Online Distribution Learning with Local Privacy Constraints

Jin Sima\textsuperscript{*}  
UIUC

Changlong Wu\textsuperscript{*}  
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. The problem may be succinctly stated as follows. Let $\mathcal{F}$ be a distribution-valued function class with an unbounded label set. Our aim is to estimate an unknown function $f \in \mathcal{F}$ in an online fashion. More precisely, at time $t$, given a sample $x_t$, we generate an estimate of $f(x_t)$ using only a privatized version of the true labels from $f(x_t)$. The objective is to minimize the cumulative KL-risk of a finite class $\mathcal{F}$.

1 INTRODUCTION

Online conditional distribution learning (a.k.a., sequential probability assignment) [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 domains, 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, interactive decision making [Foster et al. 2021]. Formally, let $\mathcal{X}$ be a set of contexts (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 $\mathcal{F} \subset \mathcal{D}(\mathcal{Y})^X$, the conditional distribution learning problem is formulated as a game between Nature and a Learner that uses the following protocol. At the start of the game, Nature selects some $f \in \mathcal{F}$; then, at each time step $t$, Nature selects a context $x_t$; the Learner predicts $\hat{p}_t \in \mathcal{D}(\mathcal{Y})$ for the distribution of the next label; Nature subsequently generates $p_t = f(x_t)$, samples $y_t \sim p_t$ and reveals the label $y_t$ to the Learner. The goal is to minimize the cumulative risk $\sum_{t=1}^T \mathcal{D}(f(x_t), \hat{p}_t)$, for an appropriately selected divergence function $\mathcal{D}$.

It can be shown that for any finite class $\mathcal{F}$, the cumulative risk is upper bounded by $\log |\mathcal{F}|$ if we take $\mathcal{D}$ to be the KL-divergence. There are many other work that extend this basic setup to account for 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].

Our work investigates a different angle for this fundamental problem in which data are revealed privately to the Learner. We employ the concept of local differential privacy (LDP) for privatizing the labels $y_t$ shared with the learner. Formally, we consider the following game between three parties, Nature, Learner and the Users: (i) at the start of the game, Nature selects $f \in \mathcal{F}$; (ii) Nature at each time step $t$ selects $x_t$ and reveals it to the Learner; the Learner makes a prediction $\hat{p}_t \in \mathcal{D}(\mathcal{Y})$; (iii) the User then samples $y_t \sim f(x_t)$ but reveals a privatized version $Y_t$ to the Learner. The main goal is to minimize the following KL-risk, $\sum_{t=1}^T \mathcal{KL}(f(x_t), \hat{p}_t)$, subject to local differential privacy constrains on the labels $Y_t$.

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 for the case that the label set $\mathcal{Y}$ is unbounded.
That is, we aim to achieve a KL-risk that is independent of the label size $|\mathcal{Y}|$.

**Theorem 1** (Lower Bound). For any $K \geq 2$ and label set $\mathcal{Y}$ with $|\mathcal{Y}| \geq K$, there exists a finite class $\mathcal{F} \subset \mathcal{D}(\mathcal{Y})^X$ of size $K$ such that for any $(\epsilon, 0)$-local differential private mechanism and any learning rule, the KL-risk is lower bounded by $\Omega(\frac{1}{\epsilon} \sqrt{KT})$.

A related setup was recently studied in Wu et al. (2023a), where the authors demonstrated that the KL-risk can be upper bounded by $O(\frac{M}{\epsilon} \sqrt{T \log K})$ via a randomized response mechanism. However, when defining their noisy kernel in a differentially private manner, the risk upper bound grows as $O(\frac{M}{\epsilon} \sqrt{T \log K})$, where $M = |\mathcal{Y}|$. This provides only vacuous bounds when $M \rightarrow \infty$. In fact, our lower bound, based on a novel Hadamard-matrix technique, demonstrates that an $\Omega(\frac{1}{\epsilon} \sqrt{KT})$ minimax lower bound is necessary for $M \geq K$. Therefore, a logarithmic dependency of the KL-risk on the size $|\mathcal{F}|$ as in Wu et al. (2023a) is not achievable for unbounded label sets.

The next result shows that the $\Omega(\frac{1}{\epsilon} \sqrt{KT})$ lower bound is essentially tight for unbounded label sets (whenever $|\mathcal{Y}| \geq K$), up to a poly-logarithmic factor.

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

Our proof technique for establishing this upper bound makes use of 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 to scale the (Laplace) noise to the order of $K/\epsilon$ so as to achieve $(\epsilon, 0)$-differential privacy for a vector of dimension $K$.

To resolve the first issue, we employ a "clipping" technique, such as Gopi et al. (2020), for controlling the log-likelihoods; still, the existing clipping approach requires careful adaption that is suitable 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, since this reduces the noise from $K/\epsilon$ to $O(1/\epsilon)$. We then apply an adaption of the regret analysis for the EXP3 algorithm to derive 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 setting of Gopi et al. (2020) with non-interactive privatization mechanisms.

### 1.2 Related Works

Our problem is related to the local differentially private hypothesis selection problem introduced in Gopi et al. (2020). The referenced result may be seen as a batch version of our setup with a class of constant functions. It is crucial to point out that analyzing our online mechanism poses substantial 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 the batch case and with different loss functions was also 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 to some approaches used in Edmonds et al. (2020).

### 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 establish a (surprising) lower bound for the KL-risk of the form $\Omega(\sqrt{KT})$, in contrast to the $O(\sqrt{T \log K})$ upper bound known 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

We are concerned with the locally differentially private setting where each entry $y_t \in \mathcal{Y}$, $t \in [T]$, is protected by a message $\tilde{y}_t \in \mathcal{Y}$, generated by some randomized mapping $A_t : \mathcal{Y} \rightarrow \tilde{\mathcal{Y}}$ that is $(\epsilon, \delta)$-differentially private. Here, we do not impose any constraints on the choices of $\mathcal{Y}$ and $\tilde{\mathcal{Y}}$. We will specify the sets $\mathcal{Y}$ and $\tilde{\mathcal{Y}}$ in the following section (Section 2.1).

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

If $\delta = 0$, the privatization scheme $(A_1, \ldots, A_T)$ is called purely $\epsilon$-locally differentially private.
We consider only the case when the $A_t$ components are independent for different $t \in [T]$, i.e., each $A_t$ uses only private coins.

### 2.1 Problem Formulation

For some positive integer $d$, let $\mathcal{X} = \mathbb{R}^d$ be the feature space and let $\mathcal{Y} = [M] = \{1, \ldots, M\}$ be the label space, again for some positive integer $M$. It is assumed that $M$ and $d$ can take arbitrary values. Next, we use $\mathcal{D}(\mathcal{Y}) = \{(u_1, \ldots, u_M) \in \mathbb{R}^M : \sum_{i=1}^M u_i = 1, u_i \geq 0, i \in [M]\}$ to denote the set of probability distributions over $\mathcal{Y}$. Let $\mathcal{F} = \{f_1, \ldots, f_K\}$ be a hypothesis set where $f_j : \mathcal{X} \rightarrow \mathcal{D}(\mathcal{Y})$, $j \in [K]$ is a function that maps a feature $x \in \mathcal{X}$ to a distribution $f_j(x) \in \mathcal{D}(\mathcal{Y})$ over the label space. For any $y \in \mathcal{Y}$, $f \in [K]$, and $x \in \mathcal{X}$, let $f_j(x)[y]$ be the probability mass of element $y$ for the distribution $f_j(x)$.

