
Online Distribution Learning with Local Private Constraints

Jin Sima*
UIUC

Changlong Wu*
Purdue University

Olgica Milenkovic
UIUC

Wojciech Szpankowski
Purdue University

Abstract

We study the problem of online conditional distribution estimation with *unbounded* label sets under local differential privacy. Let \mathcal{F} be a distribution-valued function class with unbounded label set. We aim at estimating an *unknown* function $f \in \mathcal{F}$ in an online fashion so that at time t when the context \mathbf{x}_t is provided we can generate an estimate of $f(\mathbf{x}_t)$ under KL-divergence knowing only a *privatized* version of the true *labels* sampling from $f(\mathbf{x}_t)$. The ultimate objective is to minimize the cumulative KL-risk of a finite horizon T . We show that under $(\epsilon, 0)$ -local differential privacy of the privatized labels, the KL-risk grows as $\tilde{\Theta}(\frac{1}{\epsilon}\sqrt{KT})$ upto poly-logarithmic factors where $K = |\mathcal{F}|$. This is in stark contrast to the $\tilde{\Theta}(\sqrt{T\log K})$ bound demonstrated by Wu et al. (2023a) for *bounded* label sets. As a byproduct, our results recover a nearly tight upper bound for the hypothesis selection problem of Gopi et al. (2020) established only for the *batch* setting.

1 Introduction

Online conditional distribution learning (a.k.a., sequential probability assignments) (Shtarkov, 1987; Xie and Barron, 1997; Merhav and Feder, 1995; Cesa-Bianchi and Lugosi, 2006; Bilodeau et al., 2021) is a fundamental problem that arises in many different application contexts, including universal source coding (Xie and Barron, 1997; Merhav and Feder, 1995; Drmota and Szpankowski, 2004; Szpankowski and Weinberger, 2012), portfolio optimization (Cesa-Bianchi and Lugosi, 2006), and more recently, the interactive decision making (Foster et al., 2021). Let \mathcal{X}

be a context (or features), \mathcal{Y} be a set of labels, and $\mathcal{D}(\mathcal{Y})$ be the set of all probability distributions over \mathcal{Y} . For any distribution-valued class $\mathcal{F} \subset \mathcal{D}(\mathcal{Y})^{\mathcal{X}}$, the problem is formulated as a game between *Nature* and *learner* with the following protocol: at beginning Nature selects $f \in \mathcal{F}$; at each time step t Nature selects a context \mathbf{x}_t ; the learner then predicts $\hat{p}_t \in \mathcal{D}(\mathcal{Y})$ for the distribution of the next label; Nature then generates $p_t = f(\mathbf{x}_t)$, samples $y_t \sim p_t$ and reveals label y_t to the learner. The goal is to minimize the cumulative risk $\sum_{t=1}^T D(f(\mathbf{x}_t), \hat{p}_t)$, for appropriately selected divergence function D .

It can be shown that for *any* finite class \mathcal{F} the cumulative risk is upper bounded by $\log |\mathcal{F}|$ if we take D to be the KL-divergence. There are many other work that extend this basic setup to different formulations, including non-parametric infinite classes, misspecified setting, stochastic feature generation scenarios and computational efficient methods (Rakhlin and Sridharan, 2015; Bhatt and Kim, 2021; Bilodeau et al., 2021, 2020; Wu et al., 2022; Bhatt et al., 2023).

This paper investigates a different angle for this fundamental problem in which data are shared *privately* to the learner. We employ the concept of *local differential privacy* (LDP) for privatizing the *labels* y_t s shared to the learner. Formally, we consider the following game between *three* parties, Nature, learner and the *users*: (i) at beginning Nature selects $f \in \mathcal{F}$; (ii) Nature at each time step t selects \mathbf{x}_t and reveals to learner; learner makes prediction $\hat{p}_t \in \mathcal{D}(\mathcal{Y})$; (iii) the *user* then samples $y_t \sim f(\mathbf{x}_t)$ but reveals a privatized version \tilde{Y}_t to the learner. Our main focus is on minimizing the following KL-risk $\sum_{t=1}^T \text{KL}(f(\mathbf{x}_t), \hat{p}_t)$ subject to local differential privacy constrains on the \tilde{Y}_t s, where KL is the KL-divergence.

1.1 Results and Techniques

Our main results of this paper establish nearly matching lower and upper bounds for the KL-risk of our private online conditional distribution estimation problem when the label set \mathcal{Y} are *unbounded*.

Theorem 1 (Lower Bound). *There exists a finite*

Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain. PMLR: Volume TBD. Copyright 2024 by the author(s). * Equal contribution.

class \mathcal{F} of size K with $|\mathcal{Y}| \leq K$ such that for any $(\epsilon, 0)$ -local differential private mechanism and learning rules, the KL-risk is lower bounded by $\Omega(\frac{1}{\epsilon}\sqrt{KT})$.

A pertinent setup was studied recently by Wu et al. (2023a), where the authors demonstrated that the KL-risk can be upper bounded by $\tilde{O}(\sqrt{T} \log K)$ via a randomized response mechanism. However, it was assumed implicitly that the alphabet size of \mathcal{Y} is a *constant* when defining their noisy kernel in a differentially private manner. Our lower bound demonstrates that such an logarithmic dependency on the size K cannot be achieved for *unbounded* label set \mathcal{Y} .

Our next main result shows that the lower bound is essentially tight for *unbounded* label sets, upto poly-logarithmic factors:

Theorem 2 (Upper Bound). *For any finite class \mathcal{F} of size K with any \mathcal{Y} , there exists $(\epsilon, 0)$ -local private differential private mechanism and learning rules, such that with adversarially generated features, the KL-risk is upper bounded by $O(\frac{1}{\epsilon}\sqrt{KT \log^5(KT)})$.*

Our main technique for achieving this upper bound is based on a modified version of the EXP3 algorithm by appropriately defining the loss vector using the *log-likelihoods* of distributions specified by \mathcal{F} . Unlike the conventional randomized response that perturbs the labels, we add noise directly to the log-likelihood vectors at each time step. Note that the main challenge here is that the log-likelihoods are generally *unbounded*. Moreover, we need the scale of (Laplacian) noise of order K/ϵ to achieving $(\epsilon, 0)$ -differential privacy for a vector of dimension K . To resolve the first issue, we employ a "clipping" technique as Gopi et al. (2020) for controlling the log-likelihoods, but it needs careful adaption that is suited for bounding our KL-risk. To resolve the second issue, we employ a novel randomized approach that reveals only one component of the log-likelihood vector at each time step which reduces the need of noise scale from K/ϵ to $O(1/\epsilon)$. We then apply an adaption of the regret analysis of EXP3 algorithm for deriving our desired KL-risk bound.

Finally, by applying the *online-to-batch* conversion technique and Pinsker's inequality, we show that our upper bound for the KL-risk also implies a nearly tight upper bound for the *batch* setup of Gopi et al. (2020) with *non-interactive* privatization mechanisms.

1.2 Related Work

Our setup is related to the locally differential private hypothesis selection problem, as introduced in Gopi et al. (2020). This can be seen as a *batch* version of our setup with the *constant* function class. It is crucial to emphasize that our online setup poses sub-

stantial technical challenges. This is primarily because the underlying distributions, which we are attempting to learn, are *changing* at every time step and are unknown *a-priori*. Unlike the local privacy studied in this paper, online learning with *central* differential privacy was studied extensively in the literature, see Dwork et al. (2014); Golowich and Livni (2021); Kaplan et al. (2023); Asi et al. (2023). Learning with locally private labels for *batch* case and with different loss functions was discussed in Chaudhuri and Hsu (2011); Ghazi et al. (2021); Wu et al. (2023b); Esfandiari et al. (2022). Our lower bound proof is also related in spirit to those employed in Ullman (2018).

Summary of contributions. Our main contributions can be summarized as follows: (i) we formulate a novel *online* distribution learning problem with *local* differential privacy constraints; (ii) we demonstrate a surprising lower bound for the KL-risk of $\Omega(\sqrt{KT})$, in contrast to the $\tilde{O}(\sqrt{T} \log K)$ upper bound known in literature only for *bounded* label sets; (iii) we present a novel privatization mechanism and learning rule based on EXP3 that nearly matches the lower bound.

2 Problem Setup and Preliminaries

In this paper, we are interested in the *locally* differential private setting where each entry $y_t \in \mathcal{Y}$, $t \in [T]$, is protected by a message $\tilde{y}_t \in \tilde{\mathcal{Y}}$, generated by some randomized mapping $A_t : \mathcal{Y} \rightarrow \tilde{\mathcal{Y}}$ that is (ϵ, δ) -differentially private. Here, we do not impose any constraints on the choices \mathcal{Y} and $\tilde{\mathcal{Y}}$ for the generality of the definition of differential privacy. We specify \mathcal{Y} and $\tilde{\mathcal{Y}}$ in the following section (Section 2.1).

Definition 1. *A privatization scheme (A_1, \dots, A_T) where $A_t : \mathcal{Y} \rightarrow \tilde{\mathcal{Y}}$, $t \in [T]$, is (ϵ, δ) -locally differentially private for some $\epsilon, \delta > 0$ if for any different $y_t, y'_t \in \mathcal{Y}$ and $S_t \subseteq \tilde{\mathcal{Y}}$, $t \in [T]$,*

$$\Pr(A_t(y_t) \in S_t) \leq e^\epsilon \Pr(A_t(y'_t) \in S_t) + \delta. \quad (1)$$

If $\delta = 0$, the privatization scheme (A_1, \dots, A_T) is called pure ϵ -locally differentially private.

In this paper, we consider only the case when A_t are independent for different $t \in [T]$, i.e., each A_t uses only private coins.

2.1 Problem Formulation

Let $\mathcal{X} = \mathbb{R}^d$ be the feature space for some positive integer d and $\mathcal{Y} = [M] = \{1, \dots, M\}$ be the label space for some positive integer M . It is assumed that M and d can take any value. Denote $\mathcal{D}(\mathcal{Y}) = \{(u_1, \dots, u_M) \in \mathbb{R}^M : \sum_{i=1}^M u_i = 1, u_i \geq 0, i \in [M]\}$ as the set of probability distributions over \mathcal{Y} . Let $\mathcal{F} = \{f_1, \dots, f_K\}$

be a hypothesis set where $f_j : \mathcal{X} \rightarrow \mathcal{D}(\mathcal{Y})$, $j \in [K]$ is a function that maps a feature $\mathbf{x} \in \mathcal{X}$ to a distribution $f_j(\mathbf{x}) \in \mathcal{D}(\mathcal{Y})$ over the label space. For any $y \in \mathcal{Y}$, $j \in [K]$, and $\mathbf{x} \in \mathcal{X}$, let $f_j(\mathbf{x})[y]$ be the probability mass of element y in distribution $f_j(\mathbf{x})$.

