Supplementary Material

May 2023

1 Preliminaries

Definition S1 (ϵ-DP, Dwork et al. (2006b)). A mechanism M is ϵ-DP if the following holds:

$$\sup_{S,S'\subset\mathcal{X},S\sim S'}\ \sup_{A\subset\mathcal{Z}} \mathbb{P}\left(M(S)\in A\right) - e^{\epsilon}\mathbb{P}\left(M(S')\in A\right) \leq 0,$$

where $S\sim S'$ denotes that data sets S and S' differ only by one individual.

Definition S2 ((ϵ,δ)-DP, Dwork et al. (2006a)). A mechanism M is (ϵ,δ)-DP if the following holds:

$$\sup_{S,S'\subset\mathcal{X},S\sim S'}\ \sup_{A\subset\mathcal{Z}} \mathbb{P}\left(M(S)\in A\right) - e^{\epsilon}\mathbb{P}\left(M(S')\in A\right) \leq \delta.$$

Definition S3 (Locally private mechanism, variant of Duchi et al. (2013)). For a multivariate mechanism $(Z_1,\ldots,Z_n) = M(X_1,\ldots,X_n)$ defines a collection of randomized mechanisms $\{M_i\}_{i=1}^n$ as $M_i(X_1,\ldots,X_n) = Z_i$. Here, M is a locally private mechanism if M_i takes X_i for its input. Furthermore, M is said to be sequential if M_i depends only on the realizations of $Z_1,\ldots,Z_{i-1}$ for all i.

Definition S4 ((ϵ,δ)-LDP, Asoodeh et al. (2021)). A given locally private mechanism $M:\mathcal{X}^n \longrightarrow \mathcal{P}\left(\mathcal{Z}\right)$ is (ϵ,δ)-LDP if

$$\sup_{x,x'\in\mathcal{X}}\ \sup_{A\subset\mathcal{Z}} \left|\mathbb{P}_{M_i(x)}\left(A\right) - e^{\epsilon}\mathbb{P}_{M_i(x')}\left(A\right)\right| \leq \delta$$

for $i=1,\ldots,n$.

Definition S5 (FDP, Dong et al. (2022)). Let $f:[0,1]\longrightarrow[0,1]$ be a trade-off function for some distributions P and Q. A given mechanism M is f-FDP if

$$T(M(S),M(S'))(\alpha) \geq f(\alpha)$$

for every $S\sim S'\subset\mathcal{X}^n$ and $\alpha\in[0,1]$.

Definition S6 (f-FLDP and μ-GLDP)**.** A given locally private mechanism $M : \mathcal{X}^n \longrightarrow \mathcal{P}(\mathcal{Z})$ is f-FLDP (or μ-GLDP) if

$$T(M_i(x), M_i(x'))(\alpha) \geq f(\alpha) \text{ (or } G_\mu(\alpha))$$

for every $x, x' \in \mathcal{X}$ and $i = 1, \ldots, n$.

Definition S7 (Minimax risk)**.**

$$\mathcal{R}_n(\theta(\mathcal{P}), L) := \inf_{\hat{\theta}:\mathcal{X}^n \longrightarrow \Theta} \sup_{P \in \mathcal{P}} \mathbb{E}_P \left[L\left(\hat{\theta}\left(X_1, \ldots, X_n\right), \theta(P) \right) \right].$$

2 Existing results

Proposition S1 (Le Cam's method, LeCam (1973); Yu (1997); Tsybakov (2009))**.** *For two distributions $P_1, P_2 \in \mathcal{P}$, suppose $\rho(\theta(P_1), \theta(P_2)) \geq 2\eta > 0$ holds. Then,*

$$\begin{aligned} \mathcal{R}_n\left(\theta(\mathcal{P}), \Phi \circ \rho\right) &= \inf_{\hat{\theta}:\mathcal{X}^n \longrightarrow \Theta} \sup_{P \in \mathcal{P}} \mathbb{E}_P \left[\Phi\left(\rho\left(\hat{\theta}\left(X_1, \ldots, X_n\right), \theta(P) \right) \right) \right] \\ &\geq \frac{\Phi(\eta)}{2} \left(1 - |P_1^n - P_2^n\|_{TV}\right). \end{aligned}$$

Proposition S2 (Assoaud's method, Assouad (1983); Yu (1997); Duchi et al. (2018))**.** *For a family $\mathcal{P}$ of distributions and a loss function $\Phi \circ \rho$, suppose a set of distributions $\{P_v\}_{v \in \{-1,+1\}^d} \subset \mathcal{P}$ induces 2η-Hamming separation under $\Phi \circ \rho$. Then, we have*

$$\mathcal{R}_n\left(\theta(\mathcal{P}), \Phi \circ \rho\right) \geq \eta \sum_{j=1}^{d} \left(1 - ||P_{+j}^n - P_{-j}^n||_{TV}\right)$$

where $P_{\pm j}^n(x_1, \ldots, x_n) := \frac{1}{2^{d-1}} \sum_{v:v_j = \pm 1} P_v(x_1) \cdots P_v(x_n)$.

