Prediction with noisy expert advice

Alankrita Bhatt and Victoria Kostina

California Institute of Technology, Pasadena, CA
{abhatt,vkostina}@caltech.edu

Abstract

Regret minimization in the problem of prediction with expert advice in the presence of noisy feedback is a fundamental challenge in online learning and sequential decision making. A general framework is proposed for designing and analyzing no-regret algorithms in this setting. This analysis, when specialized to several canonical channel models, is shown to lead to tight bounds on the regret thus characterizing how the noise level affects the regret and demonstrating that in some cases it is possible to achieve the same regret as with noiseless feedback.

1 Introduction

Prediction with expert advice is a fundamental problem in online learning and sequential decision making [Cesa-Bianchi et al., 1997a, Cesa-Bianchi and Lugosi, 2006, Hazan et al., 2016, Orabona, 2019] where the goal is to aggregate decisions from $m \geq 2$ "experts" and achieve performance approaching that of the best individual expert in hindsight in the long run, without any prior knowledge of the identity of this hindsight-optimal expert.

In this paper, we study a variation on the prediction with expert advice setting where the feedback available to the decision maker at each step is corrupted by noise. To formally define our problem setup, we recall first the standard prediction with experts setting (with m experts): This takes the form of a sequential game where at each time $t \in [n]$

- The decision-maker picks $p_t \in \Delta^{m-1}$, i.e. a probability distribution on the $m-$simplex based on the feedback so far, representing the weight placed on each expert.
- The loss vector $\ell_t \in \mathcal{L} := [0,1]^m$ is revealed, where ℓ_{tj} represents the loss incurred by the $j-$th expert in round t, and the decision maker incurs loss $\langle p_t(\ell^{t-1}), \ell_t \rangle$.

For a strategy p represented by a sequence of functions $\{p_t(\cdot)\}_{t=1}^n$ where $p_t \colon \mathcal{L}^{t-1} \to \Delta^{m-1}$, the *regret* incurred on a loss sequence $\ell^n := \ell_1, \ldots, \ell_n$ is defined as

$$\mathrm{Reg}(p, \ell^n) := \sum_{t=1}^n \langle p_t(\ell^{t-1}), \ell_t \rangle - \min_{j \in [m]} \sum_{t=1}^n \ell_{tj} \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. The interpretation of (1) is that at time t the decision maker picks an expert at random, as per the distribution p_t (so that the decision-maker's expected loss in the $t-$th round is $\sum_{j=1}^m p_{tj}\ell_{tj} = \langle p_t, \ell_t \rangle$) and at the end of the game the cumulative performance of the decision-maker is measured against that of the best fixed expert in hindsight. In particular, the identity of the best expert(s) $\mathrm{argmin}_{j \in [m]} \sum_{t=1}^n \ell_{tj}$ cannot be calculated before the end of the time horizon n, making approaching the performance of this best expert the key challenge.

The objective is to determine fundamental limits on the *individual-sequence* minmax regret

$$\inf_p \sup_{\ell^n} \mathrm{Reg}(p, \ell^n)$$

or in other words, to characterize the regret obtained by the optimal strategy that works even when the loss functions ℓ_1^n are chosen in an adversarial manner. A classical result [Cesa-Bianchi et al., 1997a] characterizes the order-optimal minmax regret as

$$\inf_p \sup_{\ell^n} \mathrm{Reg}(p, \ell^n) = \Theta(\sqrt{n \log m}) \tag{2}$$

where $\Theta(\cdot)$ hides fixed absolute constants independent of parameters such as n and m.

While the aforementioned prediction with expert advice setting (with the fundamental performance limit characterized in (2)) is a satisfying model in many cases, it has some limitations. In particular, in many real-world scenarios the feedback received by the decision-maker may not be exactly equal to ℓ_t as it could be corrupted by noise, errors, or be communication-constrained. As an example case, consider autonomous driving: here, decision-making is challenged by rate-constrained feedback due to the processing time required for sensor data interpretation. Additionally, the sensor data is often subject to additive noise, stemming from environmental factors and sensor inaccuracies, which introduces uncertainty into the vehicle's perception of its surroundings. Such noisy feedback in decision-making poses a significant challenge, as any decision-making strategy has to cope with both the uncertainty of the environment (i.e. the adversarially chosen losses ℓ_t) and the unreliability of the feedback. Therefore, it is practically relevant to study how to design robust and adaptive algorithms that can learn effectively from noisy feedback and still achieve low regret.

To capture this challenge, we consider the *prediction with noisy expert advice* setting where only a partial observation c_t of the loss ℓ_t is available to the decision-maker: for example, c_t could be a noise-corrupted version of ℓ_t, or it could be the finite-precision output available when ℓ_t is conveyed over a rate-constrained channel. Formally, in this noisy setting we have

- The *channel*, which is a sequence of random or deterministic transformations acting on the losses ℓ^t at time t as $c_t \colon \mathcal{L}^t \to \mathcal{C}$, where $\mathcal{C}$ is the channel output alphabet.
- The *decision*, which is a distribution $p_t(c^{t-1})$ constructed on the basis of the channel outputs c^{t-1} available so far.

The challenge in this setup is that with only knowledge of the (noisy) channel outputs c^n, we wish to perform as well, in expectation, as the best expert on the *clean* loss function ℓ^n. Formally, we are interested in the quantity

$$\mathrm{Reg}(p, P_{c|\ell}, \ell^n) := \sum_{t=1}^{n} \langle \mathsf{E}[p_t(c^{t-1})], \ell_t \rangle - \min_{j \in [m]} \sum_{t=1}^{n} \ell_{tj} \tag{3}$$

for the worst-case individual sequence ℓ^n. Here, the $\mathsf{E}[\cdot]$ is with respect to any noise induced by the channel $P_{c|\ell}$. We emphasize that this is a strictly more difficult problem than prediction with expert advice (compare (3) with (1)) since while the benchmark performance remains the same in both cases ($\min_{j\in[m]} \sum_{t=1}^{n} \ell_{tj}$), in the noisy expert advice setting this performance must be approached with only knowledge of c_t which is a degraded version of ℓ_t. Naturally, the regret would depend not only on the performance of the experts, but also on some measure of the quality of the channel.

While the outlined noisy expert advice setting is quite general, two practically-motivated classes of channels are of particular interest in this paper:

- **Memoryless noise.** Here, the c_t is the output of a *fixed known* random transformation $P_{c|\ell}$ with input ℓ_t. For example, the additive white Gaussian noise (AWGN) channel acts as $c_t = \ell_t + Z_t$ where $Z_t \sim \mathcal{N}(0, \sigma^2 I)$. In this case, we wish to devise a decision strategy p that at time step t maps the noisy outputs c^{t-1} to a decision and achieves low regret (3).

- **Quantization noise.** Here the channel output is $\mathcal{C} := \left[2^R\right]$, i.e. there are only R bits available to communicate the loss vector ℓ_t to the decision maker. The decision maker has a twofold challenge: firstly, to devise an *encoding* strategy that maps the loss functions observed so far ℓ^t into an $R-$bit message c_t; and secondly to devise a *decision* strategy that maps all the messages received so far c^{t-1} into a distribution over experts p_t.

1.1 Main results

Our main result is a characterization of the fundamental limits of minmax regret in the prediction with noisy expert advice setting, for several practically-relevant canonical channels in information theory and machine learning. To achieve this characterization, we present two very general results which are then specialized to the aforementioned channels of interest.

To establish an upper bound on the minmax regret, one must construct a strategy whose regret achieves said bound. To this end, we provide a general strategy based on the exponential weights (Hedge) algorithm [Cesa-Bianchi et al., 1997b] and analyze this strategy's regret for general channels, establishing that it depends on the mean-square error of estimating ℓ_t from c_t (see Theorem 1 for a precise statement). To achieve a lower bound on the regret, we provide a fundamental converse on the performance of any strategy. This general converse quantifies the intuition that the regret grows as the quality of the channel degrades, and in particular depends on a quantity called the contraction coefficient of the channel (see Theorem 2 for a precise statement).

Theorems 1 and 2 when specialized to particular channels yield the following tight characterization of regret in several noise models. Here, $\mathrm{Reg}(P_{c|\ell})$ represents the minmax regret of prediction with noisy expert advice for channel $P_{c|\ell}$.

- **Binary Symmetric Channel.** If $\ell_t \in \{0,1\}^m$, the binary symmetric channel with bias θ (denoted as $\mathrm{BSC}(\theta)$) has output $c_{tj} = \ell_{tj} \oplus Z_{tj}$ where $\oplus$ denotes modulo-2 addition and the Z_{tj}'s are $\mathrm{Bern}(\theta)$ i.i.d. That is, all bits are (independently) flipped with probability θ. If the channel $P_{c|\ell}$ is a binary symmetric channel $\mathrm{BSC}(\theta)$, then our results yield

$$\mathrm{Reg}(\mathrm{BSC}(\theta)) = \Theta\left(\sqrt{\frac{n \log m}{(1-2\theta)^2}}\right). \tag{4}$$

- **Binary erasure channel.** For $\ell_t \in \{0,1\}^m$, the binary erasure channel with erasure probability e has output $c_{tj} = \ell_{tj}$ with probability $1-\mathsf{e}$, and is equal to erasure symbol "?" with probability e. That is, either the channel noiselessly recovers the input, or it is erased. If the channel $P_{c|\ell}$ is a binary erasure channel $\mathrm{BEC}(\mathsf{e})$, then our results yield

$$\mathrm{Reg}(\mathrm{BEC}(\mathsf{e})) = \Theta\left(\sqrt{\frac{\log m}{1-\mathsf{e}}}\right). \tag{5}$$

- **Additive noise channels.** The general class of additive noise channels, as the name suggests, have output $c_t = \ell_t + Z_t$ where Z_t is a random variable representing the noise. We can further characterize the regret for certain (classes of) distributions.

– **Gaussian noise.** If the additive noise Z_{tj} is Gaussian with variance σ^2, which we denote by AWGN(σ), then

$$\mathrm{Reg}(\mathrm{AWGN}(\sigma)) = \Theta\left(\sqrt{(1+\sigma^2)n\log m}\right). \tag{6}$$

– **Uniform noise.** If the additive noise Z_{tj} is uniformly distributed in $[-\sigma, \sigma]$ which we denote by AddUnif(σ), then

$$\mathrm{Reg}(\mathrm{AddUnif}(\sigma)) = \Theta\left(\sqrt{(1+\sigma)n\log m}\right). \tag{7}$$

– **Symmetric log-concave noise.** If the additive noise Z_{tj} comes from an arbitrary symmetric log-concave distribution with variance σ^2, which we denote by Add(f_σ), then

$$\Omega\left(\sqrt{(1+\sigma)n\log m}\right) \leq \mathrm{Reg}(\mathrm{Add}(f_\sigma)) \leq O\left(\sqrt{(1+\sigma^2)n\log m}\right). \tag{8}$$

We note that while (8) is not a tight characterization, the lower bound is tight for uniform noise (see (7)) and the upper bound is tight for Gaussian noise (see (6)).