Consider an online game between *Nature*, *Learner* and the *Users* where Nature arbitrarily picks a function $f \in \mathcal{F}$ at time $t = 0$ and the Learner wishes to learn f over a time period $t \in [T]$ using privatized data generated by the Users. At each time $t \in [T]$, Nature arbitrarily picks a feature $\mathbf{x}_t \in \mathcal{X}$ and reveals \mathbf{x}_t to the Learner. The Learner then makes an estimate $\hat{p}_t = \Phi_t(\mathbf{x}^t = (\mathbf{x}_1, \dots, \mathbf{x}_t), \tilde{\mathbf{Y}}^{t-1} = (\tilde{\mathbf{Y}}_1, \dots, \tilde{\mathbf{Y}}_{t-1}))$ of $f(\mathbf{x}_t)$ based on the feature history \mathbf{x}^t , privatized label history $\tilde{\mathbf{Y}}^{t-1}$, and a function $\Phi_t : \mathcal{X}^t \times \tilde{\mathcal{Y}}^{t-1} \rightarrow \mathcal{D}(\mathcal{Y})$. After the Learner makes the estimate \hat{p}_t , the User then samples a label Y_t independently according to the distribution $f(\mathbf{x}_t)$, generates (independently for different $t \in [T]$) a privatized version $\tilde{\mathbf{Y}}_t = \mathcal{R}_t(Y_t) \in \mathbb{R}^K$ of Y_t using a random mapping $\mathcal{R}_t : \mathcal{Y} \rightarrow \mathbb{R}^K$, and reveals $\tilde{\mathbf{Y}}_t$ to the Learner. It is required that the privatization scheme $\mathcal{R}^T = (\mathcal{R}_1, \dots, \mathcal{R}_T)$ is (ϵ, δ) -locally differentially private. Hence,

$$\Pr(\tilde{\mathbf{Y}}_t \in S | Y_t) \leq e^\epsilon \Pr(\tilde{\mathbf{Y}}_t \in S | Y'_t) + \delta \quad (2)$$

for any $S \subseteq \mathbb{R}^K$, $Y_t, Y'_t \in \mathcal{Y}$, and $t \in [T]$. Here, we choose $\tilde{\mathcal{Y}} = \mathbb{R}^K$ for convenience of the construction of our privatization schemes as introduced in the following sections. In general, $\tilde{\mathcal{Y}}$ can be arbitrarily chosen.

Our problem is to design an (ϵ, δ) -locally differentially private scheme \mathcal{R}^T for the Nature and a prediction scheme Φ^T for the Predictor, given \mathcal{F} , such that the total estimation error, measured by the expected total Kullback-Leibler (KL) distance

$$\mathbb{E}_{\tilde{\mathcal{Y}}^T} \left[\sum_{t=1}^T \text{KL}(f(\mathbf{x}_t), \hat{p}_t = \Phi_t(\mathbf{x}_t, \tilde{\mathbf{Y}}^{t-1})) \right],$$

over the randomness of the privatization output $\tilde{\mathbf{Y}}^T = (\tilde{\mathbf{Y}}_1, \dots, \tilde{\mathbf{Y}}_T) = (\mathcal{R}_1(Y_1), \dots, \mathcal{R}_T(Y_T))$, is minimized under *arbitrary* choices of $f \in \mathcal{F}$ and \mathbf{x}^T by the Nature. The KL distance $\text{KL}(p_1, p_2)$ is defined as

$$\text{KL}(p_1, p_2) = \sum_{y \in \mathcal{Y}} p_1[y] \log \left(\frac{p_1[y]}{p_2[y]} \right)$$

for any distributions $p_1, p_2 \in \mathcal{D}(\mathcal{Y})$. We define such minimized expected total KL distance as the minimax KL-risk:

$$\begin{aligned} & r_T^{\text{KL}}(\mathcal{F}) \\ &= \inf_{\mathcal{R}^T \in \mathcal{L}(\epsilon, \delta), \Phi^T} \sup_{f \in \mathcal{F}, \mathbf{x}^T \in \mathcal{X}^T} \mathbb{E}_{\tilde{\mathcal{Y}}^T} \left[\sum_{t=1}^T \text{KL}(f(\mathbf{x}_t), \hat{p}_t) \right], \end{aligned}$$

where $\mathcal{L}(\epsilon, \delta)$ is the set of all (ϵ, δ) -locally differentially private schemes \mathcal{R}^T satisfying (2).

In addition to minimizing the expected total KL-distance, we are interested in designing a privatization scheme $\mathcal{R}^T \in \mathcal{L}(\epsilon, \delta)$ and a prediction scheme Φ^T , such that the corresponding average total variation

$$\frac{\sum_{t=1}^T |f(\mathbf{x}_t) - \hat{p}_t|_{\text{TV}}}{T},$$

under arbitrary choices of $f \in \mathcal{F}$ and $\mathbf{x}^T \in \mathcal{X}^T$ is upper bounded. Here, the total variation $|p_1 - p_2|_{\text{TV}}$ between distributions $p_1, p_2 \in \mathcal{D}(\mathcal{Y})$ is defined as $\sum_{y \in \mathcal{Y}} \max\{p_1[y] - p_2[y], 0\}$. More specifically, for any privatization scheme $\mathcal{R}^T \in \mathcal{L}(\epsilon, \delta)$ and prediction scheme Φ^T , let

$$\bar{r}_T^{\text{TV}}(\Phi^T, \mathcal{R}^T, \mathcal{F}) = \sup_{f \in \mathcal{F}, \mathbf{x}^T} \mathbb{E}_{\tilde{\mathcal{Y}}^T} \left[\frac{\sum_{t=1}^T |f(\mathbf{x}_t) - \hat{p}_t|_{\text{TV}}}{T} \right]$$

be the worst-case average total variation associated with Φ^T and \mathcal{R}^T .

3 An $\Omega(\sqrt{KT})$ Lower Bound

In this section we demonstrate an $\Omega(\frac{1}{\epsilon}\sqrt{KT})$ lower bound by constructing a *hard* hypothesis class \mathcal{F} of size K with $|\mathcal{Y}| \leq K$. This will demonstrate that an $\tilde{O}(\sqrt{T \log K})$ type upper bound, such as Wu et al. (2023a), is not achievable for *unbounded* label set \mathcal{Y} that may depend on K and T . We then provide nearly matching (upto poly-logarithmic factors) upper bounds for *any* finite class \mathcal{F} with adversarially generated features \mathbf{x}^T in Section 4 and 5.

Before presenting our lower bound, we first demonstrate how the randomized response mechanism, such as Wu et al. (2023a), fails to obtain tight KL-risk bounds for unbounded label set. Let $|\mathcal{Y}| = M$, the randomized response mechanism proceeds as follows: for any $y \in \mathcal{Y}$, we set $\tilde{Y} = y$ w.p. $1 - \eta$ and w.p. η we set \tilde{Y} be uniform in $\mathcal{Y} \setminus \{y\}$. In order to achieve $(\epsilon, 0)$ -differential privacy, one would need to set $\eta = \left(\frac{e^\epsilon}{M-1} + 1 \right)^{-1}$. It was demonstrated by Wu et al. (2023a) that the KL-risk grows as $\tilde{O}(\frac{1}{c_\eta} \sqrt{T \log K})$, where $c_\eta = 1 - \frac{M\eta}{M-1}$. This gives $c_\eta = \frac{e^\epsilon - 1}{e^\epsilon + M - 1} \sim \frac{\epsilon}{M}$ for small ϵ . Therefore, the upper bound grows as $\tilde{O}(\frac{M}{\epsilon} \sqrt{T \log K})$, which will be vacuous for $M \rightarrow \infty$.

We now ready to state our main result of this section:

Theorem 3. *There exists a finite class $\mathcal{F} \subset \mathcal{D}(\mathcal{Y})^{\mathcal{X}}$ of size K with $|\mathcal{Y}| \leq K$, such that for any $(\epsilon, 0)$ -locally private online learning scheme \mathcal{R}^T and Φ^T (depending*

on \mathcal{F}), the KL-risk $r_T^{\text{KL}}(\mathcal{F})$ is lower bounded by

$$\Omega\left(\sqrt{\frac{KT}{\min\{(e^\epsilon - 1)^2, 1\}e^\epsilon}}\right)$$

for all $T \geq \frac{K}{9 \min\{(e^\epsilon - 1)^2, 1\}e^\epsilon}$. Moreover, the bound grows as $\Omega(\frac{1}{\epsilon}\sqrt{KT})$ for sufficiently small ϵ .

Sketch of Proof. At a high level, for any K , our goal is to construct K pairs of distributions $\{p_{k,1}, p_{k,2}\}$ for $k \in [K]$, such that: (i) for any $k \in [K]$, $\inf_{\hat{p}} \sup\{\text{KL}(p_{k,1}, \hat{p}), \text{KL}(p_{k,2}, \hat{p})\} \geq \Omega(a)$ where a is of order $\frac{1}{\epsilon}\sqrt{\frac{K}{T}}$; (ii) $p_{k,1} - p_{k,2} = \frac{a}{N}\mathbf{h}_k$ where each \mathbf{h}_k corresponds to distinct columns of a Hadamard matrix of dimension $K + 1 \leq N \leq 2K$ (excluding the column with all 1s). Assume such a construction exists (see Appendix A), we then construct the class \mathcal{F} of $2K$ constant functions $p_{k,i}$ for $k \in [K]$, $i \in \{1, 2\}$. Our key technical part is to show that, for any $(\epsilon, 0)$ -local differential private mechanism (possibly depending on \mathcal{F}) there exists $k^* \in [K]$ such that $\text{KL}(\tilde{p}_{k^*,1}, \tilde{p}_{k^*,2}) \leq O(\frac{a^2 \epsilon^2}{K}) \leq \frac{c}{T}$ for some $c < 1$, where $\tilde{p}_{k^*,i}$ is the distribution of the private outcomes with input distribution $p_{k^*,i}$. This is achieved by relating the KL-divergence with χ^2 divergence and a careful application of Parseval's identity using the fact that \mathbf{h}_k s are *orthogonal*. However, this implies, by Pinsker's inequality, that $|\tilde{p}_{k^*,1}^{\otimes T} - \tilde{p}_{k^*,2}^{\otimes T}|_{\text{TV}} < 1$ and therefore by Le Cam's two point method no predictor can distinguish $p_{k^*,1}, p_{k^*,2}$ from its privatized samples of length T . Therefore no algorithm could achieve the KL-risk of order $o(\frac{1}{\epsilon}\sqrt{KT})$ due to property (i) of our construction of $p_{k,i}$ s (since any such predictor can be used to distinguish $p_{k^*,1}, p_{k^*,2}$). This will then complete the proof. See Appendix A for a detailed proof. \square

Note that our proof is similar in spirit to those employed in Ullman (2018) (see also Gopi et al. (2020)), e.g., the crucial application of the Parseval's identity. However, a distinguishing feature of our construction is that our label set \mathcal{Y} is of size K , while Ullman (2018) requires a 2^K support size. Moreover, our proof is constructive and employs methodology that is suitable for bounding KL-risk instead of the total variation, which is of independent interest.