3 Contraction inequality of FLDP

In discussing involving private mechanisms, the term *contraction* is frequently used. As pointed out by Duchi et al. (2018), one consequence of private mechanisms on statistical utility is a contraction of the effective sample size. Another usage is the contraction of the distribution mechanism. In this section, we focus on the latter aspect, especially examining the f-FLDP mechanism. Both Le Cam's and Assouad's methods obtain lower bounds by considering a random selection from a set of distributions in $\mathcal{P}$. Hence, we assume that target distributions are randomly chosen from a prior family of distributions $\{P_v\}_{v \in V}$. Subsequently, our estimate $\hat{\theta}$ is fixed for given data, so the loss relies on the disparity between the distributions $\{P_v\}_{v \in V}$ and the target parameters $\{\theta(P_v)\}_{v \in V}$. In the private setting, mechanisms impact the distributions by converting the family of distributions $\{P_v\}_{v \in V}$ into a family of closer distributions $\{M(P_v)\}_{v \in V}$, but not target parameters, which can be seen in (1) and (2) in the main paper. Consequently, the difference in minimax risk between non-private and private estimation heavily depends on the contraction ability of the mechanism. The characteristic of a mechanism that transforms input distributions into output distributions that are closer has already been studied in the context of the contraction coefficient of the Markov kernel (Asoodeh et al., 2021). The Markov kernel can be identified with a mechanism in the theoretical sense, which we will adopt in our study.

The contraction coefficient of a mechanism under D_f-divergence is defined as follows:

$$\sup_{P,Q\in\mathcal{P}(\mathcal{X})} \frac{D_f(M(P)||M(Q))}{D_f(P||Q)}$$

where the f-divergence $D_f(P||Q)$ is defined as $D_f(P||Q) = \mathbb{E}_Q\left[f\left(\frac{P}{Q}\right)\right]$ for a given convex function $f : (0, \infty) \longrightarrow \mathbb{R}$ such that $f(1) = 0$. Although it cannot be represented in the form of a contraction coefficient of a specific divergence, Duchi et al. (2018) showed $D_{kl}^{sy}(M(P)||M(Q)) \leq 2(e^{\epsilon} - 1)^2\|P - Q\|_{TV}^2$ for the ϵ-LDP mechanism M where D_{kl}^{sy} is defined in (2). This result indicates the contraction property of the ϵ-LDP mechanism. By establishing such contraction inequality for a mechanism, we can employ Le Cam's or Assouad's method to derive the minimax lower bound for a given estimation problem. Therefore, considering the contraction inequality of f-FLDP mechanisms with respect to divergence measures can enhance our understanding of private minimax risk. The contraction inequality for the f-FLDP mechanism, which is also utilized in the proofs of Theorem 1 and Theorem 3, is presented in Theorem 2 in the main paper.

For a given locally private mechanism $M : \mathcal{X} \longrightarrow \mathcal{P}(\mathcal{Z})$ and distributions $P_1, P_2 \in \mathcal{P}(\mathcal{X})$, denote m_1 and m_2 for the distributions of $M(X)$ and $M(Y)$, respectively, where $X \sim P_1$ and $Y \sim P_2$. Thus, the mechanism discussed in this section is designed for a dataset of size 1.

Recall the definition of $\delta_f(y) = \sup_{x\in[0,1]} 1 - yx - f(x)$. For given two trade-off functions f and g, if $f \geq g$ on $[0, 1]$, then $\delta_f(y) \leq \delta_g(y)$ for all y. Thus, $c_{f,\kappa} \leq c_{g,\kappa}$ if $f \geq g$ on $[0, 1]$. This implies that with less private mechanism ($g \leq f$), its resulting distribution is less contracted ($c_{g,\kappa} \geq c_{f,\kappa}$), which aligns with our intuitive expectation. We state the result of Duchi et al. (2018) for comparison.