- 1 bit/dimension quantization: If ℓ_t is subject to a communication constraint and can only be expressed with rate $R = m$, i.e. 1 bit for each expert, then our results yield

$$\max_{\ell^n} \mathrm{Reg}(p, m\text{-bit}, \ell^n) = \Theta\left(\sqrt{n\log m}\right). \tag{9}$$

In the first three channel models, there is an extra factor, not present in the noiseless setting (2) that is due to the noise in the channel. In particular, we see that as $\theta \to 1/2$, $\mathtt{e} \to 1$ and $\sigma \to \infty$, the regret cannot be sublinear in n. This makes intuitive sense since in these regimes, there is no informative feedback to base decisions on. In the fourth setting, we have established that even with a 1−bit/expert communication bottleneck, we can achieve the same regret as one would be able to achieve with full-precision feedback.

1.2 Related work

To the best of our knowledge, Weissman, Merhav and Somekh-Baruch [Weissman et al., 2001] were the first to consider noisy prediction for individual sequences ℓ^n [Feder et al., 1992, Merhav and Feder, 1998] in the presence of BSC noise. They devised achievability schemes as well as a notion of conditional finite-state predictability. They also showed achievability schemes that do not depend on the BSC bias θ, thereby establishing universality both with respect to the best expert in hindsight, and the unknown channel parameter. Weissman and Merhav [Weissman and Merhav, 2001] established general achievability schemes for noisy individual-sequence universal prediction, and subsequent work [Weissman and Merhav, 2004] established similar results for noisy prediction of stationary ergodic sources. The result in [Weissman and Merhav, 2001] was extended in [Resler and Mansour, 2019] to the adversarial bandit setting (i.e. rather than the entire loss vector ℓ_t being provided as feedback, only the loss incurred by the selected expert is presented to the decision-maker).

A closely related area where sequential decision-making with noisy feedback has been considered is control. The question examined here is how control systems can maintain stability and performance despite the presence of noise in the feedback loop. While measurement-feedback control is a classical topic [Kailath et al., 2000], the line of work [Tatikonda and Mitter, 2004a,b, Tatikonda et al., 2004, Kostina and Hassibi, 2019] examines fundamental limits of control performance when the feedback is subject to communication constraints.

[Raginsky et al., 2012] considered sequential anomaly detection and sequential probability assignment (i.e. online prediction using the logarithmic loss [Rissanen, 1984, Xie and Barron, 1997, Bhatt and Kim, 2021, Cesa-Bianchi and Lugosi, 2006]) in the presence of noise and established minmax regret guarantees. Also in the setting of sequential probability assignment, [Shkel et al., 2018] considered compressed side information available noncausally—our work considers compressed feedback available causally, in the prediction with experts setting. Decision-making with noisy feedback in the sequential classification setting has been considered in [Ben-David et al., 2009, Wu et al., 2023]. The effect of noisy observations on the equlibrium value of games was characterized in [Hsieh et al., 2022, Sun et al., 2023]. The setting where rather than the feedback ℓ_t the action p_t is communicated over a noisy channel is considered in [Donmez et al., 2015, Pase et al., 2022, Hanna et al., 2023], and minmax bounds on the regret incurred are established. The line of work [Acharya et al., 2020a,b, 2021] considers sequential statistical inference under constraints, designing optimal policies and well as establishing fundamental converse bounds.

We remark that our setting is distinct from that of i.i.d. ℓ_t and adversarially injected corruptions [Amir et al., 2020], a model which aims to bridge the distance between the case where the losses ℓ_t at chosen i.i.d. and the individual-sequence case (adversarial ℓ_t). Moreover, our choice of benchmark being $\min_{j\in[m]} \sum_{t=1}^{n} \ell_{tj}$ (see the regret definition (3)) makes our setting distinct from smoothed analysis [Haghtalab et al., 2020, Bhatt et al., 2023], where the benchmark is the best expert in hindsight on the noisy loss function—making smoothed analysis a beyond-worst case setting.

Recently, in part due to federated learning becoming a major paradigm of theoretical and practical interest [Kairouz et al., 2021], there has been a renewed interest in distributed learning with communication (rate) constraints. This setting has also been considered for sequential decision-making. In particular, recent work on stochastic bandits [Hanna et al., 2022, Mitra et al., 2023, Mayekar et al., 2023] has served as one of the motivations for our investigation. Schemes to achieve no-regret strategies and characterization of the the number of bits needed for the decision-maker to achieve same regret as in the full-precision case have been proposed for multi-arm bandits [Hanna et al., 2022] and linear bandits [Mitra et al., 2023]. Mayekar et al. [Mayekar et al., 2023], in addition to the rate constraint consider a power-constrained AWGN channel over which the feedback must be sent. They show an achievability and a converse result which together establish that the regret incurs an extra factor of $\sqrt{\frac{1}{\mathsf{SNR}}}$, where SNR denotes the signal-to-noise ratio.

In this paper, we consider the *individual-sequence* (adversarial), *full-information* (experts) setting. We provide achievability and converse results for a general noise channel, and then specialize our results to particular noise models. Interestingly, our results on 1-bit/expert quantization and the AWGN channel match those in [Hanna et al., 2022, Mayekar et al., 2023], since in the latter the regret scales linearly with σ and the former the regret is (order-wise) unchanged.

1.3 Organization

In Section 2 our decision-making strategy is described, as well as a general regret bound on this scheme is presented. Before presenting the proof of the general regret bound, it is specialized to the several channels of interest presented in Section 1.1 leading to the upper bounds on the regret presented therein. Similarly, in Section 3 a fundamental converse (lower bound) on the regret achieved by any strategy is presented for the general problem, which is then specialized to the example channels of interest complementing the upper bounds obtained in Section 2. Finally, Section 4 concludes the paper with a discussion on the results as well as directions for future work.

2 Achievability

In this section, we propose a general achievability scheme p for no-regret learning in the noisy experts problem. In this setting since ℓ_t is unavailable directly due to noise in the feedback, we must construct a strategy using a proxy for ℓ_t that only uses the (noise-corrupted) c^t. A natural and intuitive idea is to construct an *unbiased* estimator $\widehat{\ell}_t$ for ℓ_t based on c^t, plug it into no-regret strategy for the noiseless experts problem, and play with the resulting strategy. Unbiasedness of an estimator is an important property in statistical estimation, with several interesting and attractive consequences [Lehmann and Casella, 2006, Chapter 2]. We argue that if $\widehat{\ell}_t$ is an unbiased estimator, the regret incurred by using $\widehat{\ell}_t$ as a proxy for ℓ_t is close to the regret incurred in the noiseless setting. Such an idea has also been successfully employed across statistics and sequential decision-making, with two prominent appearances of a similar idea in a sequential domain being in [Weissman and Merhav, 2001] and [Auer et al., 1995]. The following proposition justifies the use of an unbiased estimator.

Proposition 1 *Let $\widehat{\ell}_t$ (where $\widehat{\ell}_t$ is a possibly noisy function of c^t) be such that $\mathsf{E}[\widehat{\ell}_t|\widehat{\ell}^{t-1}] = \ell_t$, and p be any strategy for the noiseless experts problem. Then, the strategy $\widehat{p}$ that plays $\widehat{p}_t = p_t(\widehat{\ell}^{t-1})$ achieves*

$$\operatorname{Reg}(p, P_{\mathsf{c}|\ell}, \ell^n) \leq E\left[\sum_{t=1}^{n}\langle p(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle - \min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}\right].$$

Proof.

$$\begin{aligned}\langle \mathsf{E}[p_t(\widehat{\ell}^{t-1})], \ell_t\rangle &= \mathsf{E}[\langle p_t(\widehat{\ell}^{t-1}), \ell_t\rangle] \\ &\overset{(a)}{=} \mathsf{E}[\langle p_t(\widehat{\ell}^{t-1}), \mathsf{E}[\widehat{\ell}_t|\widehat{\ell}^{t-1}]\rangle] \\ &= \mathsf{E}[\mathsf{E}[\langle p_t(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle|\widehat{\ell}^{t-1}]] \\ &\overset{(c)}{=} \mathsf{E}\left[\langle p_t(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle\right] \end{aligned} \tag{10}$$

where (a) follows by the conditional unbiasedness of $\widehat{\ell}_t$ and (b) follows by the tower property of expectation. Moreover,

$$\min_{j\in[m]}\sum_{t=1}^{n}\ell_{tj} \overset{(a)}{=} \min_{j\in[m]}\mathsf{E}\left[\sum_{t=1}^{n}\widehat{\ell}_{tj}\right] \overset{(b)}{\geq} \mathsf{E}\left[\min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}\right] \tag{11}$$

where (a) follows by the unbiasedness of $\widehat{\ell}_t$ and linearity of expectation, and (b) follows since $\mathsf{E}[\min(\cdot)] \leq \min\mathsf{E}[\cdot]$. The Proposition follows by summing up (10) over t and from (11). ■

Remark 1 (Memoryless channels) *If the channel is memoryless and the c_t is a noisy function of ℓ_t only, then constructing $\widehat{\ell}_t$ from only c_t (rather than c^t) to ensure that $\mathsf{E}[\widehat{\ell}_t] = \ell_t$ clearly satisfies the conditional unbiased assumption in Proposition 1. In fact, such a memoryless estimator will suffice for all our purposes in this section.*

Proposition 1 establishes that upon construction of an unbiased estimator $\widehat{\ell}_t$, the decision-maker can pretend that the benchmark is $\min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}$, and employ a no-regret strategy for this benchmark. To construct a scheme, we need to utilize a no-regret strategy for the noiseless setting in conjunction with an unbiased estimator $\widehat{\ell}_t$. To this end, recall the landmark exponential weights/Hedge (EW) strategy [Cesa-Bianchi et al., 1997a, Cesa-Bianchi and Lugosi, 2006, Freund and Schapire, 1999] which is minmax optimal for the noiseless experts setting, i.e. it attains the minmax bound (2). Our strategy will be an appropriate extension of EW that accounts for noise in the feedback. The EW strategy assigns a weight to each expert that is inversely proportional to the exponent of the loss incurred by the expert thus far, i.e.

$$p_{tj}^{\mathrm{EW}}(\ell^{t-1}) \propto \exp\left(-\alpha \sum_{i=1}^{t-1} \ell_{ij}\right). \tag{12}$$

where $\alpha > 0$ is the learning rate. We will need the following analysis of the exponential weights strategy p^{EW} (see for example [Luo, 2022]), which bounds the regret incurred by p^{EW} in terms of the second moment of the loss functions [Cesa-Bianchi et al., 2007, Gaillard et al., 2014]. It follows the standard idea of constructing a potential function and carefully bounding the change in the potential function at each time step, and is relegated to Appendix A.