4 Approximate-DP via WMA

In this section, we present an online learning scheme (Algorithm 1) that is (ϵ, δ) -locally differentially private and has minimax KL-risk $r_T^{\text{KL}}(\mathcal{F}) = \tilde{O}\left(\frac{1}{\epsilon}\sqrt{TK \log \frac{1}{\delta}}\right)$, where \tilde{O} omits the *poly*($\log(KT)$) factors. Our scheme

makes use of the Laplacian mechanism for local differential privacy and the weighted majority algorithm (WMA) for online learning.

4.1 The weighted majority algorithm

Before presenting our scheme, we first briefly review the well-known weighted majority algorithm for the following online game: Suppose there are K experts $[K]$, accessible by the Predictor. At each time $t \in [T]$, the Predictor picks an expert $i_t \in [K]$ and observes a loss vector $\mathbf{v}_t = (v_{t,1}, \dots, v_{t,K}) \in [0, 1]^K$, which represents the loss of choosing each of the experts. Then, the Predictor experiences loss v_{t,i_t} . Given K and T , the weighted majority algorithm is as follows

Initialization: Let $\mathbf{w}^1 = (w_1^1, \dots, w_K^1) = (1, \dots, 1)$ and $\eta = \sqrt{\frac{2 \log K}{T}}$.

For time $t = 1, \dots, T$

Step 1: For $j \in [K]$, the Predictor select $i_t = j$ with probability $\tilde{w}_j^t = \frac{w_j^t}{\sum_{i=1}^K w_i^t}$

Step 2: The Predictor observes the loss vector \mathbf{v}_t and experiences loss v_{t,i_t}

Step 3: Update $w_j^{t+1} = w_j^t * e^{-\eta v_{t,j}}$ for $j \in [K]$.

The following lemma gives an upper bound on the regret of WMA, defined as the expected loss minus the minimum total loss of a single expert

Lemma 1 (Shalev-Shwartz and Ben-David (2014)). *When $\mathbf{v}_t \in [0, 1]^K$ for $t \in [T]$, the regret of WMA satisfies*

$$\sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t v_{t,j} - \min_{i \in [K]} \sum_{t=1}^T v_{t,i} \leq \sqrt{2T \log K}.$$

4.2 Our scheme

Our key idea of leveraging the WMA algorithm in the context of minimizing KL-risk is to appropriately define the *loss* vector \mathbf{v}_t s. At a high level, we will choose $\mathbf{v}_t[j]$ to be the *log-likelihoods* $-\log f_j(\mathbf{x}_t)[Y_t]$ for $j \in [K]$ and add Laplacian noise to the loss vectors. However, instead of selecting an expert (or a distribution from a set of candidates as in the online distribution learning setting) at a time, we estimate the distribution using a *weighted average*.

We first define random functions for "clipping" the log-likelihood, which is crucial for applying the Laplacian mechanism. The "clipping" technique appeared in Gopi et al. (2020). However, here we use slightly different "clipping functions" for the purpose of bounding the KL divergence.

Clipping of distributions. Let \mathbf{x}^T be any realization of features. For each $t \in [T]$, let

$$N'_t = \sum_{y \in \mathcal{Y}} \lceil M \max_{j \in [K]} f_j(\mathbf{x}_t)[y] \rceil. \quad (3)$$

Let $\{\mathcal{S}_{y,t}\}_{y \in \mathcal{Y}}$, $\mathcal{S}_{y,t} \subset [N'_t]$, be a partition of $[N'_t]$ such that $|\mathcal{S}_{y,t}| = \lceil M \max_{j \in [K]} f_j(\mathbf{x}_t)[y] \rceil$. Define the random mapping $h'_t : \mathcal{Y} \rightarrow [N'_t]$ such that $h'_t(y)$ is uniformly distributed over $\mathcal{S}_{y,t}$, i.e.,

$$\Pr(h'_t(y) = y' | y) = \begin{cases} \frac{1}{\lceil M \max_{j \in [K]} f_j(\mathbf{x}_t)[y] \rceil}, & \text{if } y' \in \mathcal{S}_{y,t} \\ 0, & \text{else.} \end{cases} \quad (4)$$

Let $h'_t \circ f_j(\mathbf{x}_t)$, $j \in [K]$, $t \in [T]$ be the distribution of $h'_t(y)$ when $y \in \mathcal{Y}$ is sampled from distribution $f_j(\mathbf{x}_t)$.

Let also $h_t(y) : \mathcal{Y} \rightarrow [N'_t]$ be $h'_t(y)$ with probability $1 - \frac{1}{T}$ and be $U(y)$ with probability $\frac{1}{T}$, where $U(y)$ is uniformly distributed over $[N'_t]$, i.e.,

$$\Pr(h_t(y) = y' | y) = \begin{cases} \frac{1 - \frac{1}{T}}{\lceil M \max_{j \in [K]} f_j(\mathbf{x}_t)[y] \rceil} + \frac{1}{TN'_t}, & \text{if } y' \in \mathcal{S}_{y,t} \\ \frac{1}{TN'_t}, & \text{else.} \end{cases} \quad (5)$$

Similarly, define $h_t \circ f_j(\mathbf{x}_t)$, $j \in [K]$, $t \in [T]$ as the distribution of $h_t(y)$, $y \in \mathcal{Y}$ when y is drawn according to distribution $f_j(\mathbf{x}_t)$, i.e.,

$$h_t \circ f_j(\mathbf{x}_t)[y'] = \sum_{y \in \mathcal{Y}} \Pr(h_t(y) = y' | y) f_j(\mathbf{x}_t)[y]. \quad (6)$$

By definition of N'_t (3), we have $M \leq N'_t \leq KM$. Therefore,

$$\frac{1}{TKM} \leq (h_t \circ f_j(\mathbf{x}_t))[y'] \leq \frac{1}{M} \quad (7)$$

for any $j \in [K]$ and $y' \in [N'_t]$. The fact that $h_t \circ f_j(\mathbf{x}_t)$ is upper and lower bounded (7) implies that the loglikelihood $-\log(h_t \circ f_j(\mathbf{x}_t))[y']$ has sensitivity¹ $\log(KT)$ and thus can be made $(\epsilon, 0)$ differentially private by adding Laplacian noise with scale² $\frac{\log(KT)}{\epsilon}$ (Dwork et al., 2014).

The privatization scheme. In order to preserve the privacy for the loss vector defined in WMA, we need to add Laplacian noise to a vector instead of a scalar value. The following lemma shows that, by using advanced composition, one only needs to add Laplacian noises with the scale which is *square-root* of the dimension.

¹For a real-valued function $s : \mathcal{Z} \rightarrow \mathbb{R}$, its sensitivity $\Delta_1(s)$ is defined as $\max_{z \in \mathcal{Z}} s(z) - \min_{z \in \mathcal{Z}} s(z)$.

²A Laplacian random variable with scale b has density $\frac{1}{2b} e^{-\frac{|x|}{b}}$ for $x \in \mathbb{R}$.

Lemma 2 (Steinke (2022)). *Let $A_1, \dots, A_K : \mathcal{Y} \rightarrow \mathbb{R}$ be K random algorithms that are $(\epsilon', 0)$ differentially private. Then the composition $A(y) = (A_1(y), \dots, A_K(y)) : \mathcal{Y} \rightarrow \mathbb{R}^K$ where A_1, \dots, A_K run independently, is $(\frac{K(\epsilon')^2}{2} + \sqrt{2 \ln(\frac{1}{\delta}) K (\epsilon')^2}, \delta)$ differentially private for any $\delta > 0$.*

We are now ready to present our privatization scheme. Let \mathbf{x}_t and (random) Y_t be the feature and label at time t , and $\bar{p}_{t,j} = h_t \circ f_j(\mathbf{x}_t)$ for $j \in [K]$. Let $\gamma > 0$ be a number to be determined later, we define the privatized vector $\tilde{\mathbf{Y}}_t = (\tilde{Y}_{t,1}, \dots, \tilde{Y}_{t,K})$ as

$$\tilde{Y}_{t,j} = -c(\gamma)(\log(\bar{p}_{t,j}[h_t(Y_t)]) + \text{Lap}_j^t + \log M - c'(\gamma)), \quad (8)$$

where Lap_j^t , $t \in [T]$, $j \in [K]$ are independent random variables having Laplacian distribution with scale $\frac{(2\sqrt{2K \ln \frac{1}{\delta}} + \sqrt{K\epsilon}) \log(KT)}{\epsilon}$. Moreover, $c(\gamma) = \frac{1}{\log(KT) + 2c'(\gamma)}$ and $c'(\gamma)$ is

$$\frac{(2\sqrt{2K \ln \frac{1}{\delta}} + \sqrt{K\epsilon}) \log(KT)(\gamma + \log K + \log T)}{\epsilon}. \quad (9)$$

Note that, the loss vector $\tilde{\mathbf{Y}}_t = (\tilde{Y}_{t,1}, \dots, \tilde{Y}_{t,K})$, $t \in [T]$ is a privatized version of the loglikelihood vector $(\bar{p}_{t,1}[h_t(Y_t)], \dots, \bar{p}_{t,K}[h_t(Y_t)])$ and thus a privatization of Y_t . The following lemma shows that the privatized data $\tilde{\mathbf{Y}}_t$ satisfies the privacy constraints.