Consider an online game between Nature, Learner and 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 $x_t \in \mathcal{X}$ and reveals $x_t$ to the Learner. The Learner then makes an estimate $\hat{p}_t = \Phi_t(x_t) = (\hat{p}_t^1, \ldots, \hat{p}_t^K)$ based on the feature history $x^t$, privatized label history $\hat{Y}^{t-1}$, and a function $\Phi_t : \mathcal{X} \times \hat{Y}^{t-1} \rightarrow \mathcal{D}(\mathcal{Y})$. After the Learner makes the estimate $\hat{p}_t$, the User then samples a label $Y_t$ independently according to the distribution $f(x_t)$, generates (independently for different $t \in [T]$) a privatized version $Y_t = R_t(Y_t) \in \mathbb{R}^K$ of $Y_t$ using a random mapping $R_t : \mathcal{Y} \rightarrow \mathbb{R}^K$, and reveals $Y_t$ to the Learner. The privatization scheme $\mathcal{R}^T = (R_1, \ldots, R_T)$ is required to be $(\epsilon, \delta)$-locally differentially private. Therefore, we must have

$$\Pr(Y_t \in S|Y'_t) \leq e^\epsilon \Pr(Y_t \in S|Y'_t) + \delta,$$

for any $S \subseteq \mathbb{R}^K$, $Y_t, Y'_t \in \mathcal{Y}$, and $t \in [T]$. Here, we choose $\mathcal{Y} = \mathbb{R}^K$ for convenience, as it makes our construction of the privatization schemes more transparent. In general, $\mathcal{Y}$ can be arbitrary.

Our goal is to design an $(\epsilon, \delta)$-locally differentially private scheme $\mathcal{R}^T$ for the Users and a prediction scheme $\Phi^T$ for the Learner, given $\mathcal{F}$, such that the total estimation error, measured by the expected total Kullback-Leibler (KL) distance

$$E_{\mathcal{Y}^T} \left[ \sum_{t=1}^T \KL(f(x_t), \hat{p}_t | \Phi_t(x_t, \hat{Y}^{t-1})) \right],$$

over the randomness of the privatization output $\hat{Y}^T = (\hat{Y}_1, \ldots, \hat{Y}_T) = (R_1(Y_1), \ldots, R_T(Y_T))$, is minimized under arbitrary choices for $f \in \mathcal{F}$ and $x^T$. By Nature.

The KL distance $\KL(p_1, p_2)$ is defined as

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

for any two distributions $p_1, p_2 \in \mathcal{D}(\mathcal{Y})$ with the same support. We refer to a minimized expected total KL distance as the minimax KL-risk,

$$r^{KL}_T(\mathcal{F}) = \inf_{\mathcal{R}^T \in L(\epsilon, \delta)} \sup_{f \in \mathcal{F}, \mathcal{X}^T} E_{\mathcal{X}^T} \left[ \sum_{t=1}^T \KL(f(x_t), \hat{p}_t) \right],$$

where $L(\epsilon, \delta)$ is the set of all $(\epsilon, \delta)$-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 $\mathcal{R}^T \in L(\epsilon, \delta)$ and a prediction scheme $\Phi^T$, such that the corresponding average total variation

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

under arbitrary choices of $f \in \mathcal{F}$ and $x^T \in \mathcal{X}^T$ is minimized. Here, the total variation $|p_1 - p_2|_{TV}$ between two distributions $p_1, p_2 \in \mathcal{Y}$ is defined as $\sum_{y \in \mathcal{Y}} |p_1(y) - p_2(y)|$. More specifically, for any privatization scheme $\mathcal{R}^T \in L(\epsilon, \delta)$ and prediction scheme $\Phi^T$, we let

$$r^{TV}_T(\Phi^T, \mathcal{R}^T, \mathcal{F}) = \sup_{f \in \mathcal{F}, x^T} E_{\mathcal{X}^T} \left[ \frac{\sum_{t=1}^T |f(x_t) - \hat{p}_t|_{TV}}{T} \right]$$

stand for the worst-case average total variation associated with $\Phi^T$ and $\mathcal{R}^T$.

### 3 THE $\Omega(\sqrt{KT})$ LOWER BOUND

We demonstrate next an $\Omega(\frac{1}{\sqrt{KT}})$ lower bound by constructing a hard hypothesis class $\mathcal{F}$ of size $K$ with $|\mathcal{Y}| \leq K$. This result will establish that an upper bound of the form $O(\sqrt{T \log K})$, such as the one derived in [Wu et al. 2023a], is not achievable for unbounded label sets $\mathcal{Y}$. We then provide a nearly matching (up to poly-logarithmic factors) upper bound for any finite class $\mathcal{F}$ with adversarially generated features $x^T$. These results are relegated to Sections 4 and 5.

Before presenting our lower bound, we first demonstrate how randomized response mechanisms, such as the one in [Wu et al. 2023a], fail to achieve tight KL-risk bounds for unbounded label sets. Let $|\mathcal{Y}| = M$, the randomized response mechanism, operate as follows. For any $y \in \mathcal{Y}$, we set $Y = y$ w.p. $1 - \eta$ and let $Y$ be uniform in $\mathcal{Y}\{y\}$ w.p. $\eta$. In order to
achieve \((\epsilon, 0)\)-differential privacy, one would have to set \(\eta = \left(\frac{e^\epsilon - 1}{M - 1}\right)^{-1}\). It was demonstrated in \cite{Wu2023} that the KL-risk grows as \(\tilde{O}(\frac{1}{M} \sqrt{T \log K})\), where \(c_0 = 1 - \frac{M - 1}{M}\). This gives \(c_0 = \frac{e^\epsilon - 1}{e^\epsilon + M - 1} \sim \frac{e^\epsilon}{M}\), for a small enough \(\epsilon\). Therefore, the upper bound grows as \(\tilde{O}(\frac{M}{T} \sqrt{T \log K})\), which is vacuous for \(M \to \infty\).

We state next the main result of this section.

**Theorem 3.** For any \(K \geq 2\) and label set \(\mathcal{Y}\) with \(|\mathcal{Y}| \geq K\), there exists a finite class \(\mathcal{F} \subset \mathcal{D}(\mathcal{Y})^K\) of size \(K\), with \(|\mathcal{Y}| \leq K\), such that for \((\epsilon, 0)\)-locally private online learning scheme \(R^T\) and \(\Phi^T\) (depending on \(\mathcal{F}\)), the KL-risk \(r^K_T(\mathcal{F})\) is lower bounded by

\[
\Omega\left(\sqrt{\frac{K}{\min\{(e^\epsilon - 1)^2, 1\}^\epsilon}}\right)
\]