Lemma 1 *If p_{tj}^{EW} is chosen as in* (12)*, and if ℓ^n and α satisfy $-\alpha\ell_{tj} \leq 1$ for all t and j, we have*

$$\mathrm{Reg}(p^{\mathrm{EW}}, \ell^n) \leq \frac{\log m}{\alpha} + \alpha \sum_{t=1}^{n} \sum_{j=1}^{m} p_{tj}^{\mathrm{EW}} \ell_{tj}^2. \tag{13}$$

Motivated by Proposition 1, will use an unbiased estimator in conjunction with the exponential weights strategy. Our general achievability strategy $\widehat{p}^{\mathrm{EW}}$ is:

- Construct an unbiased estimator $\widehat{\ell}_t$ for ℓ_t from the channel output c_t.
- Play $p_t^{\mathrm{EW}}(\widehat{\ell}^{t-1})$.

We then have the following bound on the regret of $\widehat{p}^{\mathrm{EW}}$.

Theorem 1 *Let the channel $P_{c|\ell}$ be memoryless and let $\widehat{\ell}_t$, constructed using c_t, be an unbiased estimator. For any $\alpha > 0$ define the event*

$$\mathcal{E} := \{\exists t, j : -\alpha\widehat{\ell}_{tj} \geq 1\}. \tag{14}$$

Then,

$$\mathrm{Reg}(\widehat{p}^{\mathrm{EW}}, P_{c|\ell}, \ell^n) \leq \frac{\log m}{\alpha} + \alpha n \left(1 + \max_{j,t} \mathbb{E}[\ell_{tj} - \widehat{\ell}_{tj}]^2\right) + \sqrt{4m^2n^2\left(1 + \max_{t,j} \mathbb{E}[\widehat{\ell}_{tj} - \ell_{tj}]^2\right)\mathbb{P}(\mathcal{E})}.$$

Theorem 1 follows by a combination of Proposition 1 and Lemma 1. We note, in particular, that the regret additionally depends on the mean squared error (MSE) obtained by the estimator $\widehat{\ell}_t$, drawing an interesting connection between estimation and noisy regret minimization. We also note that while strictly speaking Lemma 1 excludes non-bounded noise (since it is possible, for non-bounded noises, that $-\alpha\ell_{tj} > 1$ is possible) under very mild conditions on the tail distribution of the noise it can be adapted for non-bounded noise distributions as well, as illustrated in Theorem 1.

2.1 Application of Theorem 1 to canonical channel models

We now instantiate Theorem 1 to the four canonical noise sources in information theory: the BSC, the BEC, additive noise channels and quantization noise. In each case, we construct an unbiased estimator, evaluate its mean square error and establish the upper bound on the regret incurred upon using that particular estimator via Theorem 1.

2.1.1 Binary Symmetric Channel

Recall that for the BSC(θ), for $\ell_t \in \{0,1\}^m$, then $c_{tj} = \ell_{tj} \oplus Z_{tj}$ where $\oplus$ denotes modulo-2 addition and the Z_{tj}'s are Bern(θ) i.i.d. In this setting, we can establish that the estimator $\widehat{\ell}_t = \frac{c_t - \theta \mathbf{1}}{1-2\theta}$ is unbiased:

$$\mathsf{E}[\widehat{\ell}_{tj}] = \frac{\mathsf{E}[c_{tj}] - \theta}{1-2\theta} = \frac{\ell_{tj}(1-\theta) + (1-\ell_{tj})\theta - \theta}{1-2\theta} = \ell_{tj}.$$

Moreover, the mean squared error $\mathsf{E}[(\widehat{\ell}_{tj} - \ell_{tj})^2]$ satisfies

$$\begin{aligned}\mathsf{E}[(\widehat{\ell}_{tj} - \ell_{tj})^2] &= \mathsf{E}\left[\left(\frac{c_{tj}-\theta}{1-2\theta} - \ell_{tj}\right)^2\right] \\ &= \mathsf{E}\left[\frac{c_{tj} - \theta - (1-2\theta)\ell_{tj}}{1-2\theta}\right] \\ &\leq \frac{1}{(1-2\theta)^2}.\end{aligned}$$

Using

$$\alpha = (1-2\theta)\sqrt{\frac{\log m}{2n}},$$

Theorem 1 applies to yield

$$\mathrm{Reg}(\widehat{p}^{\mathrm{EW}}, \mathrm{BSC}(\theta), \ell^n) \leq \frac{2\sqrt{2n\log m}}{(1-2\theta)}. \tag{15}$$

2.1.2 Binary erasure channel

Recall that for the BEC(e), for $\ell_t \in \{0,1\}^m$, and $c_{tj} = \ell_{tj}$ with probability $1-\mathsf{e}$, and is equal to erasure symbol "?" with probability e. In this setting, the estimator $\widehat{\ell}_{tj} = \frac{c_{tj}\mathbb{1}\{c_{tj}\neq ?\}}{1-\mathsf{e}} = \frac{\ell_{tj}\mathbb{1}\{c_{tj}\neq ?\}}{1-\mathsf{e}}$ is unbiased:

$$\mathsf{E}\left[\frac{\ell_{tj}\mathbb{1}\{c_{tj}\neq ?\}}{1-\mathsf{e}}\right] = \ell_{tj}\frac{\mathbb{P}[c_{tj}\neq ?]}{1-\mathsf{e}} = \ell_{tj}$$

Moreover, the mean squared error $\mathsf{E}[(\widehat{\ell}_{tj} - \ell_{tj})^2]$ satisfies

$$\begin{aligned}\mathsf{E}[(\widehat{\ell}_{tj} - \ell_{tj})^2] &= \mathsf{E}\left[\frac{(1-\mathsf{e})\ell_{tj} - \ell_{tj}\mathbb{1}\{c_{tj}\neq ?\}}{1-\mathsf{e}}\right]^2 \\ &= \frac{\mathsf{E}[(1-\mathsf{e})^2\ell_{tj}^2 + \ell_{tj}^2\mathbb{1}\{c_{tj}\neq ?\} - 2\ell_{tj}^2(1-\mathsf{e})\mathbb{1}\{c_{tj}\neq ?\}]}{(1-\mathsf{e})^2} \\ &= \frac{(1-\mathsf{e})\ell_{tj}^2 - (1-\mathsf{e})^2\ell_{tj}^2}{(1-\mathsf{e})^2} \\ &= \frac{\ell_{tj}^2\mathsf{e}}{1-\mathsf{e}} \leq \frac{1}{1-\mathsf{e}}.\end{aligned}$$

Therefore, using $\alpha = \sqrt{\frac{(1-\mathsf{e})\log m}{n}}$ and observing that $-\alpha\ell_{tj} \leq 0$, we have by Theorem 1 that

$$\mathrm{Reg}(\widehat{p}^{\mathrm{EW}}, \mathrm{BEC}(\mathsf{e}), \ell^n) \leq 2\sqrt{\frac{2n\log m}{1-\mathsf{e}}}. \tag{16}$$

2.1.3 Additive noise channels

Recall that for additive noise channels, the output $c_{tj} = \ell_{tj} + Z_{tj}$ where all the Z_{tj} are independently and identically distributed. We now establish achievability results for various (classes of) noise distributions.

Gaussian noise. Consider $c_t = \ell_t + Z_t$ where $Z_t \sim \mathcal{N}(0, \sigma^2 I)$. The most natural unbiased estimator to use is simply $\widehat{\ell}_t = c_t$, with MSE $\mathsf{E}[(c_t - \ell_t)^2] = \sigma^2$. Note that in this case since the noise is unbounded $-\alpha c_{tj}$ can be arbitrarily large—but the probability of this event occuring is exponentially small. In particular, recalling from (14) that the event $\mathcal{E}$ is defined as

$$\mathcal{E} := \{\exists t, j : -\alpha c_{tj} \geq 1\}$$

we note for $\alpha = \sqrt{\frac{\log m}{n(1+\sigma^2)}}$

$$\begin{aligned}
\mathbb{P}(\mathcal{E}) &\overset{(a)}{\leq} \sum_{t=1}^{n}\sum_{j=1}^{m} \mathbb{P}\left(c_{tj} \leq -\sqrt{\frac{n(1+\sigma^2)}{\log m}}\right) \\
&\overset{(b)}{\leq} \sum_{t=1}^{n}\sum_{j=1}^{m} \mathbb{P}\left(Z_{tj} \leq -\sqrt{\frac{n(1+\sigma^2)}{\log m}}\right) \\
&\leq mn\mathbb{P}\left(Z_{tj} \leq -\sqrt{\frac{n(1+\sigma^2)}{\log m}}\right) && (17) \\
&\overset{(c)}{\leq} mn \exp\left(-\frac{n}{2\log m}\right) && (18)
\end{aligned}$$

where (a) follows by the union bound, (b) follows since $c_{tj} = Z_{tj} + \ell_{tj}$ and $0 \leq \ell_{tj} \leq 1$, and (c) follows by using that for $Z \sim \mathcal{N}(0, \sigma^2)$ the complementary CDF $\mathbb{P}(Z \geq x) \leq \exp(-x^2/2\sigma^2)$. Thus, Theorem 1 implies that the strategy $\widehat{p}^{\text{EW}}$ which sets $p_t = p^{\text{EW}}(c^{t-1})$ with learning rate $\alpha = \sqrt{\frac{\log m}{n(1+\sigma^2)}}$ achieves regret

$$\text{Reg}(\widehat{p}^{\text{EW}}, \text{AWGN}(\sigma), \ell^n) \leq 2\sqrt{(1+\sigma^2) n \log m} + o(n). \tag{19}$$

Uniform noise. For additive channels with uniform noise, the channel output $c_{tj} = \ell_{tj} + Z_{tj}$ where $Z_{tj} \sim \text{Unif}[-\sigma, \sigma]$ (so that the noise variance is $\sigma^2/3$). Since we are interested in how the regret scales as σ increases, it suffices to assume that $\sigma \geq 1$. Then, consider the following estimator $\widehat{\ell}_t$ (which is a function of c_t):

$$\widehat{\ell}_{tj} = \begin{cases} -\sigma + \frac{1}{2} & \text{if } -\sigma \leq c_{tj} < -\sigma + 1 \\ \frac{1}{2} & \text{if } -\sigma + 1 \leq c_{tj} \leq \sigma \\ \sigma + \frac{1}{2} & \text{if } \sigma < c_{tj} \leq \sigma + 1. \end{cases} \tag{20}$$

We observe that (full calculations relegated to Appendix B)

$$\mathsf{E}[\widehat{\ell}_{tj}] = \ell_{tj} \tag{21}$$

i.e. $\widehat{\ell}_t$ is unbiased and that the MSE for this estimator satisfies

$$\mathsf{E}[\widehat{\ell}_{tj} - \ell_{tj}]^2 \leq \sigma. \tag{22}$$

For choice of learning rate $\alpha = \sqrt{\frac{\log m}{n(1+\sigma)}}$, we note that $\alpha\widehat{\ell}_{tj} \geq -\sigma\alpha = -\sigma\sqrt{\frac{\log m}{n(1+\sigma)}} \geq -1$ for large enough n. Therefore, if we use the strategy $\widehat{p}^{\mathrm{EW}}$ with the unbiased estimator $\widehat{\ell}_t$ in (20), Theorem 1 yields

$$\mathrm{Reg}(\widehat{p}^{\mathrm{EW}}, \mathrm{Unif}(\sigma), \ell^n) \leq 2\sqrt{(1+\sigma)n\log m}. \tag{23}$$

In Section 3, we show a matching converse to (23) and establish that the regret must grow as $\Omega\left(\sqrt{(1+\sigma)n\log m}\right)$, showing the tightness of (23).