Lemma 3. *The privatization $\mathcal{R}_t(Y_t) = \tilde{\mathbf{Y}}_t$, $t \in [T]$ in Algorithm 1 is (ϵ, δ) -locally differentially private.*

Proof. Note that from (7), the sensitivity of $\log(\bar{p}_{t,j}[h_t(Y_t)])$ is $\log(KT)$ for $j \in [K]$, $t \in [T]$. Hence from (8), $\tilde{Y}_{t,j}$ is $(\epsilon' = \frac{\epsilon}{2\sqrt{K \ln(\frac{1}{\delta})} + \sqrt{K\epsilon}}, 0)$ -locally differentially private. Invoking Lemma 2, $\tilde{\mathbf{Y}}_t = (\tilde{Y}_{t,1}, \dots, \tilde{Y}_{t,K})$ is $(\frac{K(\epsilon')^2}{2} + \sqrt{2 \ln(\frac{1}{\delta}) K (\epsilon')^2}, \delta)$ differentially private, and thus is (ϵ, δ) -differentially private since

$$\frac{K(\epsilon')^2}{2} + \sqrt{2 \ln(\frac{1}{\delta}) K (\epsilon')^2} \leq \epsilon.$$

Hence, \mathcal{R}^T is (ϵ, δ) -locally differentially private. \square

Distribution learning algorithm. We now introduce our private learning algorithm, as presented in Algorithm 1, using our privatization scheme discussed above. Note that the prediction \hat{p}_t at time t may be different from any of the experts $h_t \circ f_j(\mathbf{x}_t)$ and the candidate distributions $f_j(\mathbf{x}_t)$, $j \in [K]$.

We now establish an $\tilde{O}(\sqrt{TK})$ upper bound on the KL-risk of Algorithm 1. The following lemma directly follows from Lemma 1.

Algorithm 1 Locally Privatized WMA

-
- 1: **Input:** The hypothesis set \mathcal{F} , the time horizon T , probability parameter γ
 - 2: **Initialize:** Let $K = |\mathcal{F}|$, $\mathbf{w}^1 = (w_1^1, \dots, w_K^1) = (1, \dots, 1)$, and $\eta = \sqrt{\frac{2K \log K}{T}}$;
 - 3: **for** $t = 1, \dots, T$ **do**
 - 4: Receive feature \mathbf{x}_t ;
 - 5: Set $\bar{p}_t = \sum_{j=1}^K \tilde{w}_j^t \bar{p}_{t,j}$ where $\bar{p}_{t,j} = h_t \circ f_j(\mathbf{x}_t)$ is defined in (6) and $\tilde{w}_j^t = \frac{w_j^t}{\sum_{j=1}^K w_j^t}$;
 - 6: Make prediction $\hat{p}_t[y] = \sum_{y' \in \mathcal{S}_{y,t}} \bar{p}_t[y']$ for all $y \in \mathcal{Y}$, where $\mathcal{S}_{y,t}$ is defined in (4);
 - 7: Receive privatized data $\tilde{\mathbf{Y}}_t = (\tilde{Y}_{t,1}, \dots, \tilde{Y}_{t,K})$ as defined in (8);
 - 8: **Update** $w_j^{t+1} = w_j^t * e^{-\eta \tilde{Y}_{t,j}}$ for $j \in [K]$;
 - 9: **end for**
-

Lemma 4. In Algorithm 1, if $\tilde{Y}_{t,j} \in [0, 1]$ for $j \in [K]$ and $T > \log K$, then

$$\sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \tilde{Y}_{t,j} - \min_{i \in [K]} \left(\sum_{t=1}^T \tilde{Y}_{t,i} \right) \leq \sqrt{2T \log K}.$$

We now introduce the following key lemma, which establishes an upper bound for the KL-risk condition on the event that the Laplacian noises are bounded.

Lemma 5. For any $f \in \mathcal{F}$ and $\mathbf{x}^T \in \mathcal{X}^T$, with probability at least $1 - e^{-\gamma}$ on the randomness of Lap_j^t s, the output \hat{p}_t of Algorithm 1 satisfies for any $\gamma > 0$

$$\begin{aligned} & \mathbb{E}_{Y^T} \left[\sum_{t=1}^T \text{KL}(f(\mathbf{x}_t), \hat{p}_t) \right] \\ & \leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT) \\ & \quad + \sum_{t=1}^T \sum_{j=1}^K \mathbb{E}_{Y^T} [\tilde{w}_j^t] Lap_j^t - \sum_{t=1}^T Lap_j^t, \end{aligned} \quad (10)$$

where $c(\gamma)$ is given in (9) and $Y_t' \sim h_t \circ f(\mathbf{x}_t)$.

Sketch of Proof. We sketch only the high level idea here and refer to Appendix B for a detailed proof. Note that, in order to apply Lemma 4 one must ensure that the loss vector $\tilde{\mathbf{Y}}_t$ to be within $[0, 1]$ for each coordinates. We show that this happens w.p. $\geq 1 - e^{-\gamma}$ by our definition of $\tilde{Y}_{t,j}$ s in (8) and the concentration property of Laplacian distributions. Now, conditioning on such an event, we show by Lemma 4 and definition of $\tilde{\mathbf{Y}}_t$ that for any $j^* \in [K]$ we have $\sum_{t=1}^T -\log(\bar{p}_t[Y_t']) + \log(\bar{p}_{t,j^*}[Y_t']) \leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + \beta$, where β depends only on the Laplacians Lap_j^t and $Y_t' = h_t(Y_t)$. Assuming $Y_t \sim f_{j^*}(\mathbf{x}_t)$ and taking expectation on the randomness of Y_t' s,

we have $\mathbb{E}_{Y^T} \left[\sum_{t=1}^T \text{KL}(\bar{p}_{j^*,t}, \bar{p}) \right] \leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + \beta$. Here, we have used the key observation that $\mathbb{E}_{y \sim p} \log(p[y]/q[y]) = \text{KL}(p, q)$ and the law of total probability. The lemma then follows by a careful analysis that relates $\text{KL}(f_{j^*}(\mathbf{x}_t), \hat{p}_t)$ and $\text{KL}(\bar{p}_{j^*,t}, \bar{p})$ by leveraging the crucial properties for our "clipping functions". See Appendix B for details. \square

We now ready to state our main result of this section:

Theorem 4. For any class \mathcal{F} of size K there exist an (ϵ, δ) -local differential private mechanism that achieves the KL-risk $r_T^{\text{KL}}(\mathcal{F})$ upper bounded by $O\left(\sqrt{2T \log K} \left(\log(KT) + \frac{(\sqrt{K \ln \frac{1}{\delta}} + \sqrt{K\epsilon}) \log^2(KT)}{\epsilon}\right)\right)$

Proof. Let \mathcal{E} be the event that for all j, t , $|Lap_j^t| \leq c'(\gamma)$, which happens w.p. $\geq 1 - e^{-\gamma}$ and implies that Lemma 5 holds, see Appendix B. We have

$$\begin{aligned} r_T^{\text{KL}}(\mathcal{F}) &= \mathbb{E}_{\tilde{\mathbf{Y}}^T} \left[\sum_{t=1}^T \text{KL}(f_{j^*}(\mathbf{x}_t), \hat{p}_t) \right] \\ &= \Pr(\mathcal{E}) \mathbb{E}_{\tilde{\mathbf{Y}}^T} \left[\sum_{t=1}^T \text{KL}(f_{j^*}(\mathbf{x}_t), \hat{p}_t) | \mathcal{E} \right] \\ & \quad + \Pr(\mathcal{E}^c) \mathbb{E}_{\tilde{\mathbf{Y}}^T} \left[\sum_{t=1}^T \text{KL}(f_{j^*}(\mathbf{x}_t), \hat{p}_t) | \mathcal{E}^c \right]. \end{aligned}$$

Denote by (A) and (B) the two terms in the above expression that correspond to \mathcal{E} and \mathcal{E}^c , respectively. By Lemma 5, we know (A) can be upper bounded by

$$\begin{aligned} & \mathbb{E}_{\tilde{\mathbf{Y}}^T} \left[\frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT) \right. \\ & \quad \left. + \sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t Lap_j^t - \sum_{t=1}^T Lap_{j^*}^t | \mathcal{E} \right]. \end{aligned} \quad (11)$$

Note that, conditioning on \mathcal{E} the Laplacians Lap_j^t s are still *i.i.d.* distributed and the \tilde{w}_j^t s are summing to 1, so that (11) vanishes when taking (conditional) expectation. Therefore, (A) is upper bounded by $\frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT)$.

To bound the term (B), we have by (30) and (31) (in Appendix B) that $\text{KL}(f_{j^*}(\mathbf{x}_t), \hat{p}_t)$ is upper bounded by $\text{KL}(\bar{p}_{t,j^*}, \bar{p}_t) + O\left(\frac{\log(KT)}{T}\right)$. Noticing that $\frac{\bar{p}_{t,j^*}[y']}{\bar{p}_t[y']} \leq KT$ for all $y' \in [N_t']$ (see (7)), we conclude $\text{KL}(\bar{p}_{t,j^*}, \bar{p}_t) \leq \log(KT)$. Therefore, by summing over $t \in [T]$, the term (B) is upper bounded by $e^{-\gamma}(T \log(KT) + O(\log(KT)))$, since $\Pr[\mathcal{E}^c] \leq e^{-\gamma}$. Putting everything together, the KL-risk $r_T^{\text{KL}}(\mathcal{F})$ is upper bounded by

$$\frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT) + O(e^{-\gamma} T \log(KT)).$$

The theorem now follows by setting $\gamma = \log T$ and by the definition of $c(\gamma)$ in (9). \square

Observe that the KL-risk bound in Theorem 4 is *independent* of the label set size M and grows as

$$O\left(\frac{1}{\epsilon} \sqrt{TK \log^5(KT) \log \frac{1}{\delta}}\right)$$

for sufficiently small ϵ and δ .

5 Pure-DP via Modified EXP3

While Algorithm 1 attains (ϵ, δ) -local differential privacy, it is worthwhile to investigate whether it is possible to attain the stronger, pure $(\epsilon, 0)$ -differential privacy while still achieving comparable KL-risk bounds. We demonstrate in this section that the answer is affirmative, and it is possible to achieve the same $\tilde{O}\left(\frac{1}{\epsilon} \sqrt{TK}\right)$ KL-risk bound. By Theorem 3, we know that this is essentially tight for pure-DP constrains.

