ example already when $C = 1$. Hence the additional $\log K$ factor is not an artifact of the proof: the claimed bound without this additional logarithm under the one-sided condition in (Papagiannis et al., 2025, Corollary 1) does not hold as stated.

On smooth schedules and the $\log C$ factor. Remark 2 shows that explicit restarts are not essential for the guarantee in Theorem 3: a dyadic FTRL-style rate realizes the same doubling idea. However, not every smoothed adaptive schedule works. Theorem 2 shows that the FTRL-style self-confident schedule can lead to strictly worse performance in our setting, with risk growing superlinearly in C. The result also applies to FTRL schedules with the stated rate lower envelope. In Theorem 3, the $\log(e + C)$ factor originates from the $C \log K$ term in the epoch-wise guarantee of Theorem 1. This term arises from the discrepancy bound used there, and it does not appear amenable to a simple replacement by something like $O(L_m \log K)$. We therefore suspect that the $\log(e + C)$ dependence may be an artifact of the proof technique, and we leave it as an open problem to determine whether it can be removed.

3.4 Extension to non-realizable true losses

In this section, we address the scenario where the true losses are not realizable, i.e., when condition (3) is violated. To handle this case, we introduce the notion of *regret*. For any learning strategy $\hat{p}_t = \Phi_t(\tilde{\ell}_1, \ldots, \tilde{\ell}_{t-1})$, we define its worst-case *regret* as

$$\mathrm{Regret}_{T,K,C} \;=\; \min_{\{\Phi_t\}_{t=1}^T} \sup_{\ell^T, \tilde{\ell}^T} \left(\mathbb{E}\Big[\sum_{t=1}^T \ell_t[\hat{k}_t] \Big] - \inf_{k \in [K]} \sum_{t=1}^T \ell_t[k] \right).$$

Note that our regret formulation is evaluated on the *true* losses, not on the observable noisy losses.

We now state our main result for this section as follows with proof delegated to Appendix B.

Theorem 4. *For arbitrary loss vectors $\ell_1, \ldots, \ell_T$ and noisy loss vectors $\tilde{\ell}_1, \ldots, \tilde{\ell}_T$ satisfying (4), the EWA algorithm (Algorithm 1) with any learning rate $\eta > 0$ achieves regret at most*

$$\mathrm{Regret}_{T,K,C} \;\le\; O\!\left(\frac{\ln K}{\eta} + \eta(L^*_{\min} + C) + \eta C(C + L^*_{\min}) + C \ln K \right),$$

where $L^*_{\min} = \min_{k \in [K]} \sum_{t=1}^T \ell_t[k]$. In particular, if C and $L^*_{\min}$ are known, choose

$$\eta \;=\; \sqrt{\frac{\ln K}{L^*_{\min} + C + C(C + L^*_{\min})}}.$$

Then the regret of the EWA algorithm is at most

$$O\!\left(\sqrt{\ln K \cdot (L^*_{\min} + C + C(C + L^*_{\min}))} + C \ln K \right).$$

For $C \ge 1$, this simplifies to $O(\sqrt{\ln K \cdot C(C + L^*_{\min})} + C \ln K)$.

Note that the regret bound in Theorem 4 is non-trivial, as it grows sublinearly with $L^*_{\min}$. A naive extension of the realizable risk bound $O(C \log K)$ would instead yield an $O((C + L^*_{\min}) \log K)$ regret. Our sublinear guarantee benefits from the tighter analysis of EWA in Theorem 1. However, the bound still requires knowledge of both C and $L^*_{\min}$ in order to tune the learning rate. To handle unknown C and $L^*_{\min}$, we leverage a similar doubling approach as in Algorithm 2 and prove below that it achieves slightly worse regret compared to Theorem 4.