Symmetric noise with tail constraints. If the additive noise is symmetric, i.e. the distribution of noise Z and $-Z$ is the same, the most natural unbiased estimator for ℓ_t is $\widehat{\ell}_t = c_t$ (since the noise is additive and $0-$mean) which achieves mean-square error $\mathsf{E}[c_{tj} - \ell_{tj}]^2 = \sigma^2$ where σ^2 is the variance of the noise Z_{tj}. In order to apply Theorem 1 with $\alpha = \sqrt{\frac{\log m}{n(1+\sigma^2)}}$ to achieve regret scaling as $O(\sqrt{(1+\sigma^2)n\log m})$ (as in the AWGN channel setting) we need to establish a bound on $\mathbb{P}(\mathcal{E})$. Following the line of reasoning employed to reach (17) we have

$$\mathbb{P}(\mathcal{E}) \leq mn\mathbb{P}\left(\frac{Z_{tj}}{\sigma} \leq -\sqrt{\frac{n(1+\sigma^2)}{\sigma^2\log m}}\right) \leq mn\mathbb{P}\left(\frac{Z_{tj}}{\sigma} \leq -\sqrt{\frac{n}{\log m}}\right)$$

which implies that a noise density with polynomially decaying tails (in particular for $\sigma = 1$, if the random variable Z satisfies for large x that $\mathbb{P}(Z \geq x) \leq \frac{c}{x^{6+\epsilon}}$ where c is a positive absolute constant and $\epsilon > 0$) suffices to achieve regret

$$2\sqrt{(1+\sigma^2)n\log m} + o(n). \tag{24}$$

An important class of distributions that achieves this tail condition is *log-concave* distributions [Saumard and Wellner, 2014], which are distributions having density $f(z)$ for which the function $z \mapsto \log f(z)$ is concave. This class has a special significance across statistics and information theory and includes distributions such as the Gaussian distribution, the uniform distribution and the Laplace distribution. Since all log-concave distributions are subexponential (i.e. have exponentially decaying tails) these satisfy aforementioned the condition on $\mathbb{P}(\mathcal{E})$ as n grows larger. While this achievability result for a general class of noise densities is interesting, unfortunately we do not have a general matching converse.

While for the specific cases of Gaussian and Laplace densities, it is possible to achieve a matching $\Omega(\sqrt{(1+\sigma^2)n\log m})$ lower bound for the regret, the most general converse we are able to achieve is a fundamental lower bound of $\Omega(\sqrt{(1+\sigma)n\log m}$ on the regret when the class of noise densities is log-concave. While it might appear that the converse can be strengthened in general, we have seen that this fundamental lower bound can in fact be achieved for uniform noise distributions by constructing a different unbiased estimator that achieves $O(\sqrt{(1+\sigma)n\log m})$ regret.

2.1.4 1-bit/expert quantization

Unlike the previous three examples (memoryless noise), in this case, the channel is not fixed. Instead, the decision maker must select an encoding protocol that maps $\mathcal{L}^t \to \mathcal{M} := [2^m]$ and a decision protocol $\mathcal{M}^{t-1} \to \Delta^{m-1}$. The real number ℓ_{tj} must be transmitted in an unbiased manner. A randomized rounding $\widehat{\ell}_{tj} \sim \mathrm{Bern}(\ell_{tj})$ achieves this with MSE $\mathsf{E}[(\widehat{\ell}_{tj} - \ell_{tj})^2] = \ell_{tj}(1 - \ell_{tj}) \leq 1$. Since the encoding is memoryless and because for $\alpha > 0$ we have $-\alpha\widehat{\ell}_{tj} \leq 0$ we can readily apply Theorem 1 with $\alpha = \sqrt{\frac{\log m}{2n}}$ to achieve

$$\mathrm{Reg}(\widehat{p}^{\mathrm{EW}}, m\text{-bit}, \ell^n) \leq 2\sqrt{2n\log m}. \tag{25}$$

Remark 2 (Adaptive quantization) *The achievability scheme is a simple memoryless (randomized) encoding. Since this is shown to achieve the order-wise (in m and n) optimal regret (see Section 3), an adaptive encoding strategy offers no extra benefit in terms of regret. However, it is unclear if for rate strictly less that 1 bit/dimension a nonadaptive encoding strategy will continue to achieve the optimal regret.*

Remark 3 (Universality in channel parameters) *We note that for the BSC, BEC, and additive noise channels in the strategy outlined above, one needs to set the optimal learning rate α as a function of the channel parameter θ, e or σ. An intriguing question is that of devising a no-regret strategy that does not depend on these parameters— [Weissman and Merhav, 2001] establish such a strategy for the BSC that achieves, up to constants, the same regret as in* (15)*. We leave the question of designing such strategies for the BEC and additive noise channels for future work.*

2.2 Proof of Theorem 1

Define the "bad" event $\mathcal{E} := \{\exists t, j : -\alpha\widehat{\ell}_{tj} > 1\}$, which by the condition stated in the Theorem occurs with probability $\mathbb{P}(\mathcal{E})$. We will split the regret analysis into two cases: if $\mathcal{E}^C$ occurs, where Lemma 1 can be invoked, and if $\mathcal{E}$ occurs, where we will utilize a worst-case bound on regret. First, we use Proposition 1 which yields

$$\mathrm{Reg}(\widehat{p}^{\mathrm{EW}}, P_{c|\ell}, \ell^n) \le \mathsf{E}\left[\sum_{t=1}^{n}\langle p^{\mathrm{EW}}(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle - \min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}\right] \tag{26}$$

and we have

$$\begin{aligned}
&\mathsf{E}\left[\sum_{t=1}^{n}\langle p^{\mathrm{EW}}(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle - \min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}\right] \\
&= \mathsf{E}\left[\Big(\sum_{t=1}^{n}\langle p^{\mathrm{EW}}(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle - \min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}\Big)\mathbb{1}\{\mathcal{E}^C\}\right] \\
&\qquad + \mathsf{E}\left[\sum_{t=1}^{n}\langle p^{\mathrm{EW}}(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle\mathbb{1}\{\mathcal{E}\}\right] - \mathsf{E}\left[\Big(\min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}\Big)\mathbb{1}\{\mathcal{E}\}\right].
\end{aligned} \tag{27}$$

We analyze the three terms in the right hand side of (27) separately. First, note that if $\mathcal{E}^C$ occurs, then the conditions in Lemma 1 are satisfied which can be employed to get

$$\left(\sum_{t=1}^{n}\langle p^{\mathrm{EW}}(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle - \min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}\right)\mathbb{1}\{\mathcal{E}^C\} \le \frac{\log m}{\alpha} + \alpha\sum_{t=1}^{n}\sum_{j=1}^{m}p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})\widehat{\ell}_{tj}^2 \tag{28}$$

where (28) also uses that indicator is bounded by 1 and the term to be multiplied is positive. Next, note that

$$\begin{aligned}
\mathsf{E}[p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})\widehat{\ell}_{tj}^2] &= \mathsf{E}[p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})\ell_{tj}^2] + \mathsf{E}[p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})(\widehat{\ell}_{tj}-\ell_{tj})^2] + \mathsf{E}[2p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})\ell_{tj}(\widehat{\ell}_{tj}-\ell_{tj})] \\
&\overset{(a)}{=} \mathsf{E}[p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})\ell_{tj}^2] + \mathsf{E}[p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})]\,\mathsf{E}[(\widehat{\ell}_{tj}-\ell_{tj})^2] \\
&\overset{(b)}{\le} \mathsf{E}[p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})]\big(1 + \max_{t,j}\mathsf{E}[(\widehat{\ell}_{tj}-\ell_{tj})^2]\big)
\end{aligned} \tag{29}$$

where (a) follows from the fact that $\widehat{\ell}_t$ is independent of $\widehat{\ell}^{t-1}$ and that $\widehat{\ell}_t$ is unbiased, and (b) uses that $\ell_{tj}^2 \leq 1$ by assumption. Taking expectations on both sides of (28) and (29) yields

$$\begin{aligned}
\mathsf{E}\Big[\Big(\sum_{t=1}^{n}\langle p^{\mathrm{EW}}(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle - \min_{j\in[m]}\sum_{t=1}^{n}\widehat{\ell}_{tj}\Big)\mathbb{1}\{\mathcal{E}^C\}\Big] \\
&\leq \frac{\log m}{\alpha} + \alpha\sum_{t=1}^{n}\big(1+\max_{t,j}\mathsf{E}[(\widehat{\ell}_{tj}-\ell_{tj})^2]\big)\,\mathsf{E}\Big[\sum_{j=1}^{m}p_{tj}^{\mathrm{EW}}(\widehat{\ell}^{t-1})\Big] \\
&= \frac{\log m}{\alpha} + \alpha n\left(1+\max_{t,j}\mathsf{E}[(\widehat{\ell}_{tj}-\ell_{tj})^2]\right). \qquad (30)
\end{aligned}$$

To bound the second term in (27), we apply

$$\begin{aligned}
\sum_{t=1}^{n}\mathsf{E}\left[\langle p^{\mathrm{EW}}(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle\mathbb{1}\{\mathcal{E}\}\right] &\leq \sum_{t=1}^{n}\mathsf{E}\left[|\langle p^{\mathrm{EW}}(\widehat{\ell}^{t-1}), \widehat{\ell}_t\rangle|\mathbb{1}\{\mathcal{E}\}\right] \\
&\overset{(a)}{\leq} \sum_{t=1}^{n}\mathsf{E}\left[\|\widehat{\ell}_t\|_\infty\mathbb{1}\{\mathcal{E}\}\right] \\
&\overset{(b)}{\leq} \sum_{t=1}^{n}\mathsf{E}\left[\sum_{j=1}^{m}|\widehat{\ell}_{tj}|\mathbb{1}\{\mathcal{E}\}\right] \\
&\overset{(c)}{\leq} \sum_{t=1}^{n}\sum_{j=1}^{m}\sqrt{\mathsf{E}[\widehat{\ell}_{tj}]^2}\sqrt{\mathbb{P}(\mathcal{E})} \\
&\overset{(d)}{\leq} mn\sqrt{\left(1+\max_{t,j}\mathsf{E}[\widehat{\ell}_{tj}-\ell_{tj}]^2\right)\mathbb{P}(\mathcal{E})} \qquad (31)
\end{aligned}$$

where (a) uses the Holder inequality and the fact that $p^{\mathrm{EW}}(c^{t-1})$ is a probability distribution, (b) uses the fact that the absolute maximum in a vector is bounded by the sum of the absolute values, (c) uses the Cauchy–Schwartz inequality, and (d) uses unbiasedness of $\widehat{\ell}_{tj}$ along with the fact that $\ell_{tj}^2 \leq 1$. The third term in (27) will be dealt with similarly:

$$\begin{aligned}
\mathsf{E}\left[\Big(-\min_{j}\sum_{t=1}^{n}\widehat{\ell}_{tj}\Big)\mathbb{1}\{\mathcal{E}\}\right] &\leq \sum_{j=1}^{m}\sum_{t=1}^{n}\mathsf{E}\left[|\widehat{\ell}_{tj}|\mathbb{1}\{\mathcal{E}\}\right] \\
&\leq mn\sqrt{\left(1+\max_{t,j}\mathsf{E}[\widehat{\ell}_{tj}-\ell_{tj}]^2\right)\mathbb{P}(\mathcal{E})} \qquad (32)
\end{aligned}$$

where (32) follows from (31). Finally, using (30), (31) and (32) in (27) concludes the proof.