Note that the reason why Theorem 4 has a dependency on δ is due to the *advanced composition* lemma (Lemma 2) that allows us to select the scale of the Laplacians with order of \sqrt{K} , which is essential for achieving a \sqrt{KT} type bound. To resolve this issue, we introduce a new privatization mechanism in this section by selecting a single *random* component in the log-likelihood vectors and reveal only the privatized version of such a component, as in Algorithm 2. Note this reduces significantly the scale of Laplacian from \sqrt{K} to $O(1)$. However, since we only return a single component of the loss vector, the WMA algorithm will not apply. To resolve this issue, we employ an analysis similar to the EXP3 (Exponential-weight for Exploration and Exploitation) algorithm designed for bandit learning via an unbiased estimation of the loss vector. Note that the main difference with the standard EXP3 algorithm is that we *do not* reveal the loss of the expert selected by predictor and instead we reveal the loss for a *randomly* selected expert. This is crucial for making our privatization mechanism *independent* of prior history, i.e., using only private coins.

We now describe our privatization scheme. Let \mathbf{x}_t and Y_t be the feature and label at time t , and $\bar{p}_{t,j} = h_t \circ f_j(\mathbf{x}_t)$ as in (6). Let $\gamma > 0$ to be determined later, we define the vector $\mathbf{Z}_{t,j} = (Z_{t,j,1}, \dots, Z_{t,j,K})$ such that

$$Z_{t,j,i} = \begin{cases} -c_{pdp}(\gamma)(\log(\bar{p}_{t,j}[h_t(Y_t)])) + \\ \text{Lap}_j^t + \log M - c'_{pdp}(\gamma), & \text{if } i = j \\ 0, & \text{else.} \end{cases} \quad (12)$$

where Lap_j^t are independent random variables having Laplacian distribution with scale $\frac{\log(KT)}{\epsilon}$ and h_t is the

random function as in (5). Moreover, the numbers $c_{pdp}(\gamma) = \frac{1}{\log(KT) + 2c'_{pdp}(\gamma)}$ and

$$c'_{pdp}(\gamma) = \frac{\log(KT)(\gamma + \log K + \log T)}{\epsilon}. \quad (13)$$

Let $\tilde{\mathbf{Y}}_t$ be a *random* vector that equals $\mathbf{Z}_{t,j}$ with probability $\frac{1}{K}$ for each $j \in [K]$. We define \tilde{Y}_t as the privatized version of Y_t . The following lemma demonstrates that our scheme is indeed $(\epsilon, 0)$ -differential private.

Lemma 6. *The privatization $\mathcal{R}_t(Y_t) = \tilde{\mathbf{Y}}_t$, $t \in [T]$ in Algorithm 2 is $(\epsilon, 0)$ -locally differentially private.*

Proof. Since for each t , only a single entry of $\tilde{\mathbf{Y}}_t$ is non-zero, and $\log(\bar{p}_{t,j}[h_t(Y_t)])$ has sensitivity $\log(KT)$, the Laplacian mechanism with scale $\frac{\log(KT)}{\epsilon}$ provides $(\epsilon, 0)$ differential privacy (Dwork et al., 2014). \square

We present our learning algorithm in Algorithm 2.

Algorithm 2 Pure Locally Differentially Private

- 1: **Input:** The hypothesis set \mathcal{F} , the time horizon T , probability parameter γ
 - 2: **Initialize:** Let $K = |\mathcal{F}|$, $\mathbf{w}^1 = (w_1^1, \dots, w_K^1) = (1, \dots, 1)$, and $\eta = \sqrt{\frac{2K \log K}{T}}$;
 - 3: **for** $t = 1, \dots, T$ **do**
 - 4: Receive feature \mathbf{x}_t ;
 - 5: Set $\bar{p}_t = \sum_{j=1}^K \tilde{w}_j^t \bar{p}_{t,j}$, where $\tilde{w}_j^t = \frac{w_j^t}{\sum_{j=1}^K w_j^t}$.
 - 6: Make prediction $\hat{p}_t[y] = \sum_{y' \in \mathcal{S}_{y,t}} \bar{p}_t[y']$ for all $y \in \mathcal{Y}$, where $\mathcal{S}_{y,t}$ is defined in (4);
 - 7: Receive privatized data $\tilde{\mathbf{Y}}_t$, where $\tilde{\mathbf{Y}}_t = \mathbf{Z}_{t,J_t}$ as in (12) and J_t is *uniform* over $[K]$;
 - 8: **Update** $w_j^{t+1} = w_j^t * e^{-\eta \tilde{Y}_{t,j}}$ for $j \in [K]$;
 - 9: **end for**
-

We first prove the following key lemma that extends lemma 4 to the *expected* loss vectors.

Lemma 7. *In Algorithm 2, if $\mathbf{Z}_{t,j} \in [0, 1]^K$ for all $t \in [T]$, $j \in [K]$ and $T > \log K$, then*

$$\max_{i \in [K]} \mathbb{E} \left[\sum_{t=1}^T \left(\sum_{j=1}^K \tilde{w}_j^t Z_{t,j,j} - Z_{t,i,i} \right) \right] \leq \sqrt{2TK \log K}$$

where the expectation is over the randomness of J_t s.

Proof. Let $\tilde{\mathbf{Y}}_t = (\tilde{Y}_{t,1}, \dots, \tilde{Y}_{t,K})$ be as in Algorithm 2. We use a more general result of Lemma 1, which is implied in the proof of Shalev-Shwartz and Ben-David (2014, Thm 21.11), that is,

$$\sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \tilde{Y}_{t,j} - \min_{i \in [K]} \left(\sum_{t=1}^T \tilde{Y}_{t,i} \right)$$

$$\leq \frac{\log K}{\eta} + \frac{\eta}{2} \sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \tilde{Y}_{t,j}^2$$

Taking expectations over J_t s, we have

$$\begin{aligned} & \max_{i \in [K]} \mathbb{E} \left[\sum_{t=1}^T \left(\sum_{j=1}^K \tilde{w}_j^t \tilde{Y}_{t,j} - \tilde{Y}_{t,i} \right) \right] \\ & \leq \frac{\log K}{\eta} + \mathbb{E} \left[\frac{\eta}{2} \sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \tilde{Y}_{t,j}^2 \right] \\ & \stackrel{(a)}{\leq} \frac{\log K}{\eta} + \frac{\eta}{2} \sum_{t=1}^T \frac{1}{K} = \frac{\log K}{\eta} + \frac{\eta}{2} \frac{T}{K} \end{aligned} \quad (14)$$

where (a) follows by $\mathbb{E}_{J_t}[\tilde{Y}_{t,j}^2] = \frac{Z_{t,j,j}^2}{K} \leq \frac{1}{K}$ and that \tilde{w}_j^t s are independent of J_t . Taking $\eta = \sqrt{\frac{2K \log K}{T}}$ and noticing that $\mathbb{E}_{J_t}[\tilde{Y}_{t,j}] = \frac{Z_{t,j,j}}{K}$, the lemma follows. \square

The following lemma is analogous to Lemma 5, and we defer the proof to Appendix C.

Lemma 8. *For any $f \in \mathcal{F}$ and $\mathbf{x}^T \in \mathcal{X}^T$, with probability at least $1 - e^{-\gamma}$ on the randomness of Lap_j^t s, the output \hat{p}_t of Algorithm 2 satisfies*

$$\begin{aligned} & \mathbb{E}_{Y^T, J^T} \left[\sum_{t=1}^T \text{KL}(f(\mathbf{x}_t), \hat{p}_t) \right] \\ & \leq \frac{1}{c_{pdp}(\gamma)} \sqrt{2TK \log K} + 3 \log(KT) \\ & \quad + \sum_{t=1}^T \sum_{j=1}^K \mathbb{E}_{Y^T, J^T} [\tilde{w}_j^t] \text{Lap}_j^t - \sum_{t=1}^T \text{Lap}_j^{t*}, \end{aligned} \quad (15)$$

where $c_{pdp}(\gamma)$ is given in (13), $Y_t' \sim h_t \circ f(\mathbf{x}_t)$ and J_t s are the random indices as in Algorithm 2.

We are now ready to state our main result of this section. The proof essentially follows the same steps as the proof of Theorem 4 and is therefore omitted due to space constraints.

Theorem 5. *For any class \mathcal{F} of size K , there exists an $(\epsilon, 0)$ -local differential private mechanism that achieves the following KL-risk r_T^{KL} upper bound*

$$O \left(\sqrt{2TK \log K} \left(\log(KT) + \frac{\log^2(KT)}{\epsilon} \right) \right).$$

6 Bounding Averaged TV-risk

In this section, we show how the KL-risk upper bound for our $(\epsilon, 0)$ -local differential private algorithm (Theorem 5) can be applied to obtain bounds for the averaged TV-risk as introduced in Section 2.1.

The following result follows directly from Pinsker's inequality (Polyanskiy and Wu, 2022).

Theorem 6. *For any class \mathcal{F} of size K , there exists a $(\epsilon, 0)$ -local differential private mechanism that achieves the averaged TV-risk \bar{r}_T^{TV} upper bounded by*

$$\tilde{O} \left(\sqrt{\frac{1}{\epsilon} \sqrt{\frac{K}{T}}} \right),$$

for sufficiently small ϵ .