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

**Sketch of Proof.** At a high level, for any \(K\), our goal is to construct \(K\) pairs of distributions \(\{p_{i,1}, p_{i,2}\}\) for \(i \in [K]\), such that: (i) for any \(i \in [K]\), \(\inf_{\tilde{p}} \text{KL}(p_{i,1}, \tilde{p}), \text{KL}(p_{i,2}, \tilde{p}) \geq \Omega(a)\) where \(a\) is of the order of \(\sqrt{\frac{1}{\epsilon}}\); (ii) \(p_{i,1} - p_{i,2} = \frac{\sqrt{K}}{N} H_i^N\), where each \(H_i^N\) corresponds to distinct columns of a Hadamard matrix of dimension \(K + 1 \leq N \leq 2K\) (excluding the all-ones column). Assume that such a construction exists (see Appendix A). Then, we construct the class \(\mathcal{F}\) of \(2K\) constant functions \(p_{i,\ell}\) for \(i \in [K], \ell \in \{1, 2\}\). Our key technical contribution is to show that for any \((\epsilon, 0)\)-locally differentially private mechanism (possibly depends on \(\mathcal{F}\)), there exists an \(i^* \in [K]\) such that

\[
\text{KL}(p_{i^*, 1}, p_{i^*, 2}) \leq O\left(\frac{\sqrt{\epsilon}}{N}\right) \leq \frac{\sqrt{\epsilon}}{N}\text{ for some } c < 1\text{; here, } p_{i^*, \ell}\text{ stands for the distribution of the private outcomes with input distribution } p_{i^*, \ell}.
\]

This is accomplished by relating the KL-divergence to the \(\chi^2\) divergence and carefully applying Parseval’s identity using the fact that the \(H_i^N\)’s are orthogonal. However, by Pinsker’s inequality, this implies that \(\frac{1}{2} p_{i^*, 1}^2 - p_{i^*, 2}^2\) for \(i \in [K]\). Therefore, by Le Cam’s two point method, no Learner can distinguish \(p_{i^*, 1}, p_{i^*, 2}\) from its privatized samples of length \(T\). Therefore, no algorithm can achieve a KL-risk of \(O(\frac{1}{\epsilon} \sqrt{KT})\) due to property (i) of our construction of \(p_{i,\ell}\)’s (since any such Learner can be used to distinguish \(p_{i^*, 1}, p_{i^*, 2}\)). This completes the proof sketch; see Appendix A for full details.

Note that our proof of Theorem 3 can be extended to the case when \(|\mathcal{Y}| \leq K\) as well, yielding a \(\Omega(\frac{1}{\epsilon} \sqrt{MT})\) KL-risk lower bound, where \(M = |\mathcal{Y}|\). Our proof is also similar in spirit to that outlined in \cite{Edmonds2020} (see also \cite{Gopi2020}), in terms of the use of Parseval’s identity. However, a distinguishing feature of our construction is that our label set \(\mathcal{Y}\) is of size \(K\), while \cite{Edmonds2020} requires a support size of \(2^\eta\). Therefore, for small label set \(\mathcal{Y}\) with \(|\mathcal{Y}| = K\), our result yields a \(\Omega(\frac{1}{\epsilon} \sqrt{KT})\) lower bound, while the lower bound in \cite{Edmonds2020} only implies an \(\Omega(\frac{1}{\epsilon} \sqrt{KT})\) lower bound. Furthermore, our proof is constructive and employs a methodology that is suitable for bounding KL-risk instead of the total variation, which is a result of independent interest.

## 4 APPROXIMATE-DP VIA WMA

We present next an online learning scheme (Algorithm 11) that is \((\epsilon, \delta)\)-locally differentially private and has minimax KL-risk \(r^K_T(\mathcal{F}) = \tilde{O}\left(\frac{1}{\epsilon} \sqrt{TK \log \frac{1}{\delta}}\right)\), where \(\tilde{O}\) hides \(poly(\log(KT))\) factors. Our scheme makes use of the Laplace 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 Learner. At each time \(t \in [T]\), the Learner picks an expert \(i_t, i_t \in [K]\), and observes a loss vector \(v_t = (v_{t,1}, \ldots, v_{t,K}) \in [0, 1]^K\), which represents the loss of choosing each of the experts. Then, the Learner incurs a loss equal to \(v_{i_t, i_t}\). Given \(K\) and \(T\), the weighted majority algorithm executes the following steps. The following lemma gives an upper bound on the regret of WMA, as defined as the expected loss minus the minimum total loss of a single expert.

**Lemma 1** \cite{Shalev-Shwartz2014}. When \(v_t \in [0, 1]^K\) for \(t \in [T]\), the regret of the WMA satisfies

\[
\sum_{t=1}^{T} \sum_{j=1}^{K} \min_{i \in [K]} \sum_{t=1}^{T} v_{t,i} \leq \sqrt{2TK \log K}.
\]
4.2 Our Scheme

Our key idea of leveraging the WMA algorithm in the context of minimizing the KL-risk is to appropriately define the loss vector $v_t$s. At a high level, we will choose the $v_t[j]$s to be the log-likelihoods $-\log f_j(x_t)[y]$ for $j \in [K]$, and then add Laplace noise to the loss vectors. However, instead of sampling one 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 define next random functions for "clipping" the log-likelihoods, which are crucial for applying the Laplace privatization mechanism. The "clipping" technique appeared in Gopi et al. (2020). However, here we use a slightly different form of "clipping functions" for the purpose of bounding the KL divergence.

Clipping of distributions. Let $x^T$ be any realization of the features. For each $t \in [T]$, let
\[
N'_t = \sum_{y \in Y} \left[ M \max_{j \in [K]} f_j(x_t)[y] \right].
\] (3)

Let $\{S_{y,t}\}_{y \in Y}$, $S_{y,t} \subset [N'_t]$, be a partition of $[N'_t]$ such that $|S_{y,t}| = \left[ M \max_{j \in [K]} f_j(x_t)[y] \right]$. Define the random mapping $h'_t : Y \to [N'_t]$ such that $h'_t(y)$ is uniformly distributed over $S_{y,t}$, i.e.,
\[
\Pr(h'_t(y) = y' | y) = \begin{cases} \frac{1}{|S_{y,t}|}, & \text{if } y' \in S_{y,t}, \\ 0, & \text{else}. \end{cases}
\] (4)

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

Also, let $h_t(y) : Y \to [N'_t]$ equal $h'_t(y)$ with probability $1 - \frac{1}{M}$, and $U(y)$ with probability $\frac{1}{M}$; here, $U(y)$ is uniformly distributed over $[N'_t]$, so that
\[
\Pr(h_t(y) = y' | y) = \begin{cases} \frac{1}{|S_{y,t}|}, & \text{if } y' \in S_{y,t}, \\ \frac{1}{M}, & \text{else}. \end{cases}
\] (5)

Similarly, define $h_t \circ f_j(x_t)$, $j \in [K]$, $t \in [T]$, to be the distribution of $h_t(y)$, $y \in Y$, when $y$ is drawn according to distribution $f_j(x_t)$. More formally,
\[
h_t \circ f_j(x_t)[y'] = \sum_{y \in Y} \Pr(h_t(y) = y' | y)f_j(x_t)[y].
\] (6)

By definition of $N'_t$, we have $M \leq N'_t \leq KM$. Therefore,
\[
\frac{1}{TKM} \leq (h_t \circ f_j(x_t))[y'] \leq \frac{1}{M}. \tag{7}
\]

for any $j \in [K]$ and $y' \in [N'_t]$. The fact that $h_t \circ f_j(x_t)$ is upper and lower bounded \footnote{For a real-valued function $s : \mathbb{Z} \to \mathbb{R}$, its sensitivity $\Delta_1(s)$ is defined as $\max_{x \in \mathbb{Z}} s(x) - \min_{x \in \mathbb{Z}} s(x)$.} implies that the log-likelihood $-\log(h_t \circ f_j(x_t)[y])$ has sensitivity $\frac{1}{M}$ equal to $\log(KM)$ and thus can be made $(\epsilon, 0)$-differentially private by adding Laplace noise with scale $\frac{\log(KM)}{\epsilon}$ (Dwork et al., 2014).