3 Converse

In this section, we establish fundamental lower bounds on the regret $\max_{\ell^n}\mathrm{Reg}(p, P_{\mathsf{c}|\ell}, \ell^n)$ for any strategy p. Recall the converse for the standard prediction with expert advice problem [Cesa-Bianchi et al., 1997a, Cesa-Bianchi and Lugosi, 2006] where the feedback is available noiselessly from (2):

$$\max_{\ell^n}\mathrm{Reg}(p, \ell^n) \geq \Omega(\sqrt{n\log m}). \qquad (33)$$

While this naturally serves as a lower bound on the regret incurred when the feedback is noisy, we aim to achieve a tighter bound by further quantifying the extra penalty incurred due to the noise in the feedback. We take the usual approach [Cesa-Bianchi et al., 1997a, Cesa-Bianchi and Lugosi, 2006, Slivkins, 2019] of reducing the experts problem to a hypothesis testing problem via a carefully chosen ensemble of (random) loss functions and leveraging fundamental lower bounds for hypothesis testing from information theory. To quantify the factor by which the regret must increase, we need the following definition.

Definition 1 *The strong data processing constant of a binary-input channel $P_{Y|X}$ is defined as*

$$\eta(P_{Y|X}) = \sup_{P_X \neq Q_X} \frac{D(P_X \circ P_{Y|X} \| Q_X \circ P_{Y|X})}{D(P_X \| Q_X)} \tag{34}$$

where P_X and Q_X are distributions defined on $\{0, 1\}$.

Intuitively, this measure quantifies some sense of “loss of information" in a noisy channel—this interpretation is more clear by the alternate representation of $\eta(P_{Y|X})$ (see [Polyanskiy and Wu, 2022, Theorem 33.5])

$$\eta(P_{Y|X}) = \sup_{P_{UX}: U \to X \to Y} \frac{I(U;Y)}{I(U;X)} \tag{35}$$

where U is an auxiliary random variable, and $U \to X \to Y$ represents a Markov chain. The data processing inequality [Cover and Thomas, 2006] from information theory immediately implies that

$$\eta(P_{Y|X}) \leq 1;$$

often, as we show subsequently, we can establish $\eta(P_{Y|X}) < 1$. There has been much interest in characterizing $\eta(P_{Y|X})$ for various channels due to numerous applications arising in the domain of statistical inference—see [Polyanskiy and Wu, 2017, Raginsky, 2016], [Polyanskiy and Wu, 2022, Chapter 33] for a detailed survey.

We state our general converse result next.

Theorem 2 *If the noise is memoryless and component-wise independent (i.e. $P_{c|\ell} = \prod_{j=1}^{m} \mathsf{P}_{c_j|\ell_j}$) then*

$$\sup_{\ell^n} \mathrm{Reg}(p, P_{c|\ell}, \ell^n) \geq \sqrt{\frac{n \log(m/4)}{16 \eta(\mathsf{P}_{c|\ell})}} \tag{36}$$

where with some abuse of notation, $\eta(\mathsf{P}_{c|\ell})$ (as in Definition 1) restricts the channel to binary input $\{0, 1\}$.

Remark 4 *The restriction to binary inputs appears because in order to prove a converse for the individual-sequence regret $\sup_{\ell^n} \mathrm{Reg}(p, P_{c|\ell}, \ell^n)$, we need to construct an ensemble of loss functions that achieves high regret, and binary loss functions suffice for this purpose.*

3.1 Application of Theorem 2 to canonical channel models

Before presenting the proof of Theorem 2, we instantiate the achieved bound for all the example channels considered earlier. In each case, one needs to bound $\eta(\mathsf{P}_{c|\ell_{c|\ell}})$.

3.1.1 Binary symmetric channel BSC(θ)

In this case, it is easy to evaluate (see for example [Cover and Thomas, 2006, Exercise 7.7]) that

$$\eta(\mathrm{BSC}(\theta)) = (1-2\theta)^2$$

This implies for a large enough n that

$$\sup_{\ell^n} \mathrm{Reg}(p, \mathrm{BSC}(\theta), \ell^n) \geq \sqrt{\frac{n\log(m/4)}{16(1-2\theta)^2}}, \tag{37}$$

matching up-to constants the achievability result in (15). This furthermore implies, as intuitively expected, that as $\theta \to \frac{1}{2}$ it is impossible to achieve any no-regret strategy since the corrupted feedback c_{tj} is a random Bern(1/2) bit carrying no information about ℓ_{tj}.

3.1.2 Binary erasure channel BEC(e)

For the BEC, is is convenient to use the characterization in (35) to evaluate $\eta(\mathrm{BEC}(\mathsf{e}))$. For any Markov chain $U \to X \to Y$, $I(U;Y,X) = I(U;X) + I(U;Y|X) = I(U;X)$, where the first equality follows by the chain rule of mutual information, and the second follows since U and Y are independent given X (by Markovity). Furthermore, $I(U;Y,X) = I(U;Y) + I(U;X|Y) = I(U;Y) + \mathsf{e}I(U;X)$ where the second equality here follows since $Y = X$ with probability $1-\mathsf{e}$, in which case the conditional mutual information is 0. Putting these two together, we see that

$$\eta(\mathrm{BEC}(\mathsf{e})) = (1-\mathsf{e})$$

which implies that

$$\sup_{\ell^n} \mathrm{Reg}(p, \mathrm{BEC}(\mathsf{e}), \ell^n) \geq \sqrt{\frac{n\log(m/4)}{16(1-\mathsf{e})}} \tag{38}$$

for a large enough n matching up-to constants the achievability result in (16); and offering the intuitive interpretation that as $\mathsf{e} \to 1$, since most of the feedback ℓ_{tj} gets erased it becomes impossible to make any meaningful decisions.

3.1.3 Additive noise channels

We now consider channels of the form $c_{tj} = \ell_{tj} + Z_{tj}$ for (independent and identically distributed) random variables Z_{tj}. To quantify $\eta(\mathsf{P}_{\mathsf{c}|\ell})$, we will utilize the following characterization from [Polyanskiy and Wu, 2017, Theorem 21]

Theorem 3 *For a binary-input channel $P_{Y|X}$,*

$$\frac{H^2(P_{Y|X=0}, P_{Y|X=1})}{2} \leq \eta(P_{Y|X}) \leq H^2(P_{Y|X=0}, P_{Y|X=1}) \tag{39}$$

where H represents the Hellinger divergence between two distributions.

We can now use this result for the specific noise models we are interested in.

Additive white Gaussian noise. If $Z_{tj} \sim \mathcal{N}(0,\sigma^2)$, then (39) implies that

$$\eta(\mathrm{AWGN}(\sigma^2)) = H^2(\mathcal{N}(0,\sigma^2), \mathcal{N}(1,\sigma^2))$$

$$= 1 - \frac{1}{2\pi\sigma^2}\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2} - \frac{(x-1)^2}{2\sigma^2}\right) dx \tag{40}$$

$$= 2 - 2e^{-1/8\sigma^2} \tag{41}$$

$$\leq \frac{4}{1+\sigma^2} \tag{42}$$

where (42) follows since $(1 - e^{-1/8x^2})(1 + x^2) \leq 4$ for all x. Using (42) in Theorem 2 implies that

$$\sup_{\ell^n} \mathrm{Reg}(p, \mathrm{BEC}(\mathsf{e}), \ell^n) \geq \sqrt{\frac{(1+\sigma^2)n\log(m/4)}{64}} \tag{43}$$

matching up to constants the achievability result in (19).

Additive uniform noise. The uniform additive noise channel has the noise $Z_{tj} \sim \mathrm{Unif}[-\sigma, \sigma]$—note that this noise has variance $\sigma^2/3$. In this case $P_{Y|X=0} = \mathrm{Unif}[-\sigma, \sigma]$ with density $f_0(x) = \frac{1}{2\sigma}\mathbb{1}\{-\sigma \leq x \leq \sigma\}$, and $P_{Y|X=1} = \mathrm{Unif}[-\sigma+1, \sigma+1]$ with density $f_1(x) = \frac{1}{2\sigma}\mathbb{1}\{-\sigma+1 \leq x \leq \sigma+1\}$. Let us assume that $\sigma \geq 1$; in this case

$$\begin{aligned}\eta(\mathrm{Unif}(\sigma)) &\leq H^2(\mathrm{Unif}[-\sigma, \sigma], \mathrm{Unif}[-\sigma+1, \sigma+1]) \\ &= 1 - \frac{1}{2\sigma}\int_{-\sigma+1}^{\sigma+1} dx \\ &= \frac{1}{2\sigma}.\end{aligned} \tag{44}$$

Combining (44) with the trivial bound $\eta \leq 1$ yields

$$\eta(\mathrm{Unif}(\sigma)) \leq \frac{1}{\sigma+1} \tag{45}$$

for all $\sigma > 0$; implying the fundamental lower bound on the regret when the feedback is corrupted with additive uniform noise

$$\sup_{\ell^n} \mathrm{Reg}(p, \mathrm{BEC}(\mathsf{e}), \ell^n) \geq \sqrt{\frac{(1+\sigma)n\log(m/4)}{16}} \tag{46}$$

matching the achievability result obtained in (23) up to constants.

Additive symmetric, log-concanve noise. So far, in the additive noise examples we have considered (Gaussian and uniform noise), we established that noisy feedback incurs a multiplicative cost (over the noiseless case) on the regret that depends on the moments of the noise and this cost is strictly greater than 1 ($\sqrt{1+\sigma^2}$ and $\sqrt{1+\sigma}$ respectively). In light of the achievability result in (24), we might hope that for general additive noise channels with mild tail conditions on the noise one can achieve $\eta(\mathsf{P}_{\mathsf{c}|\ell}) \geq \Omega(\sigma)$. Unfortunately, this is not the case in general—consider the additive channel $Y = X + Z$ with noise distribution $Z \sim \mathrm{Uniform}\{-\sigma, \sigma\}$—this noise distribution is bounded; but still $\eta(P_{Y|X}) = 1$ since given Y, X is perfectly known. Therefore, to obtain a more general result, more conditions need to be imposed on the noise distribution.