Theorem 5. *For arbitrary loss vectors $\ell_1, \ldots, \ell_T$ and noisy loss vectors $\tilde{\ell}_1, \ldots, \tilde{\ell}_T$ satisfying only (4), Algorithm 2 with learning rates $\eta_m = 1/\sqrt{L_m}$ (instead of $\eta_m = 1/L_m$), without prior knowledge of C or $L^*_{\min}$, achieves*

$$\mathrm{Regret}_{T,K,C} \;\le\; O\Big((\ln K + C)\sqrt{L^*_{\min} + C + 1} + C \ln K \log(e + L^*_{\min} + C) \Big).$$

Proof. Let I_m be the set of rounds in epoch m, and let L_m denote the threshold used by Algorithm 2 in that epoch. Let $k^* \in \arg\min_k \sum_{t=1}^T \ell_t[k]$. Since $\sum_t \tilde{\ell}_t[k^*] \le L^*_{\min} + C$, the global minimum cumulative noisy loss

is always at most $L^*_{\min} + C$. Hence every threshold satisfies $L_m \le L^*_{\min} + C + 1$, the number of epochs is $O(\log(e + L^*_{\min} + C))$, and the dyadic growth of the thresholds gives

$$\sum_m \sqrt{L_m} \le O\Big(\sqrt{L^*_{\min} + C + 1}\Big).$$

We next bound the regret on one epoch. Let

$$\tilde{N}^*_m := \min_k \sum_{t \in I_m} \tilde{\ell}_t[k], \qquad L^*_m := \min_k \sum_{t \in I_m} \ell_t[k].$$

During epoch m, the restart rule ensures $\tilde{N}^*_m \le 2L_m + 1$. The standard noisy-loss regret bound for EWA with $\eta_m = 1/\sqrt{L_m}$ gives

$$\sum_{t \in I_m} \langle \hat{p}_t, \tilde{\ell}_t \rangle - \tilde{N}^*_m \le O(\sqrt{L_m} \ln K).$$

Moreover, if b_m is a true-loss minimizer on I_m, then $\tilde{N}^*_m \le \sum_{t \in I_m} \tilde{\ell}_t[b_m] \le L^*_m + C$, so $\tilde{N}^*_m - L^*_m \le C$.

It remains to control the discrepancy on the epoch. Repeating the log-potential argument from Theorem 1 within the restarted epoch, and using an expert attaining $\tilde{N}^*_m \le 2L_m + 1$ as reference, gives

$$\sum_{k=1}^{K} \max_{t \in I_m} \hat{p}_t[k] \le O(\ln K + \eta_m L_m) = O(\ln K + \sqrt{L_m}).$$

By (4),

$$\sum_{t \in I_m} \langle \hat{p}_t, \ell_t - \tilde{\ell}_t \rangle \le O(C \ln K + C\sqrt{L_m}).$$

Combining the three displays,

$$\sum_{t \in I_m} \langle \hat{p}_t, \ell_t \rangle - L^*_m \le O\Big((\ln K + C)\sqrt{L_m} + C \ln K\Big).$$

Finally, $\sum_m L^*_m \le L^*_{\min}$, because the best expert on each epoch can only improve over the single best expert used on all epochs. Therefore the regret to the global best expert is bounded by the sum of the epoch-wise upper bounds:

$$\mathrm{Regret}_{T,K,C} \le O\Big((\ln K + C)\sqrt{L^*_{\min} + C + 1} + C \ln K \log(e + L^*_{\min} + C)\Big).$$