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

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

We are now ready to present our privatization scheme. Let $x_t$ and (a random) $Y_t$ be the feature and label at time $t$, respectively, and $	ilde{p}_{t,j} = h_t \circ f_j(x_t)$, for $j \in [K]$. Let $\gamma > 0$ be a value to be determined later. We define the privatized vector $\tilde{Y}_t = (\tilde{Y}_{t,1},\ldots,\tilde{Y}_{t,K})$ as
\[
\tilde{Y}_{t,j} = -c(\gamma)(\log(\tilde{p}_{t,j}[h_t(Y_t)]) + \text{Lap}_j^\gamma + \log M - \epsilon'(\gamma)), \tag{8}
\]
where $h_t$s are the random functions as in \footnote{A Laplace random variable with scale $b$ has density $\frac{1}{2b} \exp(-\frac{|x|}{b})$ for $x \in \mathbb{R}$.} and Lap$_j^\gamma$s are i.i.d. Laplace random variables whose distributions have scale $(2\sqrt{2K \ln \frac{1}{\delta} + \sqrt{K} \epsilon}) \log(KM)$. Moreover, $c(\gamma) = \log(KM) + 2\epsilon(\gamma)$ equals
\[
\frac{(2\sqrt{2K \ln \frac{1}{\delta} + \sqrt{K} \epsilon}) \log(KT)(\gamma + \log K + \log T)}{\epsilon} \tag{9}
\]

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

**Lemma 3**. The privatization $\tilde{R}_t(Y_t) = \tilde{Y}_t$, $t \in [T]$, in Algorithm 2 is $(\epsilon, \delta)$-locally differentially private.

**Proof.** Note that from \footnote{For a real-valued function $s : \mathbb{Z} \to \mathbb{R}$, its sensitivity $\Delta_1(s)$ is defined as $\max_{x \in \mathbb{Z}} s(x) - \min_{x \in \mathbb{Z}} s(x)$.} the sensitivity of $\log(\tilde{p}_{t,j}[h_t(Y_t)])$ is $\log(KM)$ for $j \in [K]$, $t \in [T]$.\footnote{A Laplace random variable with scale $b$ has density $\frac{1}{2b} \exp(-\frac{|x|}{b})$ for $x \in \mathbb{R}$.}
Hence, from (8), \( \hat{Y}_{t,j} \) is \( (\epsilon', \delta') \)-locally differentially private. Invoking Lemma 2, \( Y_t = (\hat{Y}_{t,1}, \ldots, \hat{Y}_{t,K}) \) is \( (\frac{\epsilon'^2}{2} + \sqrt{2\ln(\frac{1}{\delta})}K(\epsilon')^2, \delta) \) differentially private, and thus is \( (\epsilon, \delta) \)-differentially private since
\[
\frac{K(\epsilon')^2}{2} + \sqrt{2\ln(\frac{1}{\delta})}K(\epsilon')^2 \leq \epsilon.
\]
Hence, \( \mathcal{R}^T \) is \( (\epsilon, \delta) \)-locally differentially private. \( \square \)

**Distribution Learning Algorithm.** We now introduce our private learning approach, summarized in Algorithm 1, which uses the above discussed privatization scheme. Note that the prediction \( \hat{p}_i \) at time \( t \) may be different from any of the expert opinions \( h_t \circ f_j(x_t) \) and the candidate distributions \( f_j(x_t), j \in [K] \).

**Algorithm 1 Locally Privatized WMA**

1: **Input:** The hypothesis set \( \mathcal{F} \), the time horizon \( T \), and a probability parameter \( \gamma \).
2: **Initialize:** Set \( K = |\mathcal{F}|, w^1 = (w^1_1, \ldots, w^1_K) = (1, \ldots, 1) \), and \( \eta = \sqrt{2\log \frac{K}{\delta}} \).
3: for \( t = 1, \ldots, T \) do
4: Fetch a feature \( x_t \); 
5: Set \( \hat{p}_i = \sum_{j=1}^{K} \bar{w}_j^{t} \bar{p}_{i,j} \), where \( \bar{p}_{i,j} = h_t \circ f_j(x_t) \) is defined in (6) and \( \bar{w}_j^t = \frac{w_j^t}{\sum_{j=1}^{K} w_j^t} \);
6: Make a prediction \( \hat{p}_i[y] = \sum_{y' \in \mathcal{Y}} S_{y,y'} \hat{p}_i[y'] \) for all \( y \in \mathcal{Y} \), where \( S_{y,y'} \) is defined in (4);
7: Aquire the privatized data \( \hat{Y}_t = (\hat{Y}_{t,1}, \ldots, \hat{Y}_{t,K}) \), as defined in (8);
8: Update \( w_{j}^{t+1} = w_j^t e^{-\eta\hat{Y}_{t,j}} \) for \( j \in [K] \); 
9: end for

We now introduce the following key lemma, which establishes an upper bound for the KL-risk conditioned on the event that the Laplace noise is bounded.

**Lemma 4.** In Algorithm 1, if \( \hat{Y}_{t,j} \in [0, 1] \) for \( j \in [K], t \in [T] \) and \( T > \log K \), then
\[
\sum_{t=1}^{T} \sum_{j=1}^{K} \bar{w}_j^t \hat{Y}_{t,j} - \min_{i \in [K]} \left( \sum_{t=1}^{T} \hat{Y}_{t,i} \right) \leq \sqrt{2T \log K}.
\]

We now introduce the following key lemma, which establishes an upper bound for the KL-risk conditioned on the event that the Laplace noise is bounded.

**Lemma 5.** For any \( f \in \mathcal{F} \) and \( x^T \in \mathcal{X}^T \), with probability at least \( 1 - e^{-\gamma} \) wrt the randomness of \( \text{Lap}^j \)'s, and for any \( \gamma > 0 \), the output \( \hat{p}_i \) of Algorithm 1 satisfies
\[
\mathbb{E}_{Y^T} \left[ \sum_{t=1}^{T} \mathbb{K}(f(x_t), \hat{p}_i) \right] \leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT)
\]

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

**Sketch of Proof.** We only present high-level ideas and refer to Appendix B for a detailed proof. We show that this holds true w.p. \( \geq 1 - e^{-\gamma} \) by our definition of \( \hat{Y}_{t,j} \)'s in (6) and through the use of concentration properties of Laplace distributions. Now, by conditioning on such an event, we show by Lemma 4 and definition of \( \hat{Y}_t \) that for any \( j^* \in [K] \) we have \( \sum_{t=1}^{T} \mathbb{K}^t(\hat{p}_{j^*}, \hat{p}_i) \leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + \beta \), where \( \beta \) depends only on the Laplace variable \( \text{Lap}^j \) and \( \hat{Y}_t' = h_t(Y_t) \). Assuming \( \hat{Y}_t' \sim f_{j^*}(x_t) \) and taking the expectation wrt the randomness of the \( Y_t' \)'s, we have
\[
\mathbb{E}_{Y^T} \left[ \sum_{t=1}^{T} \mathbb{K}(\hat{p}_{j^*}, \hat{p}_i) \right] \leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + \beta.
\]