Proposition S3 (Contraction in ϵ-LDP mechanism, Duchi et al. (2018))**.** *For the ϵ-LDP mechanism,*

$$D_{kl}^{sy}(m_1||m_2) \leq 2(e^{\epsilon} - 1)^2\|P_1 - P_2\|_{TV}^2$$

holds for any $P_1, P_2 \in \mathcal{P}(\mathcal{X})$.

Our bound extends the result presented in Proposition S3. If $c_{f,\kappa}$ exists for $\kappa = 1$ and a given trade-off function f, then Theorem 2 bounds $D_{kl}(m_1||m_2)$

by $\|P_1 - P_2\|_{TV}^2$. This implies that the contraction and minimax rate for inference problems under f-FLDP are analogous to those of ϵ-LDP, as shown in Theorem 1,Theorem 3 and Theorem 2. Combining Theorem 2 and Lemma 1, we obtain the following contraction inequality.

Corollary S1. *If a given trade-off function $f(x) \geq 1 - c_0 x^{c_1}$ on $[0,1]$ for some $c_0 > 0$ and $c_1 \in (0.5, 1)$, or $\delta_f(x) \leq c_3 x^{-1-c_2}$ for $x > x_0$ and some $c_3, c_2, x_0 > 0$, then $\int_0^\infty \delta_f(t)\, dt$ converges, and*

$$D_{kl}^{sy}(m_1||m_2) \leq c_{f,1} \frac{\|P_1 - P_2\|_{TV}^2}{1 - \|P_1 - P_2\|_{TV}}$$

holds for f-FLDP mechanim M and for every distributions P_1 and P_2 over $\mathcal{X}$.

Combining with the fact that G_μ satisfies the condition of Corollary S1, we can also state the contraction inequality for μ-GLDP.

Corollary S2. *For μ-GLDP mechanism M, the following holds:*

$$D_{kl}^{sy}(m_1||m_2) \leq c \frac{\|P_1 - P_2\|_{TV}^2}{1 - \|P_1 - P_2\|_{TV}}$$

for every distributions P_1 and P_2 over $\mathcal{X}$.

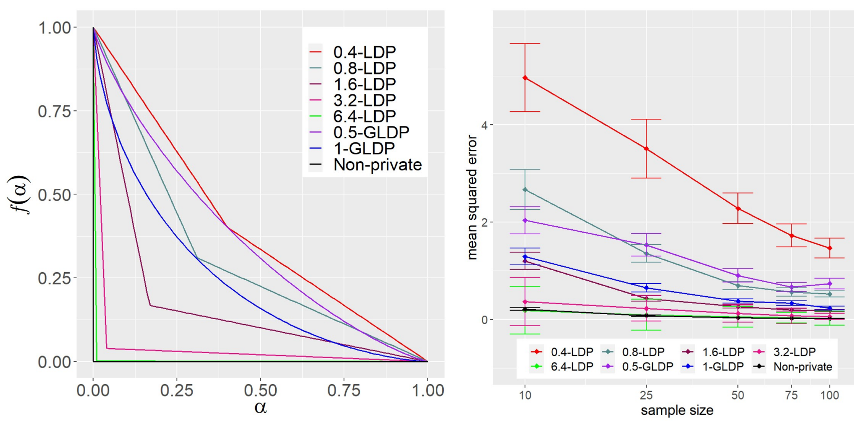
Recall that the minimax rates of μ-GLDP are equivalent to those of ϵ-LDP with respect to n. There exists a mechanism that is μ-GLDP but not ϵ-LDP for any $\epsilon \geq 0$. Hence, μ-GLDP offers a more lenient privacy guarantee, yet its optimal statistical utility does not differ from that of ϵ-LDP. This phenomenon can be explained using the concept of contraction. The contraction inequality of the μ-GLDP mechanism is not significantly different from that of ϵ-LDP, as demonstrated in Proposition S3 and Corollary S2. Additionally, we note another result regarding the contraction of the mechanism. Asoodeh et al. (2021) showed $D_f(m_1||m_2) \leq (1 - (1-\delta)e^{-\epsilon}) D_f(P_1||P_2)$ for a (ϵ, δ)-DP mechanism where D_f represents an f-divergence. This result provides an intuitive understanding of the contraction capability of the mechanism. However, as they pointed out, the bounds in their work are loose in exchange for their generality. In comparison, our result is more powerful when considering the K-L divergence, as our bound can provide a finite upper bound even for distributions P_1 and P_2 with diverging K-L divergence. Furthermore, our inequality allows for the classification of private mechanisms based on their characteristics. For example, consider Gaussian mechanisms with different variances of noise. With a fixed ϵ, these mechanisms satisfy (ϵ, δ_1)-LDP and (ϵ, δ_2)-LDP for different values of δ_1 and δ_2. While they may be treated differently from the perspective of (ϵ, δ)-LDP, they exhibit essentially the same contraction ability according to (2), as the contraction is bounded by the square of the total variation.