We will show a converse for the general class of symmetric log-concave distributions considered in Section 2.1, which encompasses the Gaussian and uniform distributions considered previously. Consider a log-concave noise distribution with variance σ^2 and let f denote its density. Then,

$$\eta(P_{Y|X}) \leq H^2(P_{Y|X=0}, P_{Y|X=1})$$

$$\overset{(a)}{\leq} 2TV(P_{Y|X=0}, P_{Y|X=1})$$
$$\overset{(b)}{=} \int_{-\infty}^{\infty} |f(z) - f(z-1)| dz \tag{47}$$

where (a) follows from the well known inequality $H^2 \leq 2TV$ between Hellinger and total variation distances, and (b) follows from the definition of the total variation distance (and, the fact that the density of $Y|X = 1$ is $f(z-1)$). Next, we further simplify (47) using the symmetry and unimodality of f (since any log-concave distribution is also unimodal). Since $f(z)$ is decreasing for $z \geq 0$ and $f(z-1)$ is increasing for $z \leq 1$, for any $z \leq \frac{1}{2}$, we have

$$f(z-1) \leq f(1/2) \leq f(z)$$

and similarly for $z > \frac{1}{2}$, $f(z) \leq f(z-1)$. Therefore,

$$\begin{aligned}
\int_{-\infty}^{\infty} |f(z) - f(z-1)| dz &= \int_{-\infty}^{1/2} (f(z) - f(z-1)) dz + \int_{1/2}^{\infty} (f(z-1) - f(z)) dz \\
&= 2 \int_{1/2}^{\infty} (f(z-1) - f(z)) dz \qquad (48) \\
&= 2 \left(\int_{1/2}^{\infty} f(z-1) dz - \int_{1/2}^{\infty} f(z) dz \right) \\
&= 2 \left(\int_{-1/2}^{\infty} f(z) dz - \int_{1/2}^{\infty} f(z) dz \right) \\
&= 2 \left(\int_{-1/2}^{1/2} f(z) dz \right) \\
&\leq \frac{4}{\sigma} \qquad (49)
\end{aligned}$$

where (49) is due to the following proposition.

Proposition 2 *For a symmetric, log-concave distribution with variance σ^2, its density satisfies $f(z) \leq \frac{2}{\sigma}$.*

Proposition 2 appears in [Marsiglietti and Kostina, 2018, Remark 6], but we provide a standalone proof in Appendix C for completeness.

Putting together (49) and (47) along with the trivial bound $\eta \leq 1$, we see that for any additive noise channel with a symmetric, log-concave density

$$\eta(P_{Y|X}) \leq \frac{8}{(1+\sigma)}. \tag{50}$$

This furthermore implies in the experts problem that if feedback is available with additive noise $c_{tj} = \ell_{tj} + Z_{tj}$ where Z_{tj} is symmetric and log-concave, then

$$\sup_{\ell^n} \mathrm{Reg}(p, \mathrm{BEC}(\mathsf{e}), \ell^n) \geq \sqrt{\frac{(1+\sigma) n \log(m/4)}{128}}. \tag{51}$$

It is interesting to note that the converse in (51) is not tight in general. In particular for Gaussian noise and Laplace (double exponential) additive noise (for both, we can establish a $\sqrt{1+\sigma^2}$ scaling by direct computation of $H^2(P_{Y|X=0}, P_{Y|X=1})$. Nonetheless, it is tight for uniform noise, which is a log-concave distribution, as we have shown a matching achievability result in (23). Thus, it is tight in the sense that it cannot be improved without imposed further restrictions on the class of noise densities.

3.1.4 1−bit/expert quantization

Strictly speaking, Theorem 2 cannot be applied to the quantization setting—we need to show a fundamental lower bound over *all* quantization strategies including strategies with memory; Theorem 2 requires the noise to be memoryless and independent for each expert. Nonetheless, we can bypass this limitation to establish a converse matching up to constants the achievability result obtained in (25) for *any* encoding strategy (possibly adaptive). This is because a fundamental lower bound on the noiseless regret is automatically a converse for the noisy setting; see Remark 5 for a more detailed argument. Theorefore, one must necessarily incur $\Omega(\sqrt{n \log m})$ regret for any quantization method—interestingly, our achievability result in (25) shows that a simple, memoryless quantizer suffices to achieve this fundamental limit.

3.2 Proof of Theorem 2

Consider the following (random) ensemble of loss vectors:

- Pick $J^* \sim \text{Uniform}[m]$.
- Given $J^* = j^*$, the loss vectors ℓ^n are generated i.i.d., with independent components as per the distribution
$$\ell_{tj} \sim \begin{cases} \text{Bern}(1/2 - \epsilon), & \text{if } j = j^* \\ \text{Bern}(1/2), & \text{otherwise} \end{cases} \tag{52}$$
for some $0 < \epsilon < 1/4$ to be determined later.

Intuitively, in order to achieve sublinear regret in n with these loss functions, the decision-maker must eventually detect the expert j^* that has the lowest bias and therefore this can be thought of as a hypothesis testing problem. To formalize this, we have

$$\sup_{\widetilde{\ell}^n} \text{Reg}(p, P_{c|\ell}, \widetilde{\ell}^n) \geq \mathsf{E}\left[\sum_{t=1}^{n} \langle p_t(c^{t-1}), \ell_t \rangle\right] - \mathsf{E}\left[\min_{j\in[m]} \sum_{t=1}^{n} \ell_{tj}\right]. \tag{53}$$

Now, note that

$$\mathsf{E}\left[\min_{j\in[m]} \sum_{t=1}^{n} \ell_{tj}\right] \overset{(a)}{\leq} \mathsf{E}\left[\min_{j\in[m]} \mathsf{E}\left[\sum_{t=1}^{n} \ell_{tj} \middle| J^*\right]\right] \tag{54}$$
$$\overset{(b)}{=} n\left(\frac{1}{2} - \epsilon\right) \tag{55}$$

where (a) follows since $\mathsf{E}[\min(\cdot)] \leq \min \mathsf{E}[\cdot]$ and (b) follows since by the distribution on the losses in (52)

$$\mathsf{E}[\ell_{tj}|J^*] = \begin{cases} \frac{1}{2}, & \text{if } j = J^* \\ \frac{1}{2} - \epsilon. & \text{otherwise} \end{cases} \tag{56}$$

To further bring out the analogy between hypothesis testing and the regret, we note that for the random variable distributed as $J_t \sim p_t(c^{t-1})$ conditional on c^{t-1} (i.e. a random expert is chosen as per the distribution $p_t(c^{t-1})$)

$$\mathsf{E}[\langle p_t(c^{t-1}), \ell_t \rangle | c^{t-1}, \ell_t] = \mathsf{E}[\ell_{tJ_t} | c^{t-1}, \ell_t],$$

and therefore

$$\mathsf{E}[\langle p_t(c^{t-1}), \ell_t \rangle] = \mathsf{E}[\ell_{tJ_t}].$$

Then,

$$\begin{aligned}\mathsf{E}[\langle p_t(c^{t-1}), \ell_t \rangle] &= \mathsf{E}[\mathsf{E}[\ell_{tJ_t} | J_t]] \\ &\overset{(a)}{=} \mathsf{E}\left[\frac{1}{2}\mathbb{1}\{J_t \neq J^*\} + \left(\frac{1}{2} - \epsilon\right)\mathbb{1}\{J_t = J^*\}\right] \\ &= \frac{1}{2} - \epsilon \mathbb{P}[J_t = J^*] \end{aligned} \tag{57}$$

where (a) follows from (56). Using (57) along with (55) and (53) yields

$$\sup_{\widetilde{\ell}^n} \operatorname{Reg}(p, P_{c|\ell}, \widetilde{\ell}^n) \geq \epsilon \sum_{t=1}^{n} \mathbb{P}[J_t \neq J^*]. \tag{58}$$

To further lower bound the regret, we apply the Fano inequality to each term in the right hand side of (58)

$$\begin{aligned}\mathbb{P}[J_t(c^{t-1}) \neq J^*] &\geq 1 - \frac{I(J^*; J_t) + \log 2}{\log m} &(59)\\ &\geq 1 - \frac{I(J^*; c^{t-1}) + \log 2}{\log m}, &(60)\end{aligned}$$

where (60) follows by the data processing inequality since $J^* \to c^{t-1} \to J_t$.

Since the noise is memoryless by assumption,

$$\begin{aligned}I(J^*; c^t) &= H(c^t) - H(c^t | J^*) \\ &\overset{(a)}{\leq} \sum_{i=1}^{t} H(c_i) - H(c^t | J^*) \\ &\overset{(b)}{=} \sum_{i=1}^{t} (H(c_i) - H(c_i | J^*)) \\ &= \sum_{i=1}^{t} I(J^*; c_i) \\ &\overset{(c)}{\leq} tI(J^*; c_1). \end{aligned} \tag{61}$$

where (a) follows by the subadditivity of entropy, (b) follows since given J^*, c^t are independent (because given J^*, ℓ^t are independent as per (52) and the channel is memoryless by assumption), and finally (c) follows by symmetry (l^t are identically distributed, therefore so are c^t). Next, we have

$$I(J^*; c_1) = D\left(P_{c_1|J^*} \| P_{c_1} | P_{J^*}\right)$$

$$
\begin{aligned}
&= \frac{1}{m}\sum_{j=1}^{m} D\left(P_{c_1|J^*=j}\|P_{c_1}\right) \\
&= \frac{1}{m}\sum_{j=1}^{m} D\left(P_{c_1|J^*=j}\Bigg\|\frac{1}{m}\sum_{j'=1}^{m} P_{c_1|J^*=j'}\right) \\
&\overset{(a)}{\leq} \frac{1}{m^2}\sum_{j=1}^{m}\sum_{j'=1}^{m} D\left(P_{c_1|J^*=j}\|P_{c_1|J^*=j'}\right) \\
&\overset{(b)}{=} \frac{m^2-m}{m^2} D\left(P_{c_1|J^*=1}\|P_{c_1|J^*=2}\right) \\
&\leq D\left(P_{c_1|J^*=1}\|P_{c_1|J^*=2}\right) \\
&\overset{(c)}{=} \sum_{j=1}^{m} D(P_{c_{1j}|J^*=1}\|P_{c_{1j}|J^*=2}) \\
&\overset{(d)}{=} D(P_{c_{11}|J^*=1}\|P_{c_{12}|J^*=2}) + D(P_{c_{12}|J^*=1}\|P_{c_{12}|J^*=2}) \\
&= D\left(\text{Bern}(1/2-\epsilon)\circ \mathsf{P}_{\mathsf{c}|\ell}\|\text{Bern}(1/2)\circ \mathsf{P}_{\mathsf{c}|\ell}\right) \\
&+ D\left(\text{Bern}(1/2)\circ \mathsf{P}_{\mathsf{c}|\ell}\|\text{Bern}(1/2-\epsilon)\circ \mathsf{P}_{\mathsf{c}|\ell}\right) \qquad (62)
\end{aligned}
$$