**Theorem 4.** For any class \( \mathcal{F} \) of size \( K \) there exist an \( (\epsilon, \delta) \)-local differential private mechanism that achieves a KL-risk \( r^K_{\text{KL}}(\mathcal{F}) \) upper bounded by
\[
O \left( \sqrt{2T \log K} \left( \log(KT) + \frac{(\sqrt{K \ln \frac{1}{\delta}} + \sqrt{KT}) \log^2(KT)}{c(\gamma)} \right) \right).
\]

**Proof.** Let \( \mathcal{E} \) be the event that for all \( j, t \), \( |\text{Lap}^j_t| \leq c'(\gamma) \), which happens w.p. \( \geq 1 - e^{-\gamma} \) and implies that Lemma 5 holds (see Appendix B). We have
\[
r^K_{\text{KL}}(\mathcal{F}) = \mathbb{E}_{Y^T} \left[ \sum_{t=1}^{T} \mathbb{K}(f_{j^*}(x_t), \hat{p}_i) \right]
\]
\[
= \Pr(\mathcal{E}) \mathbb{E}_{Y^T} \left[ \sum_{t=1}^{T} \mathbb{K}(f_{j^*}(x_t), \hat{p}_i) | \mathcal{E} \right]
\]
\[
+ \Pr(\mathcal{E}^c) \mathbb{E}_{Y^T} \left[ \sum_{t=1}^{T} \mathbb{K}(f_{j^*}(x_t), \hat{p}_i) | \mathcal{E}^c \right]
\]
\[
\leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT)
\]

where \( j^* \in [K] \) is the underlying truth. 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 that (A) can be upper bounded by
\[
\mathbb{E}_{Y^T} \left[ \frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT) \right]
\]
\[
+ \sum_{t=1}^{T} \sum_{j=1}^{K} \tilde{w}_j^t \text{Lap}_j - \sum_{t=1}^{T} \text{Lap}_j, |E| \].
\]

Also note that even when conditioning on \( \mathcal{E} \), the Laplace random variables \( \text{Lap}_j \)'s remain i.i.d. and the \( \tilde{w}_j^t \)'s still sum up 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 analyze the term (B), by (30) and (31) (Appendix [B]) we have that \( \text{KL}(f_t, (x_t), \tilde{\mu}_t) \) is upper bounded by \( \text{KL}(\tilde{\mu}_{t,j}^*, \tilde{\mu}_t) + O \left( \frac{\log(KT)}{\epsilon} \right) \). Noting that \( \frac{\tilde{\mu}_{t,j}^*[y]}{\tilde{\mu}_t[y]} \leq KT \) for all \( y' \in [N_t'] \) (see (7)), we conclude that \( \text{KL}(\tilde{\mu}_{t,j}^*, \tilde{\mu}_t) \leq \log(KT) \). Therefore, by summing over \( t \in [T] \), we see that the term (B) is upper bounded by \( e^{-\gamma}(T \log(KT) + O(\log(KT))) \), since \( \Pr[\mathcal{E}] \leq \epsilon^{-\gamma} \). Putting everything together, the KL-risk \( \tau_{KL}^t(F) \) is upper bounded by

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

The theorem now follows by setting \( \gamma = \log T \) and from 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

\[
O \left( \frac{1}{\epsilon} \sqrt{TK \log^2(KT) \log \frac{1}{\delta}} \right)
\]

for sufficiently small \( \epsilon \) and \( \delta \).

5 Pure-DP via Modified EXP3

While Algorithm 1 offers (\( \epsilon, \delta \))-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 that it is possible to achieve the same \( O \left( \frac{1}{\epsilon} \sqrt{TK} \right) \) KL-risk bound. By Theorem 3, we known that this is essentially tight for pure-DP constrains.

Note that the reason why Theorem 4 has a dependency on \( \delta \) is due to the advanced composition lemma (Lemma 2) that allows us to select the scale of the Laplace distribution with order of \( \sqrt{K} \), which is essential for achieving a \( \sqrt{KT} \)-type bound. To resolve this issue, we now introduce a new privatization mechanism by selecting a single random component in the log-likelihood vectors and then revealing only the privatized version of such a component, as in Algorithm 2.

Note that this significantly reduces the scale of the Laplace distribution from \( \sqrt{K} \) to \( O(1) \). However, since we only return a single component of the loss vector, the WMA algorithm cannot be used directly. To resolve this issue, we instead perform an analysis similar to that for 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 compared to the standard EXP3 algorithm is that we do not reveal the loss of the expert selected by the Learner but instead 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 \( x_t \) and \( Y_t \) be the feature and label at time \( t \), and \( \tilde{\mu}_{t,j} = h_t \circ f_j(x_t) \) as in [6]. Let \( \gamma > 0 \) be a value to be determined later, and define the vector \( Z_{t,j} = (Z_{t,j,1}, \cdots, Z_{t,j,K}) \) according to

\[
Z_{t,j,i} = \begin{cases} 
-\epsilon_{\text{pdf}}(\gamma)(\log (\tilde{\mu}_{t,j}[h_t(Y_t)])) + \text{Lap}_j + \log M - \epsilon_{pdf}(\gamma), & \text{if } i = j \\
0, & \text{else,}
\end{cases}
\]

where \( \text{Lap}_j \) are independent Laplace random variables with scale \( \log(KT) \), and where \( h_t \) is the random function from [5]. Moreover, \( \epsilon_{pdf}(\gamma) = \frac{1}{\log(KT)} + 2\epsilon_{pdf}(\gamma) \epsilon \)

and

\[
\epsilon_{pdf}(\gamma) = \frac{\log(KT)(\gamma + \log K + \log T)}{\epsilon}.
\]

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

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

Proof. Since for each \( t \), only a single entry of \( \tilde{Y}_t \) is nonzero, and \( \log(\tilde{\mu}_{t,j}[h_t(Y_t)]) \) has sensitivity \( \log(KT) \), the Laplace mechanism with scale \( \log(KT) \) ensures \( \epsilon, 0 \)-differential privacy. \( \square \)

We present our learning algorithm in Algorithm 2. The following key lemma bounds the regret that extends Lemma 4 for expected loss vectors.