4 Additional experiments

For our experiments, we utilized a CPU with an 11th Gen Intel(R) Core(TM) i7-1185G7 processor. Also, we employed the Laplace random number generator proposed by Chen (2023) for private mean estimation.

4.1 Univariate mean estimation

This experiment compares the univariate mean estimation for ϵ-LDP, μ-GLDP, and non-private mechanisms. Our proposed mechanism in Section 3.1 under μ-GLDP is compared with the mechanism in Duchi et al. (2018) and the non-private mechanism, which is obtained by the sample mean. We generate the data from $\chi^2(1)$. We set $\epsilon \in \{0.4, 0.8, 1.6, 3.2, 6.4\}$ for ϵ-LDP and $\mu \in \{0.5, 1\}$ for μ-GLDP. Also, we vary the sample size from $n = 10$ to $n = 100$ and the moment from k=2 to $k = 60$. Estimation errors are calculated by taking the square of the difference between the true and estimated means. We repeat the experiment 100 times for each mechanism.



(a) Trade-off functions

(b) Mean squared errors

Figure S1: (a) Trade-off functions of each ϵ-LDP for $\epsilon \in \{0.4, 0.8, 1.6, 3.2, 6.4\}$ and μ-GLDP for $\mu \in \{0.5, 1\}$. (b) The mean squared errors for univariate mean estimation with the best assumed moment k for each privacy mechanism.

Fig. S1a presents a graphical comparison of the trade-off function for each assumed local differential privacy mechanism. In the mean estimation, we observe that the higher-utility-less-private rule remains consistent across all mechanisms, unlike in the density estimation, as can be seen in Fig. S1b. This implies that if an optimal mechanism is partially more private than the others (exhibiting smaller trade-off function values for some given Type I errors), its estimation errors are worse, even when selecting the kth moment that yields the smallest error for each mechanism, except for $n = 10$.

5 Proofs

Notation. For a given mechanism $M : \mathcal{X} \longrightarrow \mathcal{P}(\mathcal{Z})$ we denote $q_M(z)$ as the density of the random variable $M(x)$ for $x \in \mathcal{X}$. When the specific mechanism being referenced is clear, we omit the subscript and use $q(z|x)$. Also, for a distribution P on $\mathcal{X}$, we denote $M(P)$ as the distribution of $M(X)$ with $X \sim P$. The density of $M(P)$ can be expressed as $\int_{\mathcal{X}} q(z|x)dP(x)$.

Proposition S4 (Asoodeh et al. (2021))**.** *If a mechanism $M : \mathcal{X} \longrightarrow \mathcal{Z}$ is (ϵ, δ)-DP, then*

$$E_{e^\epsilon}\left[M(P)||M(Q)\right] \leq \delta E_{e^\epsilon}\left[P||Q\right] \leq \delta$$

holds for every distributions P and Q over $\mathcal{X}$.

5.1 Proof of Theorem 2

We prove the following lemma first before proving Theorem 2.

Lemma S1.

$$\log(1+t) \leq \left(\frac{1-\kappa}{\kappa}\right)^{1-\kappa} t^\kappa$$

holds for $t > 0$ and $0 \leq \kappa \leq 1$.

Proof. The generalized AM-GM inequality states that

$$x^\kappa y^{1-\kappa} \leq \kappa x + (1-\kappa)y$$

for $x, y \geq 0$ and $\kappa \in [0,1]$. Thus, for $\kappa \in [0,1]$ and $t \geq 0$,

$$\left(\frac{1}{\kappa}\frac{1}{1+t}\right)^\kappa \left(\frac{1}{1-\kappa}\frac{t}{1+t}\right)^{1-\kappa} \leq 1$$

holds by the generalized AM-GM inequality. Then,

$$\frac{d}{dt}\left(\left(\frac{1-\kappa}{\kappa}\right)^{1-\kappa} t^\kappa - \log(1+t)\right) = (1-\kappa)^{1-\kappa}\kappa^\kappa t^{\kappa-1} - \frac{1}{1+t} \geq 0.$$