Proof. Let \hat{p}^T be the predictors that achieve the KL-risk bound of Theorem 5. For any \mathbf{x}^T and $f \in \mathcal{F}$, we have by Pinsker's inequality (Polyanskiy and Wu, 2022) that $|f(\mathbf{x}_t) - \hat{p}_t|_{\text{TV}} \leq \sqrt{\text{KL}(f(\mathbf{x}_t), \hat{p}_t)/2}$ for all $t \in [T]$. Therefore, we have

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T |f(\mathbf{x}_t) - \hat{p}_t|_{\text{TV}} & \leq \frac{1}{T} \sum_{t=1}^T \sqrt{\frac{1}{2} \text{KL}(f(\mathbf{x}_t), \hat{p}_t)} \\ & \leq \frac{1}{T} \sqrt{\frac{T}{2} \sum_{t=1}^T \text{KL}(f(\mathbf{x}_t), \hat{p}_t)} \\ & \leq \tilde{O} \left(\sqrt{\frac{1}{\epsilon} \sqrt{\frac{K}{T}}} \right), \end{aligned}$$

where the second inequality follows by Cauchy-Schwartz inequality and the last inequality follows by Theorem 5. \square

Finally, we find T_α that makes the TV-risk bounded by some α . Let $\mathcal{Q} = \{p_1, \dots, p_K\}$ be K distributions. We construct a class \mathcal{F} of K constant functions with values in \mathcal{Q} . Denote by \hat{p}^T the predictors that achieves the upper bound of Theorem 6 and $\hat{p} = \frac{1}{T} \sum_{t=1}^T \hat{p}_t$. We have by convexity of TV-distance and Jensen's inequality that for any underlying truth $p^* \in \mathcal{Q}$, $|p^* - \hat{p}|_{\text{TV}} \leq \tilde{O} \left(\sqrt{\frac{1}{\epsilon} \sqrt{\frac{K}{T}}} \right)$.

Taking $T_\alpha = \tilde{O} \left(\frac{K}{\epsilon^2 \alpha^2} \right)$ one can make the TV-risk upper bounded by α . This can be boosted to high probability via a median trick, and therefore recovers the nearly tight upper bound in Gopi et al. (2020, Thm 3) with *non-interactive* mechanisms.

Remark 1. *We conjecture that the averaged TV-risk in Theorem 6 may not be tight. We leave it as an open problem to determine if such an upper bound can be improved to $\tilde{O} \left(\frac{1}{\epsilon} \sqrt{\frac{K}{T}} \right)$. Note that, it was demonstrated by Gopi et al. (2020) via suitable comparison schemes that the sample complexity of the batch setting is upper bounded by $\tilde{O} \left(\frac{K}{\epsilon^2 \alpha^2} \right)$ using interactive private mechanisms. It is therefore interesting to investigate if such comparison based arguments can be extended to our online case.*

References

- Asi, H., Feldman, V., Koren, T., and Talwar, K. (2023). Private online prediction from experts: Separations and faster rates. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 674–699. PMLR.
- Bhatt, A., Haghtalab, N., and Shetty, A. (2023). Smoothed analysis of sequential probability assignment. *arXiv preprint arXiv:2303.04845*.
- Bhatt, A. and Kim, Y.-H. (2021). Sequential prediction under log-loss with side information. In *Algorithmic Learning Theory*, pages 340–344. PMLR.
- Bilodeau, B., Foster, D. J., and Roy, D. M. (2021). Minimax rates for conditional density estimation via empirical entropy. *arXiv preprint arXiv:2109.10461*.
- Bilodeau, B., Negrea, J., and Roy, D. M. (2020). Relaxing the iid assumption: Adaptively minimax optimal regret via root-entropic regularization. *arXiv preprint arXiv:2007.06552*.
- Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, Learning and Games*. Cambridge University Press.
- Chaudhuri, K. and Hsu, D. (2011). Sample complexity bounds for differentially private learning. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 155–186. JMLR Workshop and Conference Proceedings.
- Drmota, M. and Szpankowski, W. (2004). Precise minimax redundancy and regrets. *IEEE Trans. Information Theory*, (50):2686–2707.
- Dwork, C., Roth, A., et al. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.
- Esfandiari, H., Mirrokni, V., Syed, U., and Vassilvitskii, S. (2022). Label differential privacy via clustering. In *International Conference on Artificial Intelligence and Statistics*, pages 7055–7075. PMLR.
- Foster, D. J., Kakade, S. M., Qian, J., and Rakhlin, A. (2021). The statistical complexity of interactive decision making. *arXiv preprint arXiv:2112.13487*.
- Ghazi, B., Golowich, N., Kumar, R., Manurangsi, P., and Zhang, C. (2021). Deep learning with label differential privacy. *Advances in Neural Information Processing Systems*, 34:27131–27145.
- Golowich, N. and Livni, R. (2021). Littlestone classes are privately online learnable. *Advances in Neural Information Processing Systems*, 34:11462–11473.
- Gopi, S., Kamath, G., Kulkarni, J., Nikolov, A., Wu, Z. S., and Zhang, H. (2020). Locally private hypothesis selection. In *Conference on Learning Theory*, pages 1785–1816. PMLR.
- Kaplan, H., Mansour, Y., Moran, S., Nissim, K., and Stemmer, U. (2023). On differentially private online predictions. *arXiv preprint arXiv:2302.14099*.
- Merhav, N. and Feder, M. (1995). A strong version of the redundancy-capacity theorem of universal coding. *IEEE Trans. Information Theory*, 41(3):714–722.
- Polyanskiy, Y. and Wu, Y. (2022). *Information Theory: From Coding to Learning*. Cambridge University Press.
- Rakhlin, A. and Sridharan, K. (2015). Sequential probability assignment with binary alphabets and large classes of experts. *arXiv preprint arXiv:1501.07340*.
- Shalev-Shwartz, S. and Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge university press.
- Shtarkov, Y. M. (1987). Universal sequential coding of single messages. *Problems of Information Transmission*, 23(3):3–17.
- Steinke, T. (2022). Composition of differential privacy & privacy amplification by subsampling. *arXiv preprint arXiv:2210.00597*.
- Szpankowski, W. and Weinberger, M. (2012). Minimax pointwise redundancy for memoryless models over large alphabets. *IEEE Trans. Information Theory*, (58):4094–4104.
- Ullman, J. (2018). Tight lower bounds for locally differentially private selection. *arXiv preprint arXiv:1802.02638*.
- Wu, C., Heidari, M., Grama, A., and Szpankowski, W. (2022). Expected worst case regret via stochastic sequential covering. *arXiv preprint arXiv:2209.04417*.
- Wu, C., Wang, Y., Grama, A., and Szpankowski, W. (2023a). Learning functional distributions with private labels. In *International Conference on Machine Learning (ICML)*, pages 37728–37744. PMLR 202.
- Wu, R., Zhou, J. P., Weinberger, K. Q., and Guo, C. (2023b). Does label differential privacy prevent label inference attacks? In *International Conference on Artificial Intelligence and Statistics*, pages 4336–4347. PMLR.
- Xie, Q. and Barron, A. (1997). Minimax redundancy for the class of memoryless sources. *IEEE Trans. Information Theory*, pages 647–657.

Supplementary Material for "Online Distribution Learning with Local Private Constraints"

A Proof of Theorem 3

Let $N = 2^n$ for some positive integer n such that $\frac{N}{2} \leq K \leq N - 1$. Let

$$H^N = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{\otimes n}$$

be a Hadamard matrix and H_i^N be the i th column of H^N . Let $\mathcal{Y} = [N]$ and consider the following $2K$ distributions:

$$p_{i,1}[y] = \begin{cases} 0, & \text{if } H_{i+1,y}^N = 1, \\ \frac{2}{N}, & \text{if } H_{i+1,y}^N = -1 \end{cases}$$

$$p_{i,2}[y] = p_{i,1}[y] + \frac{H_{i+1,y}^N a}{N}, \quad a = \sqrt{\frac{K}{9T \min\{(e^\epsilon - 1)^2, 1\} e^\epsilon}} \quad (16)$$

for $i \in [K]$, $y \in [N]$ Consider learning a hypothesis set $\mathcal{F} = \{f_{1,1}, f_{1,2}, f_{2,1}, \dots, f_{K,1}, f_{K,2}\}$, where $f_{i,\ell}(\mathbf{x}) = p_{i,\ell}$ for all $\mathbf{x} \in \mathcal{X}$, $i \in [K]$, and $\ell \in [2]$. The samples Y^T are i.i.d. random variables generated according to some distribution $p \in \{p_{i,\ell}\}_{i \in [K], \ell \in [2]}$. The goal is to give an estimate of \hat{p} based on locally privatized data $\tilde{\mathbf{Y}}^T$, where $\tilde{\mathbf{Y}}_t = \mathcal{R}_t(Y_t)$, such that the KL distance $\text{KL}(p, \hat{p})$ is minimized. Note that one can use a private online learning predictor Φ^T to make the estimate \hat{p} as follows: set $\hat{p} = \frac{1}{T} \sum_{t=1}^T \hat{p}_t$ where $\hat{p}_t = \Phi_t(\mathbf{x}^t, \tilde{\mathbf{Y}}^{t-1})$. Let f_{j^*, ℓ^*} be the ground truth function. Then

$$\begin{aligned} \mathbb{E}_{\tilde{\mathbf{Y}}^T}[\text{KL}(p, \hat{p})] &\leq \frac{\mathbb{E}_{\tilde{\mathbf{Y}}^T}[\sum_{t=1}^T \text{KL}(f_{j^*, \ell^*}[\mathbf{x}_t], \hat{p}_t)]}{T} \\ &= \frac{r_T^{\text{KL}}(\mathcal{F})}{T}. \end{aligned} \quad (17)$$

In the following, we show that

$$\mathbb{E}_{\tilde{\mathbf{Y}}^T}[\text{KL}(p, \hat{p}(\tilde{\mathbf{Y}}^T))] \geq \frac{a}{33 * 4} = \frac{a}{132} \quad (18)$$

for any estimator $\hat{p} : \tilde{\mathcal{Y}}^T \rightarrow \mathcal{D}([N])$, which proves the theorem.

We use the Le Cam's two point method. Note that for any $\hat{p} \in \mathcal{D}([N])$,

$$\max\{\text{KL}(p_{i,1}, \hat{p}), \text{KL}(p_{i,2}, \hat{p})\} \stackrel{(a)}{\geq} \frac{\text{KL}(p_{i,1}, \frac{p_{i,1} + p_{i,2}}{2})}{2} = \frac{\log(\frac{2}{2 - \frac{a}{2}})}{2} \stackrel{(b)}{\geq} \frac{a}{16}, \quad (19)$$

where (a) follows by Wu et al. (2023a, Lemma C.1) and (b) follows since $0 \leq a \leq 1$ and $\log(1+x) \geq \frac{x}{2}$ when $x \leq 1$. In the following we show that there exists $i \in [K]$ such that

$$|\tilde{p}_{i,1}^{\otimes T} - \tilde{p}_{i,2}^{\otimes T}|_{TV} \leq \frac{1}{3},$$

where $\tilde{p}_{i,\ell}$ is the distribution of $\mathcal{R}_t(Y_t)$ when $Y_t \sim p_{i,\ell}$ for $i \in [K]$ and $\ell \in \{1, 2\}$.