Lemma 7. In Algorithm 2 if \( Z_{t,j} \in [0,1]^K \) 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.
Algorithm 2 Locally Pure-DP Algorithm

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

**Proof.** Let \( \tilde{Y}_t = (\tilde{Y}_{t,1}, \ldots, \tilde{Y}_{t,K}) \) be as in Algorithm 2. We use a more general result of Lemma 3 which is implied in the proof of Shalev-Shwartz and Ben-David (2014) Thm 21.11, that is,

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

Taking expectations over \( J_t \)s, we have

\[
\max_{i \in [K]} \mathbb{E} \left[ \sum_{t=1}^T \left( \sum_{j=1}^K \tilde{w}_t^j \tilde{Y}_{t,j} - \tilde{Y}_{t,i} \right) \right] \leq \frac{\log K}{\eta} + \mathbb{E} \left[ \frac{\eta}{2} \sum_{t=1}^T \sum_{j=1}^K \tilde{w}_t^j \tilde{Y}_{t,j}^2 \right] \leq \frac{(a) \log K}{\eta} + \frac{\eta}{2} \sum_{t=1}^T \frac{1}{K} = \frac{\log K}{\eta} + \frac{\eta T}{2 K}.
\]

where \((a)\) follows from \( \mathbb{E}_{J_t}[\tilde{Y}_{t,j}^2] = \frac{2\tilde{w}_t^j}{K} \leq \frac{1}{K} \) and the fact that the \( \tilde{w}_t^j \)'s are independent of \( J_t \). Setting \( \eta = \sqrt{2K \log K} \) and noticing that \( \mathbb{E}_{J_t}[\tilde{Y}_{t,j}] = \frac{2\tilde{w}_t^j}{K} \), the result follows.

The following lemma is analogous to Lemma 3 and we defer its proof to Appendix C.

**Lemma 8.** For any \( f \in \mathcal{F} \) and \( x^T \in \mathcal{X}^T \), with probability at least \( 1 - e^{-\gamma} \) wrt the randomness of Laplace noise \( \text{Lap}_\gamma \), the output \( \hat{p}_t \) of Algorithm 2 satisfies

\[
\mathbb{E}_{Y^t \sim J_t} \left[ \sum_{i=1}^T \text{KL}(f(x_i), \hat{p}_t) \right] \leq \frac{1}{c_{dp}(\gamma)} \sqrt{2TK \log K} + 3 \log(KT)
\]

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

We are now ready to state the main result of this section. The proof essentially follows the same steps as the proof of Theorem 3 and is therefore omitted due to space constraints.

**Theorem 5.** For any class \( \mathcal{F} \) of size \( K \), there exists an \( (\epsilon, 0) \)-locally differentially private mechanism that achieves a KL-risk \( r_T^{KL} \) upper bounded by

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

6 BOUNDING THE AVERAGED TV-RISK

We conclude our exposition by showing how the KL-risk upper bound for our \((\epsilon, 0)\)-local differential private algorithm (Theorem 5) can be used to obtain bounds for the averaged TV-risk introduced in Section 2.

The next result follows directly from Pinsker’s inequality (Polyanskiy and Wu 2022).

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

\[
\tilde{O} \left( \sqrt{\frac{1}{\epsilon^2} \frac{1}{T}} \right),
\]

provided that \( \epsilon \) is sufficiently small.

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

\[
\frac{1}{T} \sum_{t=1}^T |f(x_t) \hat{p}_t|_{TV} \leq \frac{1}{T} \sum_{t=1}^T \sqrt{\frac{1}{2} KL(f(x_t), \hat{p}_t)} \leq \frac{1}{T} \sum_{t=1}^T \frac{KL(f(x_t), \hat{p}_t)}{2} \leq \tilde{O} \left( \sqrt{\frac{1}{\epsilon^2} \frac{1}{T}} \right),
\]


where the second inequality follows from the Cauchy-Schwartz inequality, while the last inequality follows from Theorem 5.

Finally, we determine $T_\alpha$ that makes the TV-risk bounded by some given $\alpha$. Let $Q = \{p_1, \ldots, p_K\}$ be $K$ distributions. We construct a class $\mathcal{F}$ of $K$ constant functions with values in $Q$. Denote by $\hat{p}_n$ the learners that achieve the upper bound of Theorem 5 and write $\hat{p} = \frac{1}{T}\sum_{t=1}^T \hat{p}_t$. By convexity of the TV-distance and Jensen’s inequality, we have that for any ground truth distribution $p^* \in Q$, $|p^* - \hat{p}|_{TV} \leq \tilde{O}\left(\sqrt{\frac{1}{\epsilon} \sqrt{\frac{K}{T}}}\right)$. Taking $T_\alpha = \tilde{O}\left(\frac{K}{\alpha^2}\right)$ one can make the TV-risk upper bounded by $\alpha$. This can be boosted to high probability via a median trick (Wu et al., 2023a), 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 bound in Theorem 5 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 in Gopi et al. (2020) (via suitable comparison schemes) that the sample complexity of the batch setting is upper bounded by $\tilde{O}\left(\frac{K}{\epsilon^2 \alpha^2}\right)$ when using interactive private mechanisms. It is therefore also interesting to investigate if such comparison-based arguments can be extended to our online case.

**Acknowledgements**

This work was partially supported by the NSF Center for Science of Information (CSoI) Grant CCF-0939370, and also by NSF Grants CCF-2006440, CCF-2007238, and CCF-2211423.

**References**


and large classes of experts.  


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

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

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

for $i \in [K], y \in [N]$. Consider learning a hypothesis set $\mathcal{F} = \{f_{1,1}, f_{1,2}, f_{2,1}, \ldots, f_{K,1}, f_{K,2}\}$, where $f_{i,\ell}(x) = p_{i,\ell}$ for all $x \in \mathcal{X}$, $i \in [K]$, 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 $Y^T$, where $Y_t = \mathcal{R}_t(Y_t)$, such that the KL distance $\text{KL}(p, \hat{p})$ is minimized. Note that one can use a private online learning learner $\Phi^T$ to make the estimate $\hat{p}$ as follows: set $\hat{p} = \frac{1}{T} \sum_{t=1}^T \hat{p}_t$ where $\hat{p}_t = \Phi_t(x^t, \hat{Y}^{t-1})$. Let $f_{j,\ell}$ be the ground truth function. Then

$$\mathbb{E}_{Y^T}[\text{KL}(p, \hat{p})] \leq \frac{\mathbb{E}_{Y^T}[\sum_{t=1}^T \text{KL}(f_{j,\ell} Y^t, \hat{p}_t)]}{T} = \frac{r_{KL}(\mathcal{F})}{T}. \quad (17)$$

In the following, we show, via the Le Cam’s two point method (Polyanskiy and Wu, 2022), that

$$\mathbb{E}_{Y^T}[\text{KL}(p, \hat{p}(Y^T))] \geq a \leq \frac{a}{2 \cdot 132} = \frac{a}{132} \quad (18)$$

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

Note that for any $\hat{p} \in \mathcal{D}([N])$,

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

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

$$|\hat{p}_{i^*,1} - p_{i^*,1}^0|_{TV} \leq \frac{1}{3},$$

where $\hat{p}_{i,\ell}$ 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 $\tilde{y} \in \tilde{Y}$, where $\tilde{Y}$ is the output space of $\mathcal{R}_t$. For each $\tilde{y} \in \tilde{Y}$, let $q(\tilde{y}) = (q(\tilde{y}|1), \ldots, q(\tilde{y}|N))$ be a vector and let