Therefore, the inequality holds for $t \geq 0$. □

Note that

$$\begin{aligned}
D_{kl}^{sy}(m_1||m_2) &= \int_{\mathcal{Z}} m_1 \log\frac{m_1}{m_2}dz + \int_{\mathcal{Z}} m_2 \log\frac{m_2}{m_1}dz \\
&= \int_{\mathcal{Z}} (m_1 - m_2)\log\frac{m_1}{m_2}dz \\
&= \int_{\mathcal{Z}} (\max\{m_1, m_2\} - \min\{m_1, m_2\})\log\frac{\max\{m_1, m_2\}}{\min\{m_1, m_2\}}dz \\
&= \int_{\mathcal{Z}} |m_1 - m_2| \log\left(1 + \frac{|m_1 - m_2|}{\min\{m_1, m_2\}}\right)dz.
\end{aligned}$$

Then, by Lemma S1,

$$D_{kl}^{sy}(m_1||m_2) \le \left(\frac{1-\kappa}{\kappa}\right)^{1-\kappa} \int_{\mathcal{Z}} \frac{|m_1-m_2|^{1+\kappa}}{\min\{m_1,m_2\}^{\kappa}} dz.$$

Define distributions P_+ and P_- following the distributions $\frac{(P_1-P_2)_+}{\|P_1-P_2\|_{TV}}$ and $\frac{(P_2-P_1)_+}{\|P_1-P_2\|_{TV}}$, respectively, where $(x)_+ = \max\{x,0\}$. Note that P_+ and P_- are distributions since their densities are non-negative and integrals are 1 by definition. Then, we can derive the densities of $m_+ = M(P_+)$ and $m_- = M(P_-)$ as follows:

$$\begin{aligned} m_+(z) &= \frac{1}{\|P_1-P_2\|_{TV}} \int_{\mathcal{X}} q(z|x) \left(dP_1(x) - dP_2(x)\right)_+, \\ m_-(z) &= \frac{1}{\|P_1-P_2\|_{TV}} \int_{\mathcal{X}} q(z|x) \left(dP_2(x) - dP_1(x)\right)_+. \end{aligned}$$

Thus, we have

$$\begin{aligned} m_+(z) - m_-(z) &= \frac{1}{\|P_1-P_2\|_{TV}} \int_{\mathcal{X}} q(z|x) \left(dP_1(x) - dP_2(x)\right) \\ &= \frac{m_1(z)-m_2(z)}{\|P_1-P_2\|_{TV}}, \end{aligned}$$

since $\max\{x,0\} - \max\{-x,0\} = x$. Then,

$$\begin{aligned} D_{kl}^{sy}(m_1||m_2) &\le \left(\frac{1-\kappa}{\kappa}\right)^{1-\kappa} \|P_1-P_2\|_{TV}^{1+\kappa} \int_{\mathcal{Z}} \frac{|m_+-m_-|^{1+\kappa}}{\min\{m_1,m_2\}^{\kappa}} dz \\ &= \left(\frac{1-\kappa}{\kappa}\right)^{1-\kappa} \|P_1-P_2\|_{TV}^{1+\kappa} \left(\int_{\mathcal{Z}} \frac{m_+^{1+\kappa} + m_-^{1+\kappa}}{\min\{m_1,m_2\}^{\kappa}} dz\right). \end{aligned}$$

Also, if we define $m_{\min} = M(P_{\min})$ where $P_{\min}$ is the distribution defined by $\frac{\min\{P_1,P_2\}}{\int_{\mathcal{X}} \min\{dP_1,dP_2\}}$, we have

$$\begin{aligned} m_{\min}(z) &= \frac{1}{\int_{\mathcal{X}} \min\{dP_1,dP_2\}} \int_{\mathcal{X}} q(z|x) \min\{dP_1,dP_2\} \\ &\le \frac{1}{\int_{\mathcal{X}} \min\{dP_1,dP_2\}} \min\left\{\int_{\mathcal{X}} q(z|x) dP_1, \int_{\mathcal{X}} q(z|x) dP_2\right\} \\ &= \frac{1}{1-\|P_1-P_2\|_{TV}} \min\{m_1,m_2\}. \end{aligned}$$