where (a) follows since $D(P\|Q)$ is convex in the pair P and Q, (b) follows by symmetry, (c) follows since the vector c_1 has a product distribution given J^* (because ℓ_1 has a product distribution and the noise is component-wise independent) and (d) follows since all the other components except the first and second have the same distribution ($\text{Bern}(1/2)\circ \mathsf{P}_{\mathsf{c}|\ell}$). Recalling the definition of $\eta(\mathsf{P}_{\mathsf{c}|\ell})$ in Definition 1, we have

$$
\begin{aligned}
D\left(\text{Bern}(1/2-\epsilon)\circ \mathsf{P}_{\mathsf{c}|\ell}\|\text{Bern}(1/2)\circ \mathsf{P}_{\mathsf{c}|\ell}\right) &\leq \eta(\mathsf{P}_{\mathsf{c}|\ell})\left(d\left(\frac{1}{2}-\epsilon\Big\|\frac{1}{2}\right)\right) \\
&\leq \eta(\mathsf{P}_{\mathsf{c}|\ell})\epsilon^2 \qquad (63)
\end{aligned}
$$

where $d(\cdot\|\cdot)$ denotes the binary KL divergence, and the final inequality follows since $\frac{d\left(\frac{1}{2}-x\|\frac{1}{2}\right)}{x^2} \leq 1$ for $x < 1/4$ and $\epsilon < 1/4$ by assumption. Using the same reasoning for the second term of (62), and using (63) in (61) we have

$$
I(J^*; c^t) \leq 2t\eta(\mathsf{P}_{\mathsf{c}|\ell})\epsilon^2
$$

and therefore from (58) and (60) we get

$$
\begin{aligned}
\sup_{\widetilde{\ell}^n} \text{Reg}(p, P_{c|\ell}, \widetilde{\ell}^n) &\geq \epsilon \sum_{t=1}^{n}\left(1 - \frac{2(t-1)\eta(\mathsf{P}_{\mathsf{c}|\ell})\epsilon^2 + \log 2}{\log m}\right) \\
&\geq n\epsilon\left(1 - \frac{2n\eta(\mathsf{P}_{\mathsf{c}|\ell})\epsilon^2 + \log 2}{\log m}\right). \qquad (64)
\end{aligned}
$$

Finally, the choice of $\epsilon = \sqrt{\frac{\log(m/4)}{4n\eta(\mathsf{P}_{\mathsf{c}|\ell})}}$ (which guarantees $\epsilon \leq 1/4$ for a large enough n) in (64) yields

$$
\sup_{\widetilde{\ell}^n} \text{Reg}(p, P_{c|\ell}, \widetilde{\ell}^n) \geq \sqrt{\frac{n\log(m/4)}{16\eta(\mathsf{P}_{\mathsf{c}|\ell})}} \qquad (65)
$$

as claimed.

Remark 5 (Converse for the noiseless problem) *From* (60)*, and since* $J^* \to \ell^t \to c^t$*, we see that* $I(J^*; c^t) \leq I(J^*; \ell^t)$*. Following the single-letterization argument in* (61) *and using the arguments leading up to* (62) *we can recover the the converse for the noiseless prediction with experts problem. We note in particular that since* (60) *applies for* any *channel (not necessarily memoryless) the outlined argument establishes a converse for the 1-bit/expert converse.*

4 Discussion

This paper addresses the problem of prediction with expert advice in the presence of noisy feedback. We propose a general achievability framework, analyze its regret (Theorem 1) and show a converse result (Theorem 2) on the regret incurred by any strategy. We then specialize these general results to canonical memoryless noise channels such as the BSC (4), BEC (5), additive noise channels (6)(7)(8), and to the noise induced due to quantization (9), for which the achievability and converse bounds match. The paper characterizes how the noise level affects the regret, and demonstrates that in some cases, it is possible to achieve the same regret as with noiseless feedback.

A natural question that arises is whether we can derive minmax regret for an arbitrary memoryless channel. Doing this would require us to relate the mean square error (as appears in the achievability bound in Theorem 1) to the error in a multiple hypothesis test (as appears in the proof of the converse in Theorem 2). An intriguing question to ask is which figure of merit of the channel characterizes the factor by which the regret increases.

One of the limitations of our analysis is that it assumes that the noise channels are memoryless, i.e., the corrupted feedback at each round is independent of the previous rounds. However, this may not be realistic in some scenarios, where the noise could exhibit temporal correlations or dependencies. For example, if the feedback is transmitted over a wireless channel, the channel state could vary over time and affect the noise level. In such cases, our achievability scheme and our converse result will need to be modified to account for the channel memory. This is an interesting and challenging direction for future work, as it requires developing new techniques for learning from noisy feedback with memory.

Another important direction for future work is to study quantization noise, where the rate is strictly less than 1 bit per expert. This is relevant in applications where the communication bandwidth is very limited or costly. Unlike the 1-bit/dimension randomized rounding presented here, an optimal quantization scheme will likely need to be adaptive rather than memoryless. Such adaptivity introduces memory to our framework, as it will have to exploit the temporal structure of the losses to achieve better compression. It would be interesting to explore the trade-off between the rate and the regret in this setting, and to design efficient and robust algorithms for learning adaptively from quantized feedback as well as from other channels with memory.

Acknowledgements

This work was funded by NSF grants CCF-1751356 and CCF-1956386 and a Caltech Center for the Mathematics of Information fellowship.

A Proof of Lemma 1

Define

$$Z_t := \sum_{j=1}^{m} \exp\big(-\alpha \sum_{i=1}^{t-1} \ell_{ij}\big),$$

the normalizing term in p_t^{EW}, so that $p_{tj}^{\mathrm{EW}} = \exp\left(-\alpha \sum_{i=1}^{t-1} \ell_{ij}\right)/Z_t$. Then, we will consider $\log Z_t$ to be the potential function and bound the difference in the potential function at each step. We have Note that

$$\begin{aligned}
\log Z_{t+1} - \log Z_t &= \log \frac{\sum_{j=1}^{m} \exp\big(-\alpha \sum_{i=1}^{t} \ell_{ij}\big)}{\sum_{j=1}^{m} \exp\big(-\alpha \sum_{i=1}^{t-1} \ell_{ij}\big)} \\
&= \log \frac{\sum_{j=1}^{m} \exp\big(-\alpha \sum_{i=1}^{t-1} \ell_{ij}\big) \exp(-\alpha \ell_{tj})}{\sum_{j=1}^{m} \exp\big(-\alpha \sum_{i=1}^{t-1} \ell_{ij}\big)} \\
&\overset{(a)}{=} \log \left(\sum_{j=1}^{m} p_{tj}^{\mathrm{EW}} \exp(-\alpha \ell_{tj})\right) \\
&\overset{(b)}{\leq} \log \left(\sum_{j=1}^{m} p_{tj}^{\mathrm{EW}} \left(1 - \alpha \ell_{tj} + \alpha^2 \ell_{tj}^2\right)\right) \\
&= \log \left(1 - \alpha \sum_{j=1}^{m} p_{tj}^{\mathrm{EW}} \ell_{tj} + \alpha^2 \sum_{j=1}^{m} p_{tj} \ell_{tj}^2\right) \\
&\overset{(c)}{\leq} -\alpha \sum_{j=1}^{m} p_{tj}^{\mathrm{EW}} \ell_{tj} + \alpha^2 \sum_{j=1}^{m} p_{tj}^{\mathrm{EW}} \ell_{tj}^2 \\
&= -\alpha \langle p_t^{\mathrm{EW}}, \ell_t \rangle + \alpha^2 \sum_{j=1}^{m} p_{tj}^{\mathrm{EW}} \ell_{tj}^2
\end{aligned} \tag{66}$$

where (a) follows by the definition of p^{EW} in (12), (b) follows since $e^x \leq 1 + x + x^2$ for $x \leq 1$ (and $-\alpha \ell_{tj} \leq 1$), and (c) follows since $\log(1+x) \leq x$ for all x. Now, we observe that

$$\begin{aligned}
\log Z_{n+1} &= \log \left(\sum_{j=1}^{m} \exp\big(-\alpha \sum_{t=1}^{n} \ell_{tj}\big)\right) \\
&\geq \max_{j \in [m]} \log \left(\exp\big(-\alpha \sum_{t=1}^{n} \ell_{tj}\big)\right) \\
&= -\alpha \min_{j \in [m]} \sum_{t=1}^{n} \ell_{tj}
\end{aligned} \tag{67}$$

and that $Z_1 = m$. Summing up (66) over all $t \in [n]$, using (67) and rearranging yields the Lemma.

B Achievability for uniform additive noise

We first show that the estimator $\widehat{\ell}_{tj}$ in (20) is unbiased. Note that

$$\begin{aligned}\mathsf{E}[\widehat{\ell}_t] &= \mathsf{E}\Bigg[\left(-\sigma+\frac{1}{2}\right)\mathbb{1}\{-\sigma\le c_{tj}<-\sigma+1\}+\frac{1}{2}\mathbb{1}\{-\sigma+1\le c_{tj}<\sigma\} \\ &\qquad\qquad +\left(\sigma+\frac{1}{2}\right)\mathbb{1}\{\sigma\le c_{tj}\le\sigma+1\}\Bigg] \\ &= \left(-\sigma+\frac{1}{2}\right)\mathbb{P}\left(-\sigma\le c_{tj}<-\sigma+1\right)+\frac{1}{2}\mathbb{P}\left(-\sigma+1\le c_{tj}<\sigma\right) \\ &\qquad\qquad +\left(\sigma+\frac{1}{2}\right)\mathbb{P}\left(\sigma\le c_{tj}<\sigma+1\right) \end{aligned} \tag{68}$$

Since $c_{tj}=\ell_{tj}+Z_{tj}$ is distributed as $\mathrm{Unif}[-\sigma+\ell_{tj},\sigma+\ell_{tj}]$, we have

$$\mathbb{P}\left(-\sigma\le c_{tj}<-\sigma+1\right)=\frac{1-\ell_{tj}}{2\sigma} \tag{69}$$

$$\mathbb{P}\left(-\sigma+1\le c_{tj}<\sigma\right)=\frac{2\sigma-1}{2\sigma} \tag{70}$$

$$\mathbb{P}\left(\sigma\le c_{tj}\le\sigma+1\right)=\frac{\ell_{tj}}{2\sigma}. \tag{71}$$

Substituting (69), (70) and (71) in (68) yields

$$\begin{aligned}\mathsf{E}[\widehat{\ell}_t] &= \frac{(-2\sigma+1)(1-\ell_{tj})}{4\sigma}+\frac{2\sigma-1}{4\sigma}+\frac{(2\sigma+1)\ell_{tj}}{4\sigma} \\ &= \ell_{tj}\end{aligned} \tag{72}$$

The MSE for this estimator satisfies

$$\begin{aligned}\mathsf{E}[\widehat{\ell}_{tj}-\ell_{tj}]^2 &= \mathsf{E}[\widehat{\ell}_{tj}^2]-\ell_{tj}^2 \\ &= \left(-\sigma+\frac{1}{2}\right)^2\mathbb{P}\left(-\sigma\le c_{tj}<-\sigma+1\right)+\frac{1}{4}\mathbb{P}\left(-\sigma+1\le c_{tj}<\sigma\right) \\ &\qquad\qquad +\left(\sigma+\frac{1}{2}\right)^2\mathbb{P}\left(\sigma\le c_{tj}<\sigma+1\right)-\ell_{tj}^2 \\ &\overset{(a)}{=} \frac{(2\sigma-1)^2(1-\ell_{tj})}{8\sigma}+\frac{(2\sigma-1)}{8\sigma}+\frac{(2\sigma+1)^2\ell_{tj}}{8\sigma}-\ell_{tj}^2 \\ &= \frac{\sigma}{2}-\left(\ell_{tj}-\frac{1}{2}\right)^2 \\ &\le \sigma \end{aligned} \tag{73}$$

where (a) uses (69), (70) and (71).