Let \mathcal{R}_t be described by conditional distribution $q_t(\tilde{y}|y)$ for $y \in [N]$ and $\tilde{y} \in \tilde{\mathcal{Y}}$, where $\tilde{\mathcal{Y}}$ is the output space of \mathcal{R}_t . For each $\tilde{y} \in \tilde{\mathcal{Y}}$, let $\mathbf{q}(\tilde{y}) = (q(\tilde{y}|1), \dots, q(\tilde{y}|N))$ be a vector and let

$$\mathbf{q}(\tilde{y}) = \sum_{y=1}^N b(\tilde{y})_y H_y^N \quad (20)$$

where $b(\tilde{y})_y$ are the coefficients associated with the orthogonal base $\{H_y^N\}_{y \in \mathcal{Y}}$. We have

$$\tilde{p}_{i,\ell}(\tilde{y}) = \sum_{y=1}^N q_t(\tilde{y}|y) p_{i,\ell}(y). \quad (21)$$

From (16), (20), and (21), we know that

$$\begin{aligned} \tilde{p}_{i,1}(\tilde{y}) &= \frac{2 \sum_{y: H_{i+1,y}^N = -1} q_t(\tilde{y}|y)}{N}, \text{ and} \\ \tilde{p}_{i,2}(\tilde{y}) &= \frac{2 \sum_{y: H_{i+1,y}^N = -1} q_t(\tilde{y}|y)}{N} + ab(\tilde{y})_{i+1}, \end{aligned}$$

for $i \in [N] \setminus \{1\}$. The KL distance

$$\begin{aligned} \text{KL}(\tilde{p}_{i,1}, \tilde{p}_{i,2}) &= \int_{\tilde{\mathcal{Y}}} -\tilde{p}_{i,1}(\tilde{y}) \log \left(\frac{\tilde{p}_{i,2}(\tilde{y})}{\tilde{p}_{i,1}(\tilde{y})} \right) d\tilde{y} \\ &\stackrel{(a)}{\leq} \int_{\tilde{\mathcal{Y}}} -\tilde{p}_{i,1}(\tilde{y}) \left(\frac{\tilde{p}_{i,2}(\tilde{y}) - \tilde{p}_{i,1}(\tilde{y})}{\tilde{p}_{i,1}(\tilde{y})} - \left(\frac{\tilde{p}_{i,2}(\tilde{y}) - \tilde{p}_{i,1}(\tilde{y})}{\tilde{p}_{i,1}(\tilde{y})} \right)^2 \right) d\tilde{y} \\ &= \int_{\tilde{\mathcal{Y}}} \frac{(\tilde{p}_{i,2}(\tilde{y}) - \tilde{p}_{i,1}(\tilde{y}))^2}{\tilde{p}_{i,1}(\tilde{y})} d\tilde{y} \\ &\leq \int_{\tilde{\mathcal{Y}}} \frac{a^2 b(\tilde{y})_{i+1}^2}{\left(\min_y q_t(\tilde{y}|y) \right)} d\tilde{y}, \end{aligned} \quad (22)$$

where (a) follows from the fact that $x - x^2 \leq \log(1+x)$ for $x \geq 0$.

One the other hand, we have that

$$\begin{aligned} \sum_{y \in [N-1]} b(\tilde{y})_{y+1}^2 &= \sum_{y \in [N-1]} \left(\frac{\sum_{i \in [N]} q_t(\tilde{y}|i) H_{y+1,i}^N}{N} \right)^2 \\ &\stackrel{(a)}{=} \frac{\sum_{i \in [N]} q_t(\tilde{y}|i)^2}{N} - \left(\frac{\sum_{i \in [N]} q_t(\tilde{y}|i)}{N} \right)^2 \\ &\stackrel{(b)}{\leq} \frac{\sum_{i \in [N]} \left(\sum_{j \in [K]} (q_t(\tilde{y}|i) - q_t(\tilde{y}|j))^2 \right)}{2N^2} \\ &\stackrel{(c)}{\leq} \frac{\min\{(e^\epsilon - 1)^2, 1\} \sum_{i \in [N]} q_t(\tilde{y}|i)^2}{2N} \\ &\leq \frac{\min\{(e^\epsilon - 1)^2, 1\} e^\epsilon \sum_{i \in [N]} \min_y q_t(\tilde{y}|y) q_t(\tilde{y}|i)}{2N}, \end{aligned} \quad (23)$$

where (a) follows from the Parseval's identity and the fact that $\{H_y^N\}_{y \in \mathcal{Y}}$ form orthogonal bases, (b) follows from the elementary identity $N \sum_{i=1}^N a_n^2 - (\sum_{i=1}^N a_n)^2 = \frac{1}{2} \sum_{i,j \leq N} (a_i - a_j)^2$, and (c) follows from property of $(\epsilon, 0)$ -differential privacy. Now, (22) and (23) imply

$$\begin{aligned} \sum_{i \in [K]} \text{KL}(\tilde{p}_{i,1}, \tilde{p}_{i,2}) &\leq \int_{\tilde{\mathcal{Y}}} \frac{\sum_{i+1 \in [N]} a^2 b(\tilde{y})_{i+1}^2}{\min_y q_t(\tilde{y}|y)} d\tilde{y} \\ &\leq a^2 \int_{\tilde{\mathcal{Y}}} \frac{\min\{(e^\epsilon - 1)^2, 1\} e^\epsilon \sum_{i \in [N]} q_t(\tilde{y}|i)}{2N} d\tilde{y} \end{aligned}$$

$$\leq \frac{a^2}{2} \min\{(e^\epsilon - 1)^2, 1\}e^\epsilon.$$

Therefore, there exists $i^* \in [K]$ such that

$$\text{KL}(\tilde{p}_{i^*,1}, \tilde{p}_{i^*,2}) \leq \frac{a^2 \min\{(e^\epsilon - 1)^2, 1\}e^\epsilon}{2K},$$

which implies that

$$\begin{aligned} |\tilde{p}_{i^*,1}^{\otimes T} - \tilde{p}_{i^*,2}^{\otimes T}|_{TV} &\leq \sqrt{2\text{KL}(\tilde{p}_{i^*,1}^{\otimes T}, \tilde{p}_{i^*,2}^{\otimes T})} \\ &\leq \sqrt{2T\text{KL}(\tilde{p}_{i^*,1}, \tilde{p}_{i^*,2})} \\ &< \frac{1}{3}. \end{aligned} \quad (24)$$

The rest of the argument is standard. Let p be one of $p_{i,1}$ and $p_{i,2}$ and let $\phi : \mathcal{Y}^T \rightarrow \{p_{i,1}, p_{i,2}\}$ be a classification function deciding if p is $p_{i,1}$ or $p_{i,2}$ based on the privatized data $\tilde{\mathbf{Y}}^T$. From (24) and Le Cam's two point lemma, the classification error

$$\Pr(\phi(\tilde{\mathbf{Y}}^T) \neq p_{i,\ell^*}) \geq \frac{1 - |\tilde{p}_{i^*,1}^{\otimes T} - \tilde{p}_{i^*,2}^{\otimes T}|_{TV}}{2} \geq \frac{1}{3}, \quad (25)$$

where $p = p_{i,\ell^*}$.

On the other hand, suppose there exists a private hypothesis testing estimator \hat{p} satisfying $\text{KL}(p_{i,\ell^*}, \hat{p}(\tilde{\mathbf{Y}}^T)) \leq \frac{a}{33}$ with probability at least $\frac{3}{4}$. Then, from (19) we conclude that a minimum KL distance classifier $\phi(\tilde{\mathbf{Y}}^T) = \text{argmin}_{p^* \in \{p_{i^*,1}, p_{i^*,2}\}} \text{KL}(p^*, \hat{p}(\tilde{\mathbf{Y}}^T))$ is correct with probability at least $\frac{3}{4}$, contradicting (25). Therefore, we have (18), and thus the theorem.

B Proof of Lemma 5

Let \mathcal{E} denote the event that there exists $j \in [K]$, $t \in [T]$ such that

$$|\text{Lap}_j^t| \geq c'(\gamma), \quad (26)$$

where $c'(\gamma)$ is given in (9). Note that for each $t \in [T]$ and $j \in [K]$, (26) occurs with probability $e^{-\gamma - \log K - \log T}$, $t \in [T]$, $j \in [K]$. Hence, by the union bound, the probability of \mathcal{E} is at most $e^{-\gamma}$.

In the following, we show that Lemma 5 holds when \mathcal{E} does not occur, which has probability at least $1 - e^{-\gamma}$.

Note that from (7), we have $\log(\bar{p}_{t,j}[h_t(Y_t)]) \in [\log(\frac{1}{TKM}), \log(\frac{1}{M})]$. Hence,

$$\log(\bar{p}_{t,j}[h_t(Y_t)]) + \text{Lap}_j^t + \log M - c'(\gamma) \in [-\log(KT) - 2c'(\gamma), 0],$$

and thus that $\tilde{Y}_{t,j} \in [0, 1]$. According to Lemma 4, we have

$$\sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \left(-\log(\bar{p}_{t,j}[h_t(Y_t)]) - \text{Lap}_j^t \right) - \min_{i \in [K]} \left(\sum_{t=1}^T \left(-\log(\bar{p}_{t,i}[h_t(Y_t)]) - \text{Lap}_i^t \right) \right) \leq \frac{1}{c(\gamma)} \sqrt{2T \log K}.$$

Let $j^* \in [K]$ be any index, and assume that $Y_t \sim f_{j^*}(\mathbf{x}_t)$. Taking expectation over the distribution of $Y_t' \sim h_t(Y_t)$ we find that

$$\mathbb{E}_{Y^T} \left[\sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \text{KL}(\bar{p}_{t,j^*}, \bar{p}_{t,j}) \right] \leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + \sum_{t=1}^T \sum_{j=1}^K \mathbb{E}_{Y^T} [\tilde{w}_j^t] \text{Lap}_j^t - \sum_{t=1}^T \text{Lap}_{j^*}^t, \quad (27)$$

where we used the fact that Y_t' distributed as $\bar{p}_{t,j^*} = h_t \circ f_{j^*}(\mathbf{x}_t)$ and $\mathbb{E}_{Y \sim p} \log\left(\frac{p[Y]}{q[Y]}\right) = \text{KL}(p, q)$. Moreover, we have by Jensen's inequality and convexity of $-\log(x)$ that

$$\sum_{t=1}^T \left(-\log \left(\sum_{j=1}^K \tilde{w}_j^t \bar{p}_{t,j} [h_t(Y_t)] \right) \right) - \sum_{t=1}^T \left(-\log(\bar{p}_{t,j^*} [h_t(Y_t)]) \right)$$

$$\leq \sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \left(-\log(\bar{p}_{t,j}[h_t(Y_t)]) \right) - \sum_{t=1}^T \left(-\log(\bar{p}_{t,j^*}[h_t(Y_t)]) \right).$$