Hence,

$$D_{kl}^{sy}(m_1||m_2) \le \left(\frac{1-\kappa}{\kappa}\right)^{1-\kappa} \frac{\|P_1-P_2\|_{TV}^{1+\kappa}}{(1-\|P_1-P_2\|_{TV})^{\kappa}} \int_{\mathcal{Z}} \frac{m_+^{1+\kappa} + m_-^{1+\kappa}}{m_{\min}^{\kappa}} dz.$$

For each term in the right-hand side, we can obtain the following inequality:

$$\begin{aligned}\int_{\mathcal{Z}} \frac{m_+^{1+\kappa}}{m_{\min}^{\kappa}} dz &= \int_{\mathcal{Z}} \int_0^{\infty} (1+\kappa)\, t^{\kappa} 1\{\frac{m_+}{m_{\min}} > t\} m_{\min} dt dz \\ &= \int_{\mathcal{Z}} \int_0^{\infty} \kappa t^{\kappa-1} 1\{\frac{m_+}{m_{\min}} > t\} m_+ dt dz,\end{aligned}$$

which yields the following bound:

$$\begin{aligned}\int_{\mathcal{Z}} \frac{m_+^{1+\kappa}}{m_{\min}^{\kappa}} dz &= (1+\kappa) \int_{\mathcal{Z}} \frac{m_+^{1+\kappa}}{m_{\min}^{\kappa}} dz - \kappa \int_{\mathcal{Z}} \frac{m_+^{1+\kappa}}{m_{\min}^{\kappa}} dz \\ &= \int_{\mathcal{Z}} \int_0^{\infty} \kappa(1+\kappa) t^{\kappa-1} 1\{\frac{m_+}{m_{\min}} > t\} \left(m_+ - t m_{\min}\right) dt dz \\ &= \int_0^{\infty} \kappa(1+\kappa) t^{\kappa-1} \mathbb{E}_t \left[m_+ || m_{\min}\right] dt \\ &\leq \int_0^{\infty} \kappa(1+\kappa) t^{\kappa-1} \delta(t) dt.\end{aligned}$$

Then, the integral converges by the assumption of Theorem 2. Also, the inequality symmetrically holds for $\frac{m_-^{1+\kappa}}{m_{\min}^{\kappa}}$. Therefore, we obtain the desired inequality:

$$D_{kl}^{sy}(m_1 || m_2) \leq 2\kappa^{\kappa} \left(1-\kappa\right)^{1-\kappa} \frac{\|P_1 - P_2\|_{TV}^{1+\kappa}}{\left(1 - \|P_1 - P_2\|_{TV}\right)^{\kappa}} \int_0^{\infty} (1+\kappa) t^{\kappa-1} \delta(t) dt.$$

5.2 Proof of Corollary S1

First, we consider the case when a trade-off function $f(x) \geq 1 - c_0 x^{c_1}$ for some $c_0 > 0$ and $c_1 \in (0.5, 1)$. Recall the definition of $\delta(y) := \sup_{x \in [0,1]} 1 - yx - f(x)$. Thus, if we denote

$$\delta_f(y) = \sup_{x \in [0,1]} 1 - yx - f(x), \quad \delta_g(y) = \sup_{x \in [0,1]} 1 - yx - g(x)$$

for trade-off functions f and g such that $f(t) \geq g(t)$ on $[0,1]$, respectively, then

$$\delta_g(y) = \sup_{x \in [0,1]} 1 - yx - g(x) \geq \sup_{x \in [0,1]} 1 - yx - f(x) = \delta_f(y).$$

Now assume that a given trade-off function f satisfies $f(x) \geq 1 - c_0 x^{c_1}$ for some $c_0 > 0$ and $c_1 \in (0.5, 1)$ and define $g(x) = \max\{1 - c_0 x^{c_1}, 0\}$. Then, $\delta_f(y) \leq \delta_g(y)$. Thus, if $\int \delta_g(t)\, dt$ converges, $\int \delta_f(t)\, dt$ also converges. By performing specific calculations of δ_g, we can obtain the desired conclusion:

$$\delta_g(y) = \sup_{x \in [0,1]} 1 - yx - g(x) = \sup_{0 \leq x \leq c_0^{-\frac{1}{c_1}}} -yx + c_0 x^{c_1}$$

for $y \geq 0$. By differentiating the objective function, we can determine that it has a local maximum in $[0,1]$ at $x = (c_0 c_1)^{\frac{1}{1-c_1}} y^{-\frac{1}{1-c_1}}$, and their boundary values are non-positive on $[0, c_0^{-1/c_1}]$. For sufficiently large $y \geq y_0$, $(c_0 c_1)^{\frac{1}{1-c_1}} y^{-\frac{1}{1-c_1}} \leq c_0^{-1/c_1}$ and