C Proof of Proposition 2

Following the argument in [Marsiglietti and Kostina, 2018], define

$$g(z):=f(0)\exp(-2xf(0)).$$

Then, $f(0) = g(0)$ and $\int_0^\infty (f(z) - g(z))dz = \int_0^\infty f(z)dz - \int_0^\infty g(z)dz = \frac{1}{2} - \frac{1}{2} = 0$. Since $f, g \to 0$ and $z \to \infty$, this implies that the function $f(z) - g(z)$ crosses the origin at least once in $z > 0$. Moreover, any solution of $f(z) - g(z) = 0 \implies f(z) = g(z)$ must satisfy also $\log f(z) - \log g(z) = 0$. Since $z \mapsto \log f(z) - \log g(z)$ is a concave function (by virtue of $f(z)$ being log-concave and $g(z)$ being log-affine), this implies that $\log f(z) - \log g(z)$ crosses the origin at most once in $z > 0$. Therefore, putting the two together implies that $f(z) - g(z) = 0$ occurs exactly at one point in $0 < z < \infty$. Let us call this point t, so that $f(t) = g(t)$. Therefore, for all $z \leq t$, $f(z) \geq g(z)$ and for all $z > t$, $f(z) \leq g(z)$. Putting these two together, we have

$$(f(z) - g(z))(t^2 - z^2) \geq 0$$

which implies that

$$\int_0^\infty z^2 f(z)dz \leq \int_0^\infty z^2 g(z)dz \tag{74}$$

Since $\int_0^\infty z^2 f(z)dz = \frac{\sigma^2}{2}$ and

$$\int_0^\infty z^2 g(z)dz = \int_0^\infty z^2 \exp(-2zf(0))dz = \frac{1}{8f(0)^2} \int_0^\infty z^2 \exp(-z)dz = \frac{1}{4f(0)^2},$$

(74) yields

$$f(0)^2 \leq \frac{1}{2\sigma^2} \tag{75}$$

which leads to the required Proposition.

References

Nicolo Cesa-Bianchi, Yoav Freund, David Haussler, David P Helmbold, Robert E Schapire, and Manfred K Warmuth. How to use expert advice. *Journal of the ACM (JACM)*, 44(3):427–485, 1997a.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.* Cambridge university press, 2006.

Elad Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019.

Nicolo Cesa-Bianchi, Yoav Freund, David Haussler, David P Helmbold, Robert E Schapire, and Manfred K Warmuth. How to use expert advice. *Journal of the ACM (JACM)*, 44(3):427–485, 1997b.

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.

Meir Feder, Neri Merhav, and Michael Gutman. Universal prediction of individual sequences. *IEEE Transactions on Information Theory*, 38(4):1258–1270, 1992.

Neri Merhav and Meir Feder. Universal prediction. *IEEE Transactions on Information Theory*, 44 (6):2124–2147, 1998.

Tsachy Weissman and Neri Merhav. Universal prediction of individual binary sequences in the presence of noise. *IEEE Transactions on Information Theory*, 47(6):2151–2173, 2001.

Tsachy Weissman and Neri Merhav. Universal prediction of random binary sequences in a noisy environment. *The Annals of Applied Probability*, 14(1):54–89, 2004.

Alon Resler and Yishay Mansour. Adversarial online learning with noise. In *International Conference on Machine Learning*, pages 5429–5437. Proceedings of Machine Learning Research, 2019.

Thomas Kailath, Ali H Sayed, and Babak Hassibi. *Linear estimation*. Prentice Hall, 2000.

Sekhar Tatikonda and Sanjoy Mitter. Control under communication constraints. *IEEE Transactions on automatic control*, 49(7):1056–1068, 2004a.

Sekhar Tatikonda and Sanjoy Mitter. Control over noisy channels. *IEEE Transactions on Automatic Control*, 49(7):1196–1201, 2004b.

Sekhar Tatikonda, Anant Sahai, and Sanjoy Mitter. Stochastic linear control over a communication channel. *IEEE Transactions on Automatic Control*, 49(9):1549–1561, 2004.

Victoria Kostina and Babak Hassibi. Rate-cost tradeoffs in control. *IEEE Transactions on Automatic Control*, 64(11):4525–4540, 2019.

Maxim Raginsky, Rebecca M Willett, Corinne Horn, Jorge Silva, and Roummel F Marcia. Sequential anomaly detection in the presence of noise and limited feedback. *IEEE Transactions on Information Theory*, 58(8):5544–5562, 2012.

Jorma Rissanen. Universal coding, information, prediction, and estimation. *IEEE Transactions on Information theory*, 30(4):629–636, 1984.

Qun Xie and Andrew R Barron. Minimax redundancy for the class of memoryless sources. *IEEE Transactions on Information Theory*, 43(2):646–657, 1997.

Alankrita Bhatt and Young-Han Kim. Sequential prediction under log-loss with side information. In *Algorithmic Learning Theory*, pages 340–344. Proceedings of Machine Learning Research, 2021.

Yanina Shkel, Maxim Raginsky, and Sergio Verdú. Sequential prediction with coded side information under logarithmic loss. In *Algorithmic Learning Theory*, pages 753–769. Proceedings of Machine Learning Research, 2018.

Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. Agnostic online learning. In *Conference on Learning Theory*, 2009.

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

Yu-Guan Hsieh, Kimon Antonakopoulos, Volkan Cevher, and Panayotis Mertikopoulos. No-regret learning in games with noisy feedback: Faster rates and adaptivity via learning rate separation. *Advances in Neural Information Processing Systems*, 2022.

Ke Sun, Samir M Perlaza, and Alain Jean-Marie. $2\times$ 2 zero-sum games with commitments and noisy observations. In *IEEE International Symposium on Information Theory*, 2023.

Mehmet A Donmez, Maxim Raginsky, and Andrew C Singer. Online optimization under adversarial perturbations. *IEEE Journal of Selected Topics in Signal Processing*, 10(2):256–269, 2015.

Francesco Pase, Deniz Gündüz, and Michele Zorzi. Rate-constrained remote contextual bandits. *IEEE Journal on Selected Areas in Information Theory*, 2022.

Osama A Hanna, Merve Karakas, Lin F Yang, and Christina Fragouli. Multi-arm bandits over action erasure channels. In *2023 IEEE International Symposium on Information Theory*, pages 1312–1317. IEEE, 2023.

Jayadev Acharya, Clément L Canonne, and Himanshu Tyagi. Inference under information constraints I: Lower bounds from chi-square contraction. *IEEE Transactions on Information Theory*, 66(12):7835–7855, 2020a.

Jayadev Acharya, Clément L Canonne, and Himanshu Tyagi. Inference under information constraints II: Communication constraints and shared randomness. *IEEE Transactions on Information Theory*, 66(12):7856–7877, 2020b.

Jayadev Acharya, Clément L Canonne, Yuhan Liu, Ziteng Sun, and Himanshu Tyagi. Interactive inference under information constraints. *IEEE Transactions on Information Theory*, 68(1):502–516, 2021.

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.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis of online and differentially private learning. *Advances in Neural Information Processing Systems*, 33, 2020.

Alankrita Bhatt, Nika Haghtalab, and Abhishek Shetty. Smoothed analysis of sequential probability assignment. *Advances in Neural Information Processing Systems*, 36, 2023.

Peter Kairouz, H Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Kallista Bonawitz, Zachary Charles, Graham Cormode, Rachel Cummings, et al. Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, 14(1–2):1–210, 2021.

Osama A Hanna, Lin Yang, and Christina Fragouli. Solving multi-arm bandit using a few bits of communication. In *International Conference on Artificial Intelligence and Statistics*. Proceedings of Machine Learning Research, 2022.

Aritra Mitra, Hamed Hassani, and George J Pappas. Linear stochastic bandits over a bit-constrained channel. In *Learning for Dynamics and Control Conference*, pages 1387–1399. Proceeding of Machine Learning Research, 2023.

Prathamesh Mayekar, Jonathan Scarlett, and Vincent YF Tan. Communication-constrained bandits under additive gaussian noise. *International Conference on Machine Learning*, 2023.

Erich L Lehmann and George Casella. *Theory of point estimation.* Springer Science & Business Media, 2006.

Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of IEEE 36th Annual Foundations of Computer Science*, pages 322–331, 1995.

Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.

Haipeng Luo. Adversarial bandits: Theory and algorithms, 2022. Available online at `https://simons.berkeley.edu/sites/default/files/docs/22250/dddp22-bcslides-haipengluo.pdf`.

Nicolo Cesa-Bianchi, Yishay Mansour, and Gilles Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66:321–352, 2007.

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

Adrien Saumard and Jon A Wellner. Log-concavity and strong log-concavity: a review. *Statistics surveys*, 8:45, 2014.

Aleksandrs Slivkins. Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2):1–286, 2019.

Yury Polyanskiy and Yihong Wu. Information theory: From coding to learning. *Book draft*, 2022.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA, 2006. ISBN 0471241954.

Yury Polyanskiy and Yihong Wu. Strong data-processing inequalities for channels and bayesian networks. In *Convexity and Concentration*, pages 211–249. Springer, 2017.

Maxim Raginsky. Strong data processing inequalities and ϕ-sobolev inequalities for discrete channels. *IEEE Transactions on Information Theory*, 62(6):3355–3389, 2016.

Arnaud Marsiglietti and Victoria Kostina. A lower bound on the differential entropy of log-concave random vectors with applications. *Entropy*, 20(3):185, 2018.