Hence, by taking expectation over the distribution of $Y'_t \sim h_t(Y_t)$

$$\mathbb{E}_{Y'^T} \left[\sum_{t=1}^T \text{KL}(\bar{p}_{t,j^*}, \bar{p}_t) \right] \leq \mathbb{E}_{Y'^T} \left[\sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \text{KL}(\bar{p}_{t,j^*}, \bar{p}_{t,j}) \right]. \quad (28)$$

Combining (27) and (28), we arrive at

$$\mathbb{E}_{Y'^T} \left[\sum_{t=1}^T \text{KL}(\bar{p}_{t,j^*}, \bar{p}_t) \right] \leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + \sum_{t=1}^T \sum_{j=1}^K \mathbb{E}_{Y'^T} [\tilde{w}_j^t] \text{Lap}_j^t - \sum_{t=1}^T \text{Lap}_{j^*}^t. \quad (29)$$

On the other hand,

$$\begin{aligned} \text{KL}(f_{j^*}(\mathbf{x}_t), \hat{p}_t) &= \sum_{y \in \mathcal{Y}} f_{j^*}(\mathbf{x}_t)[y] \log \left(\frac{f_{j^*}(\mathbf{x}_t)[y]}{\hat{p}_t[y]} \right) \\ &\stackrel{(a)}{=} \sum_{y \in \mathcal{Y}} \left(\sum_{y' \in \mathcal{S}_{y,t}} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \right) \cdot \log \left(\frac{\sum_{y' \in \mathcal{S}_{y,t}} h'_t \circ f_{j^*}(\mathbf{x}_t)[y']}{\sum_{y' \in \mathcal{S}_{y,t}} \left(\sum_{j \in [K]} \tilde{w}_j^t h_t \circ f_j(\mathbf{x}_t)[y'] \right)} \right) \\ &\stackrel{(b)}{=} \sum_{y \in \mathcal{Y}} \sum_{y' \in \mathcal{S}_{y,t}} \left(h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \cdot \log \left(\frac{h'_t \circ f_{j^*}(\mathbf{x}_t)[y']}{\sum_{j \in [K]} \tilde{w}_j^t h_t \circ f_j(\mathbf{x}_t)[y']} \right) \right) \\ &= \text{KL}(h'_t \circ f_{j^*}(\mathbf{x}_t), \bar{p}_t), \end{aligned} \quad (30)$$

where (a) follows from the definition of h'_t and the fact that $\hat{p}_t[y] = \sum_{y' \in \mathcal{S}_y} \bar{p}_t[y']$ where $\bar{p}_t = \sum_{j \in [K]} \tilde{w}_j^t h_t \circ f_j(\mathbf{x}_t)$ (see Algorithm 1), and (b) follows from the fact that $h'_t \circ f_{j^*}(\mathbf{x}_t)[y']$ and $h_t \circ f_{j^*}(\mathbf{x}_t)[y']$ are *constants* for all $y' \in \mathcal{S}_{y,t}$. Moreover,

$$\begin{aligned} &\text{KL}(h'_t \circ f_{j^*}(\mathbf{x}_t), \bar{p}_t) \\ &= \sum_{y' \in [N'_t]} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \log \left(h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \right) - \sum_{y' \in [N'_t]} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \log \left(\bar{p}_t[y'] \right) \\ &= \sum_{y' \in [N'_t]} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \log \left(\left(1 - \frac{1}{T} \right) h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \right) - \sum_{y' \in [N'_t]} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \log \left(\bar{p}_t[y'] \right) - \log \left(1 - \frac{1}{T} \right) \\ &\leq \sum_{y' \in [N'_t]} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \cdot \log \left(\left(1 - \frac{1}{T} \right) h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] + \frac{1}{TN'_t} \right) - \sum_{y' \in [N'_t]} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] \log \bar{p}_t[y'] - \log \left(1 - \frac{1}{T} \right) \\ &= \sum_{y' \in [N'_t]} \left(\left(1 - \frac{1}{T} \right) h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] + \frac{1}{TN'_t} \right) \cdot \log \left(\frac{\left(1 - \frac{1}{T} \right) h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] + \frac{1}{TN'_t}}{\bar{p}_t[y']} \right) \\ &\quad + \sum_{y' \in [N'_t]} \left(\frac{1}{T} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] - \frac{1}{TN'_t} \right) \cdot \log \left(\frac{\left(1 - \frac{1}{T} \right) h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] + \frac{1}{TN'_t}}{\bar{p}_t[y']} \right) - \log \left(1 - \frac{1}{T} \right) \\ &= \text{KL}(\bar{p}_{t,j^*}, \bar{p}_t) + \sum_{y' \in [N'_t]} \left(\frac{1}{T} h'_t \circ f_{j^*}(\mathbf{x}_t)[y'] - \frac{1}{TN'_t} \right) \cdot \log \left(\frac{\bar{p}_{j^*}[y']}{\bar{p}_t[y']} \right) - \log \left(1 - \frac{1}{T} \right) \\ &\stackrel{(a)}{\leq} \text{KL}(\bar{p}_{t,j^*}, \bar{p}_t) + \frac{\log(KT)}{T} + \frac{\log(KT)}{KTM} - \log \left(1 - \frac{1}{T} \right) \end{aligned} \quad (31)$$

where (a) follows because of (7) and the fact that \bar{p}_t is a linear combination of \bar{p}_j , $j \in [K]$. Combining (29),

(30), and (31), we have

$$\begin{aligned}
 \mathbb{E}_{Y^{\prime T}} \left[\sum_{t=1}^T \text{KL}(f_{j^*}(\mathbf{x}_t), \hat{p}_t) \right] &= \mathbb{E}_{Y^{\prime T}} \left[\sum_{t=1}^T \text{KL}(h'_t \circ f_{j^*}(\mathbf{x}_t), \bar{p}_t) \right] \\
 &\leq \mathbb{E}_{Y^{\prime T}} \left[\sum_{t=1}^T \text{KL}(\bar{p}_{t,j^*}, \bar{p}_t) - T \log\left(1 - \frac{1}{T}\right) + 2 \log(KT) \right] \\
 &\leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT) + \sum_{t=1}^T \sum_{j=1}^K \mathbb{E}_{Y^{\prime T}} [\tilde{w}_j^t] \text{Lap}_j^t - \sum_{t=1}^T \text{Lap}_{j^*}^t.
 \end{aligned}$$

This completes the proof.

C Proof of Lemma 8

The proof follows the footsteps of the proof of Lemma 5 by using Lemma 7. Let \mathcal{E} denote the event that there exists $j \in [K]$, $t \in [T]$ such that

$$|\text{Lap}_j^t| \geq c'_{\text{dpr}}(\gamma). \quad (32)$$

Then, the probability of \mathcal{E} is at most $e^{-\gamma}$. We show that Lemma 8 holds when \mathcal{E} does not occur. Note that from (7), we have $\log(\bar{p}_{t,j}[h_t(Y_t)]) \in [\log(\frac{1}{TKM}), \log(\frac{1}{M})]$. Therefore,

$$\log(\bar{p}_{t,j}[h_t(Y_t)]) + \text{Lap}_j^t + \log M - c'(\gamma) \in [-\log(KT) - 2c'(\gamma), 0],$$

and thus that $\mathbf{Z}_{t,j} \in [0, 1]^K$ for all $t \in [T]$ and $j \in [K]$.

Invoking Lemma 7, we have

$$\begin{aligned}
 \max_{i \in [K]} \mathbb{E}_{J^T} \left[\sum_{t=1}^T \left(\sum_{j=1}^K \tilde{w}_j^t Z_{t,j,j} - \sum_{t=1}^T Z_{t,i,i} \right) \right] &= \max_{i \in [K]} \mathbb{E}_{J^T} \left[\sum_{t=1}^T \left(\sum_{j=1}^K \tilde{w}_j^t (-c_{\text{dpr}}(\gamma)(\log(\bar{p}_{t,j}[h_t(Y_t)]) + \text{Lap}_j^t)) \right. \right. \\
 &\quad \left. \left. - c_{\text{dpr}}(\gamma)(\log(\bar{p}_{t,j}[h_t(Y_t)]) + \text{Lap}_j^t) \right) \right] \\
 &\leq \sqrt{2TK \log K}. \quad (33)
 \end{aligned}$$

Let $j^* \in [K]$ be any index and assume $Y_t \sim f_{j^*}(\mathbf{x}_t)$. Taking expectations over the distributions of $Y'_t \sim h_t \circ f_{j^*}(\mathbf{x}_t)$ for $t \in [T]$, we find (similar to (27))

$$\mathbb{E}_{Y^{\prime T}, J^T} \left[\sum_{t=1}^T \sum_{j=1}^K \tilde{w}_j^t \text{KL}(\bar{p}_{t,j^*}, \bar{p}_{t,j}) \right] \leq \frac{1}{c_{\text{dpr}}(\gamma)} \sqrt{2TK \log K} + \sum_{t=1}^T \sum_{j=1}^K \mathbb{E}_{Y^{\prime T}, J^T} [\tilde{w}_j^t] \text{Lap}_j^t - \sum_{t=1}^T \text{Lap}_{j^*}^t. \quad (34)$$

Combining (28), (30), (31), and (34), we conclude

$$\mathbb{E}_{Y^{\prime T}, J^T} \left[\sum_{t=1}^T \text{KL}(f_{j^*}(\mathbf{x}_t), \hat{p}_t) \right] \leq \frac{1}{c_{\text{dpr}}(\gamma)} \sqrt{2TK \log K} + 3 \log(KT) + \sum_{t=1}^T \sum_{j=1}^K \mathbb{E}_{Y^{\prime T}, J^T} [\tilde{w}_j^t] \text{Lap}_j^t - \sum_{t=1}^T \text{Lap}_{j^*}^t.$$

This completes the proof.