This completes the proof. □

4 Conclusion

We developed a minimax theory for online learning from expert advice under adversarial feedback corruption, where the learner observes only noisy losses but is evaluated against the hidden *true* losses subject to a per-expert budget C. Our central finding is that the optimal risk is of order $\tilde{\Theta}(C \log K)$ and is achievable *without prior knowledge of* C, resolving an open questions of Papagiannis et al. (2025). This follows from a novel potential-based analysis of EWA (Theorem 1) that holds for an arbitrary learning rate, together with an epoch-based "guess-and-double" scheme (Theorem 3) that adapts to *unknown* C. Beyond the achievability results, we also expose a sharp distinction within the family of adaptive EWA schedules: restart-based and dyadic-FTRL rules inherit the near-optimal guarantee, whereas a broad lower-envelope class of smooth FTRL-style schedules—including the classical self-confident rule of Auer et al. (2002a)—provably incurs risk $\Omega(C^\beta)$ for every $1 < \beta < \frac{3}{2}$ (Theorem 2). This separation reflects the fundamental gap between optimizing for observable losses and for the true losses against which performance is measured. We further moved beyond the realizable case, establishing sublinear regret bounds on the true losses (Theorems 4 and 5) that require knowledge of neither C nor $L^*_{\min}$.

Several questions remain open. Most concretely, we do not know whether the $\log(e + C)$ overhead in Theorem 3 is intrinsic to the unknown-C problem or merely an artifact of the epoch-wise discrepancy analysis. Relatedly, our FTRL lower bound targets a rate lower-envelope class and leaves untouched harmonic and dyadic rates of order $1/(1+\tilde{L}_t^*)$, so characterizing exactly which smooth adaptive schedules achieve the optimal rate is a natural next step. More broadly, we expect the potential-based discrepancy control developed here to be useful for other corrupted-feedback settings in adversarial online learning, where performance is judged on a quantity the learner never directly observes.

References

Idan Amir, Idan Attias, Tomer Koren, Yishay Mansour, and Roi Livni. Prediction with corrupted expert advice. *Advances in Neural Information Processing Systems*, 33:14315–14325, 2020.

P. Auer, N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 2002a.

Peter Auer, Nicolo Cesa-Bianchi, and Claudio Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002b.

Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. Agnostic online learning. In *COLT*, volume 3, pp. 1, 2009.

Alankrita Bhatt and Victoria Kostina. Prediction with noisy expert advice. In *2024 IEEE International Symposium on Information Theory (ISIT)*, pp. 3546–3551. IEEE, 2024.

Alankrita Bhatt and Victoria Kostina. Prediction with expert advice under additive noise. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL `https://openreview.net/forum?id=2rupjsmVpp`.

Nader H Bshouty, Nadav Eiron, and Eyal Kushilevitz. Pac learning with nasty noise. *Theoretical Computer Science*, 288(2):255–275, 2002.

N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning and Games*. Cambridge University Press, 2006.

Pierre Gaillard, Gilles Stoltz, and Tim Van Erven. A second-order bound with excess losses. In *Conference on Learning Theory*, pp. 176–196. PMLR, 2014.

Sham Kakade and Adam T Kalai. From batch to transductive online learning. *Advances in Neural Information Processing Systems*, 18, 2005.

Jaouad Mourtada and Stéphane Gaïffas. On the optimality of the hedge algorithm in the stochastic regime. *Journal of Machine Learning Research*, 20(83):1–28, 2019.

Nikolaos Papagiannis, Wojciech Szpankowski, and Changlong Wu. Online learning with nasty experts. In *2025 IEEE International Symposium on Information Theory (ISIT)*, 2025.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Random averages, combinatorial parameters, and learnability. *NIPS*, 2010.

S. Shalev-Shwartz. *Online Learning: Theory, Algorithms, and Applications*. PhD thesis, The Hebrew University of Jerusalem, 2007.

Tsachy Weissman, Neri Merhav, and Anelia Somekh-Baruch. Twofold universal prediction schemes for achieving the finite-state predictability of a noisy individual binary sequence. *IEEE Transactions on Information Theory*, 47(5):1849–1866, 2001.

Changlong Wu, Ananth Grama, and Wojciech Szpankowski. Robust online classification: From estimation to denoising. *arXiv preprint arXiv:2309.01698*, 2023.

Changlong Wu, Ananth Grama, and Wojciech Szpankowski. Information-theoretic limits of online classification with noisy labels. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024. URL `https://openreview.net/forum?id=Ke3MSP8Nr6`.