$$q(\tilde{y}) = \sum_{y=1}^N b(\tilde{y}|y) H^N_y$$

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

$$\hat{p}_{i,\ell}(\tilde{y}) = \sum_{y=1}^N q_t(\tilde{y}|y)p_{i,\ell}(y).$$
From \[16, 20, \] and \[21, \] we know that
\[
\hat{p}_{i,1}(\tilde{y}) = \frac{2 \sum_{y: H_{y+1}^N, y = -1} q_t(\tilde{y}|y)}{N}, \quad \text{and} \quad \hat{p}_{i,2}(\tilde{y}) = \frac{2 \sum_{y: H_{y+1}^N, y = -1} q_t(\tilde{y}|y)}{N} + ab(\tilde{y})_{i+1},
\]
for \(i \in [K]\). The KL distance
\[
\begin{align*}
\text{KL}(\hat{p}_{i,1}, \hat{p}_{i,2}) &= \int_{\tilde{y}} -\hat{p}_{i,1}(\tilde{y}) \log \left( \frac{\hat{p}_{i,2}(\tilde{y})}{\hat{p}_{i,1}(\tilde{y})} \right) d\tilde{y} \\
&\leq (a) \int_{\tilde{y}} -\hat{p}_{i,1}(\tilde{y}) \left( \frac{\hat{p}_{i,2}(\tilde{y}) - \hat{p}_{i,1}(\tilde{y})}{\hat{p}_{i,1}(\tilde{y})} \right)^2 d\tilde{y} \\
&= \int_{\tilde{y}} (\hat{p}_{i,2}(\tilde{y}) - \hat{p}_{i,1}(\tilde{y}))^2 d\tilde{y} \\
&\leq \int_{\tilde{y}} a^2 b(\tilde{y})_{i+1}^2 d\tilde{y}, 
\end{align*}
\]
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
\[
\sum_{y \in [N-1]} b(\tilde{y})_{i+1}^2 = \sum_{y \in [N-1]} \left( \frac{\sum_{i \in [N]} q_t(\tilde{y}|i) H_{y+1}^N}{N} \right)^2 \\
\overset{(a)}{=} \sum_{i \in [N]} q_t(\tilde{y}|i)^2 - \left( \frac{\sum_{i \in [N]} q_t(\tilde{y}|i)}{N} \right)^2 \\
\overset{(b)}{\leq} \sum_{i \in [N]} \left( \sum_{j \in [K]} (q_t(\tilde{y}|i) - q_t(\tilde{y}|j))^2 \right) \\
\overset{(c)}{\leq} \min\{\epsilon^2 - 1, 1\} \sum_{i \in [N]} q_t(\tilde{y}|i)^2 \\
\leq \frac{1}{2N} \sum_{i \in [N]} \min_{y} q_t(\tilde{y}|y) q_t(\tilde{y}|i), 
\]
where \((a)\) follows from the Parseval’s identity and the fact that \(\{H_{y+1}^N\}_{y \in [N-1]}\) form orthogonal base with the all-ones vector \(H_{1}^N\) being excluded, \((b)\) follows from the elementary identity \(N \sum_{i=1}^{N} a_i^2 - (\sum_{i=1}^{N} a_i)^2 = \frac{1}{4} \sum_{i,j \leq N} (a_i - a_j)^2\), and \((c)\) follows from property of \((\epsilon, 0)\)-differential privacy. Now, \([22]\) and \([23]\) imply
\[
\begin{align*}
\sum_{i \in [K]} \text{KL}(\hat{p}_{i,1}, \hat{p}_{i,2}) &\leq \int_{\tilde{y}} \sum_{i \in [N]} a^2 b(\tilde{y})_{i+1}^2 \min_{y} q_t(\tilde{y}|y) d\tilde{y} \\
&\leq a^2 \int_{\tilde{y}} \min\{\epsilon^2 - 1, 1\} \sum_{i \in [N]} q_t(\tilde{y}|i) d\tilde{y} \\
&\leq a^2 \frac{1}{2} \min\{\epsilon^2 - 1, 1\} \epsilon. 
\end{align*}
\]
Therefore, there exists \(i^* \in [K]\) such that
\[
\text{KL}(\hat{p}_{i^*,1}, \hat{p}_{i^*,2}) \leq \frac{a^2 \min\{\epsilon^2 - 1, 1\} \epsilon}{2K},
\]
which implies that
\[
|\hat{p}_{i^*,1}^{\otimes T} - \hat{p}_{i^*,2}^{\otimes T}|_{TV} \leq \sqrt{2 \text{KL}(\hat{p}_{i^*,1}, \hat{p}_{i^*,2})}
\]
Hence, by taking expectation over the distribution of
\[ \leq \sqrt{2TKL(\tilde{p}_{i,1}, \tilde{p}_{i,2})} \]
\[ \leq \frac{1}{3} \]  
\[ (24) \]

The rest of the argument is standard. Let \( p \) be one of \( p_{i,1} \) and \( p_{i,2} \) and let \( \phi: \tilde{Y}^T \to \{p_{i,1}, p_{i,2}\} \) be a classification function deciding if \( p \) is \( p_{i,1} \) or \( p_{i,2} \) based on the privatized data \( \tilde{Y}^T \). From (24) and Le Cam’s two point lemma, the classification error
\[
\Pr(\phi(\tilde{Y}^T) \neq p_{i,t^*}) \geq \frac{1 - |\tilde{p}_{i,1} - \tilde{p}_{i,2}|_TV}{2} \geq \frac{1}{3},
\]
where \( p = p_{i,t^*} \).

On the other hand, suppose there exists a private hypothesis testing estimator \( \hat{p} \) satisfying \( KL(p_{i,t^*}, \hat{p}(\tilde{Y}^T)) \leq \frac{\gamma}{2} \) with probability at least \( \frac{3}{4} \). Then, from (19) we conclude that a minimum KL distance classifier \( \phi(\tilde{Y}^T) = \text{argmin}_{p \in \{p_{i,1}, p_{i,2}\}} KL(p, \hat{p}(\tilde{Y}^T)) \) 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 \( E \) denote the event that there exists \( j \in [K], t \in [T] \) such that
\[
|\text{Lap}_j^t| \geq c'(\gamma),
\]
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 \( E \) is at most \( e^{-\gamma} \).

In the following, we show that Lemma 4 holds when \( E \) does not occur, which has probability at least \( 1 - e^{-\gamma} \).

Note that from (7), we have \( \log(\tilde{p}_{i,j}[h_t(Y_t)]) + \text{Lap}_j^t + \log M - c'(\gamma) \in [-\log(KT) - 2c'(\gamma), 0] \),

and thus that \( \tilde{Y}_{t,j} \in [0, 1] \). According to Lemma 4, we have
\[
\sum_{t=1}^{T} \sum_{j=1}^{K} \bar{w}_{j}^t \left( - \log(\tilde{p}_{i,j}[h_t(Y_t)]) - \text{Lap}_j^t \right) - \min_{i \in [K]} \left( \sum_{t=1}^{T} \left( - \log(\tilde{p}_{i,j}[h_t(Y_t)]) - \text{Lap}_j^t \right) \right) \leq \frac{1}{c(\gamma)} \sqrt{2T\log K}.
\]