$$\delta_g(y) = C y^{-\frac{c_1}{1-c_1}},$$

where $C = (c_0 c_1)^{\frac{1}{1-c_1}} (c_1^{-1} - 1)$. Note that $\delta_g(y) \in [0,1]$ by definition, which gives

$$\int_0^{y_0} \delta_g(y) dy \leq y_0.$$

Thus, we can show the convergence of the integral

$$\begin{aligned} \int_0^{\infty} \delta_g(y) dy &\leq y_0 + \int_{y_0}^{\infty} C y^{-\frac{c_1}{1-c_1}} dy \\ &= y_0 + C \frac{1-c_1}{1-2c_1} y_0^{-\frac{2c_1-1}{1-c_1}} \\ &< \infty \end{aligned}$$

if $c_1 \in (0.5, 1)$. Now we consider the case when $\delta(x) \leq c_3 x^{-1-c_2}$ for $x > x_0$ for some $c_3, c_2, x_0 > 0$. Since $\delta(y) \in [0,1]$, we have

$$\int_0^1 \delta(y) dy \leq 1$$

and

$$\int_1^{\infty} \delta(y) dy \leq \int_1^{\infty} c_3 y^{-1-c_2} dy = \frac{c_3}{c_2},$$

thus the integral converges and Theorem 2 yields the desired result.

5.3 Proof of GLDP holding the condition of Lemma 1

Dong et al. (2022) showed that if a trade-off function is given as $f(x) = G_\mu(x)$ for some $\mu > 0$, the corresponding $\delta(y) := \sup_{x \in [0,1]} 1 - yx - f(x)$ has the following closed form:

$$\delta\left(e^{\epsilon}\right) = \Phi\left(-\frac{\epsilon}{\mu} + \frac{\mu}{2}\right) - e^{\epsilon} \Phi\left(-\frac{\epsilon}{\mu} - \frac{\mu}{2}\right).$$

Denote $y = e^{\epsilon}$ and note that

$$\Phi(-t) = \int_{-\infty}^{-t} \frac{1}{\sqrt{2\pi}} e^{-x^2} dx \leq \int_{-\infty}^{-t} \frac{|x|}{|t|\sqrt{2\pi}} e^{-x^2} dx = \frac{1}{2t\sqrt{2\pi}} e^{-t^2}.$$

Then,

$$\begin{aligned}
\delta(y) &\leq \Phi\left(-\frac{\log y}{\mu}+\frac{\mu}{2}\right) \leq \frac{1}{2\left(\frac{\log y}{\mu}-\frac{\mu}{2}\right)\sqrt{2\pi}} e^{-\left(\frac{\log y}{\mu}-\frac{\mu}{2}\right)^2} \\
&\leq \frac{1}{4\sqrt{2\pi}} e^{-\frac{\log^2 y}{\mu^2}+\log y} \\
&= \frac{1}{4\sqrt{2\pi}} y^{-\frac{\log y}{\mu}+1} \\
&< \frac{1}{4\sqrt{2\pi}} y^{-2}
\end{aligned}$$

for $y > e^{3\mu}$. Therefore, $\int \delta(t)dt$ converges by Lemma 1.

5.4 Proof of Theorem 1

For a given mechanism M, denote the estimator $\hat{\theta}(Z) = \hat{\theta}\left(M(X_1, \ldots, X_n)\right)$. Then,

$$\begin{aligned}
\sup_{P\in\mathcal{P}} \mathbb{E}\left[\Phi\left(\rho\left(\hat{\theta}, \theta(P)\right)\right)\right] &\geq \frac{1}{2}\mathbb{E}\left[\Phi\left(\rho\left(\hat{\theta}, \theta(P_1)\right)\right)|Z \sim M(P_1)\right] \\
&+ \frac{1}{2}\mathbb{E}\left[\Phi\left(\rho\left(\hat{\theta}, \theta(P_2)\right)\right)|Z \sim M(P_2)\right]. \qquad \text{(S1)}
\end{aligned}$$

Since ρ is a metric,

$$\rho\left(\hat{\theta}, \theta(P_1)\right) + \rho\left(\hat{\theta}, \theta(P_2)\right) \geq \rho\left(\theta(P_1), \theta(P_2)\right) \geq 2\eta.$$