Changlong Wu, Ananth Grama, and Wojciech Szpankowski. Online universal learning from an information-theoretic perspective. *Foundations and Trends in Communication and Information Theory*, 23(3–4):225–443, 2026. ISSN 1567-2190.

A Proof of Lemma 2

Our goal is to construct hard instances $\ell^T, \tilde{\ell}^T$ using a *probabilistic* argument. We partition the time horizon into $\log K$ epochs, each of size C (with the other $T - C\log K$ positions padded with 0 losses). Here, we assume $\log K$ is an integer; otherwise, we use the largest power of 2, K', such that $\log K' \geq \log K - 1$. For each epoch $j \in [\log K]$, we maintain a *random* expert class $S^j \subset [K]$ with $|S^j| = \frac{K}{2^j}$. To construct this, we first select $S^0 := [K]$ and sample S^{j+1} as a random subset of S^j such that $|S^{j+1}| = |S^j|/2$ for all $j \geq 0$.

Having constructed the sets S^j, we define the true and noisy losses as follows. For any time step t in epoch j:

- $\ell_t[k] = 0$ if $k \in S^j$ and $\ell_t[k] = 1$ otherwise;
- $\tilde{\ell}_t[k] = 0$ if $k \in S^{j-1}$ and $\tilde{\ell}_t[k] = 1$ otherwise.

It is straightforward to verify that both constraints (3) and (4) are satisfied by this construction. This is because any expert accumulates discrepancies between the true and noisy losses only within a *single* epoch, and each epoch has size C.

Now, consider *any* prediction rule. At each epoch $j \leq \log K$, the predictor must choose $\hat{k}_t \in S^{j-1}$, otherwise the true loss is 1. Moreover, since the noisy losses are 0 for all $k \in S^{j-1}$, the predictor gains no information about S^j during epoch j. Therefore, the expected error at *every* step, taken over the randomness of both S^j and the predictor's internal randomness, is at least $\frac{1}{2}$ (since $|S^j| = |S^{j-1}|/2$).

This implies that the expected cumulative risk is at least $\frac{1}{2}C \cdot \log K$, since the predictor incurs an expected error of $1/2$ at *every* step, and by the linearity of expectation. The lemma then follows from the fact that there must *exist* realizations $\ell^T, \tilde{\ell}^T$ that achieve this expected cumulative risk bound.

B Proof of Theorem 4

Proof. Let $k^* \in \arg\min_{k\in[K]} \sum_{t=1}^T \ell_t[k]$, and write $L^*_{\min} := \sum_{t=1}^T \ell_t[k^*]$. By (4),

$$\sum_{t=1}^T \tilde{\ell}_t[k^*] \leq L^*_{\min} + C.$$

We use the standard first-order regret bound for fixed-rate EWA on the noisy losses:

$$\sum_{t=1}^T \langle \hat{p}_t, \tilde{\ell}_t \rangle - \sum_{t=1}^T \tilde{\ell}_t[k^*] \leq O\left(\frac{\ln K}{\eta} + \eta \sum_{t=1}^T \tilde{\ell}_t[k^*]\right) \leq O\left(\frac{\ln K}{\eta} + \eta(L^*_{\min} + C)\right).$$

It remains to control the discrepancy term. As in the proof of Theorem 1,

$$\sum_{t=1}^T \langle \hat{p}_t, \ell_t - \tilde{\ell}_t \rangle \leq C \sum_{k=1}^K \max_{t\in[T]} \hat{p}_t[k].$$

Relabel k^* as the reference expert in the log-potential argument. Since its cumulative noisy loss is at most $L^*_{\min} + C$, the same telescoping calculation gives

$$\sum_{k \neq k^*} \max_{t \in [T]} \hat{p}_t[k] \leq \ln K + \eta(L^*_{\min} + C).$$