Let \( j^* \in [K] \) be any index, and assume that \( Y_t \sim f_{j^*}(x_t) \). Taking expectation over the distribution of \( Y_t' \sim h_t(Y_t) \) we find that
\[
E_{Y^T} \left[ \sum_{t=1}^{T} \sum_{j=1}^{K} \bar{w}_{j}^t KL(\tilde{p}_{i,j^*}, \tilde{p}_{i,j}) \right] \leq \frac{1}{c(\gamma)} \sqrt{2T\log K} + \sum_{t=1}^{T} E_{Y^T} [\bar{w}_{j}^t] \text{Lap}_j^t - \sum_{t=1}^{T} \text{Lap}_j^t,
\]
where we used the fact that \( Y_t' \) distributed as \( \tilde{p}_{i,j^*} = h_t \circ f_{j^*}(x_t) \) and \( E_{Y \sim p} \log \left( \frac{\bar{w}_{j}^t}{\bar{w}_{j}^{T'}} \right) = KL(p, q) \). Moreover, we have by Jensen’s inequality and convexity of \( -\log(x) \) that
\[
\sum_{t=1}^{T} \left( - \log \left( \sum_{j=1}^{K} \bar{w}_{j}^t \tilde{p}_{i,j}[h_t(Y_t)] \right) \right) - \sum_{t=1}^{T} \left( - \log \left( \tilde{p}_{i,j^*}[h_t(Y_t)] \right) \right) \leq \sum_{t=1}^{T} \sum_{j=1}^{K} \bar{w}_{j}^t \left( - \log \left( \tilde{p}_{i,j}[h_t(Y_t)] \right) \right) - \sum_{t=1}^{T} \left( - \log \left( \tilde{p}_{i,j^*}[h_t(Y_t)] \right) \right).
\]

Hence, by taking expectation over the distribution of \( Y_t' \sim h_t(Y_t) \)
\[
E_{Y^T} \left[ \sum_{t=1}^{T} KL(\tilde{p}_{i,j^*}, \tilde{p}_\tilde{H}) \right] \leq E_{Y^T} \left[ \sum_{t=1}^{T} \sum_{j=1}^{K} \bar{w}_{j}^t KL(\tilde{p}_{i,j^*}, \tilde{p}_{i,j}) \right].
\]
Combining (27) and (28), we arrive at

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

(29)

On the other hand,

\[
\text{KL}(f_j^*\cdot(\mathbf{x}_t), \tilde{p}_t) = \sum_{y \in Y} f_j^*\cdot(\mathbf{x}_t)[y] \log \left( \frac{f_j^*\cdot(\mathbf{x}_t)[y]}{\tilde{p}_t[y]} \right)
\]

\[
= \sum_{y \in Y} \sum_{y' \in S_{y,t}} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] \cdot \log \left( \frac{\sum_{y'' \in S_{y,t}} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y]}{\sum_{y'' \in [K]} \tilde{w}_j^{t} h'_t \circ f_j(\mathbf{x}_t)[y']} \right)
\]

\[
= \text{KL}(h'_t \circ f_j^*\cdot(\mathbf{x}_t), \tilde{p}_t),
\]

(30)

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

\[
\text{KL}(h'_t \circ f_j^*\cdot(\mathbf{x}_t), \tilde{p}_t)
\]

\[
= \sum_{y \in [N]} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] \log \left( h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] \right) - \sum_{y' \in [N]} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] \log \left( \tilde{p}_t[y'] \right)
\]

\[
= \sum_{y' \in [N]} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] \log \left( \left(1 - \frac{1}{T}\right) h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] + \frac{1}{T N_j} \right) - \sum_{y' \in [N]} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] \log \left( \tilde{p}_t[y'] \right) - \log \left(1 - \frac{1}{T}\right)
\]

\[
\leq \sum_{y' \in [N]} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] \cdot \log \left( \left(1 - \frac{1}{T}\right) h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] + \frac{1}{T N_j} \right) - \sum_{y' \in [N]} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] \log \left( \tilde{p}_t[y'] \right) - \log \left(1 - \frac{1}{T}\right)
\]

\[
= \sum_{y' \in [N]} \left( \left(1 - \frac{1}{T}\right) h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] + \frac{1}{T N_j} \right) \log \left( \left(1 - \frac{1}{T}\right) h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] + \frac{1}{T N_j} \right)
\]

\[
+ \sum_{y' \in [N]} \left( \frac{1}{T} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] + \frac{1}{T N_j} \right) \log \left( \frac{1}{T} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] + \frac{1}{T N_j} \right) - \log \left(1 - \frac{1}{T}\right)
\]

\[
= \text{KL}(\tilde{p}_{t,j^*}, \tilde{p}_t) + \sum_{y' \in [N]} \left( \frac{1}{T} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] + \frac{1}{T N_j} \right) \log \left( \frac{1}{T} h'_t \circ f_j^*\cdot(\mathbf{x}_t)[y'] + \frac{1}{T N_j} \right) - \log \left(1 - \frac{1}{T}\right)
\]

(31)

where (a) follows because of (7) and the fact that \( \tilde{p}_t \) is a linear combination of \( \tilde{p}_j, j \in [K] \). Combining (29), (30), and (31), we have

\[
E_{Y^T} \left[ \sum_{t=1}^{T} \text{KL}(f_j^*\cdot(\mathbf{x}_t), \tilde{p}_t) \right] = E_{Y^T} \left[ \sum_{t=1}^{T} \text{KL}(h'_t \circ f_j^*\cdot(\mathbf{x}_t), \tilde{p}_t) \right]
\]

\[
\leq E_{Y^T} \left[ \sum_{t=1}^{T} \text{KL}(\tilde{p}_{t,j^*}, \tilde{p}_t) - T \log(1 - \frac{1}{T}) + 2 \log(KT) \right]
\]

\[
\leq \frac{1}{c(\gamma)} \sqrt{2T \log K} + 3 \log(KT) + \sum_{t=1}^{T} \sum_{j=1}^{K} E_{Y^T} [\tilde{w}_j^{t}] \text{Lap}_j^t - \sum_{t=1}^{T} \text{Lap}_j^t.
\]

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

$$|\text{Lap}_t^j| \geq c_{pdp}'(\gamma).$$

(32)

Then, the probability of $E$ is at most $e^{-\gamma}$. We show that Lemma 8 holds when $E$ does not occur. Note that from (7), we have

$$\log(\bar{p}_{t,j}[h_t(Y_t)]) \in [\log(\frac{1}{TKM}), \log(\frac{1}{MT})].$$

Therefore,

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

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

Invoking Lemma 7, we have

$$\max_{i \in [K]} E_{j}^T \left[ \sum_{t=1}^{T} \left( \sum_{j=1}^{K} \tilde{w}_t^j Z_{t,j,j} - \sum_{i=1}^{T} Z_{t,i,i} \right) \right] = \max_{i \in [K]} E_{j}^T \left[ \sum_{t=1}^{T} \left( \sum_{j=1}^{K} \tilde{w}_t^j (-c_{pdp}(\gamma)(\log(\bar{p}_{t,j}[h_t(Y_t)]) + \text{Lap}_t^j)) \right) \right.

$$

$$\left. - c_{pdp}(\gamma)(\log(\bar{p}_{t,j}[h_t(Y_t)]) + \text{Lap}_t^j) \right] \leq \sqrt{2TK \log K}.$$  (33)

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

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

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

$$E_{Y^{T},j^{T}} \left[ \sum_{t=1}^{T} \text{KL}(f_{j^*}(x_t), \hat{p}_t) \right] \leq \frac{1}{c_{pdp}(\gamma)} \sqrt{2TK \log K} + 3\log(KT) + \sum_{t=1}^{T} \sum_{j=1}^{K} E_{Y^{T},j^{T}} \left[ \tilde{w}_t^j \text{Lap}_t^j - \sum_{t=1}^{T} \text{Lap}^j \right].$$

This completes the proof.