Together with $\max_t \hat{p}_t[k^*] \leq 1$, this bounds the discrepancy by $O(C \ln K + \eta C(L^*_{\min} + C) + C)$. Finally,

$$\sum_{t=1}^{T} \tilde{\ell}_t[k^*] - L^*_{\min} \leq C.$$

Combining the noisy regret, discrepancy, and comparator terms proves

$$\sum_{t=1}^{T} \langle \hat{p}_t, \ell_t \rangle - L^*_{\min} \leq O\left(\frac{\ln K}{\eta} + \eta(L^*_{\min} + C) + \eta C(C + L^*_{\min}) + C \ln K \right),$$

where the lower-order C term is absorbed by $C \ln K$. □

C Empirical comparison of FTRL rate schedules

We evaluate FTRL-style EWA updates of the form

$$\hat{p}_{t+1}(k) \propto \exp\bigl(-\eta_t \tilde{L}_t(k)\bigr)$$

on the adversarial construction from Theorem 2, with $\beta = 1.25$. The experiment varies the corruption level C through the corresponding values of K and plots the resulting true risk on a log–log scale. We compare three schedules:

- **Harmonic FTRL:** $\eta_t = 1/(1 + \tilde{L}^*_t)$.
- **Auer-type FTRL:** $\eta_t = \sqrt{\log K / \max\{1, \tilde{L}^*_t\}}$.
- **Dyadic (Doubling) FTRL:** $\eta_t = 2^{-\lceil \log_2(1+\tilde{L}^*_t) \rceil}$.

Finding. Figure 1 shows that the dyadic FTRL schedule (which achieves the same theoretical guarantee as Algorithm 2 in Theorem 3) exhibits the best empirical scaling (slope ≈ 1.05), while the Auer-type square-root schedule is clearly suboptimal (slope ≈ 1.32). Interestingly, although the harmonic schedule can be viewed as a smoother analogue of the dyadic rule, in the sense that both rates scale on the order of $1/(1 + \widetilde{L}^\star_t)$, it still exhibits a larger empirical slope (about 1.13).

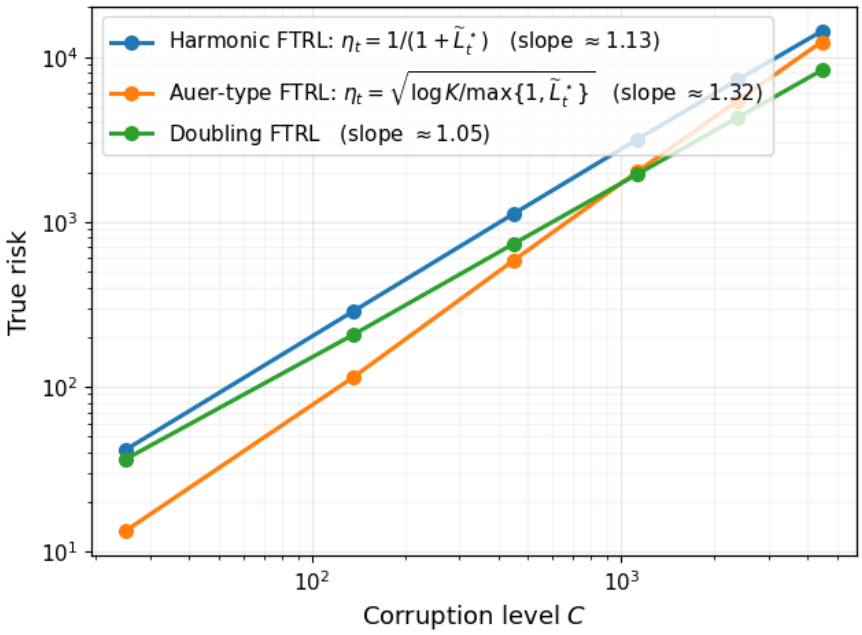

Figure 1: **Comparison of FTRL-type rates on the hard instance of Theorem 2.** The plot reports true risk versus corruption level C on a log–log scale.