Thus, at least one of the terms in the left-hand side is greater or equal to η. Since Φ is increasing, we have

$$\Phi\left(\rho\left(\hat{\theta}, \theta(P_1)\right)\right) + \Phi\left(\rho\left(\hat{\theta}, \theta(P_2)\right)\right) \geq \Phi(\eta).$$

For $X = (X_1, \ldots, X_n) \in \mathcal{X}^n$, let $L_i(Z) = \Phi\left(\rho\left(\hat{\theta}(Z), \theta(P_i)\right)\right)$ for i =1 and 2, $w_1(Z) = \frac{L_1(Z)}{L_1(Z)+L_2(Z)}$, and $w_2(Z) = \frac{L_2(Z)}{L_1(Z)+L_2(Z)}$. Then, (S1) becomes

$$\begin{aligned}
\sup_{P\in\mathcal{P}} \mathbb{E}\left[\Phi\left(\rho\left(\hat{\theta}, \theta(P)\right)\right)\right] &\geq \frac{\Phi(\eta)}{2}\left(\int_{\mathcal{Z}^n} w_1 dM(P_1^n) + \int_{\mathcal{Z}^n} w_2 dM(P_2^n)\right) \\
&\geq \frac{\Phi(\eta)}{2}\int_{\mathcal{Z}^n} \min\{dM(P_1^n), dM(P_2^n)\} \\
&= \frac{\Phi(\eta)}{2}\int_{\mathcal{Z}^n} dM(P_1^n) - (dM(P_1^n) - dM(P_2^n))_+ \\
&= \frac{\Phi(\eta)}{2}\left(1 - \|M(P_1^n) - M(P_2^n)\|_{TV}\right).
\end{aligned}$$

Pinsker's inequality (Tsybakov, 2009) states that

$$\|P-Q\|_{TV}^2 \leq \frac{1}{2} D_{kl}\left(P\|Q\right). \tag{S2}$$

Using Pinsker's inequality in (S2), we have

$$\|M(P_1^n)-M(P_2^n)\|_{TV} \leq \sqrt{\frac{1}{2} D_{kl}\left(M(P_1^n)\|M(P_2^n)\right)}. \tag{S3}$$

Note that we can decompose the K-L divergence for multivariate distributions as:

$$\begin{aligned}
& D_{kl}\left(f(x_1,\cdots,x_n)\|g(x_1,\cdots,x_n)\right) \\
= & \int_{\mathcal{X}^n} f(x_{1:n}) \log \frac{f(x_{1:n})}{g(x_{1:n})} dx_{1:n} \\
= & \int_{\mathcal{X}^n} f(x_{1:n}) \sum_{i=1}^n \log \frac{f(x_i|x_{1:i-1})}{g(x_i|x_{1:i-1})} dx_{1:n} \\
= & \sum_{i=1}^n \int_{\mathcal{X}^n} f(x_{1:n}) \log \frac{f(x_i|x_{1:i-1})}{g(x_i|x_{1:i-1})} dx_{1:n} \\
= & \sum_{i=1}^n \int_{\mathcal{X}^i} f(x_{1:i}) \log \frac{f(x_i|x_{1:i-1})}{g(x_i|x_{1:i-1})} dx_{1:i} \\
= & \sum_{i=1}^n \int_{\mathcal{X}^{i-1}} \int_{\mathcal{X}} f(x_i|x_{1:i-1}) \log \frac{f(x_i|x_{1:i-1})}{g(x_i|x_{1:i-1})} dx_i f(x_{1:i-1}) dx_{1:i-1} \\
= & \sum_{i=1}^n \int_{\mathcal{X}^{i-1}} D_{kl}\left(f(x_i|x_{1:i-1})\|g(x_i|x_{1:i-1})\right) f(x_{1:i-1}) dx_{1:i-1}
\end{aligned}$$

where $f(x_1,\ldots,x_i)$ and $f(x_i|x_{1:i-1})$ denote the marginal distribution of $(X_1,\ldots,X_i)$ and conditional distribution of $X_i|X_1,\ldots,X_{i-1}$, respectively, under the distribution with density $f(x_1,\cdots,x_n)$. Hence, the K-L divergence is decomposed into the summation of the expected K-L divergences:

$$D_{kl}\left(f(x_{1:n})\|g(x_{1:n})\right) = \sum_{i=1}^n \mathbb{E}\left[D_{kl}\left(f(x_i|X_{1:i-1})\|g(x_i|X_{1:i-1})\right)|X_{1:n}\sim f\right]. \tag{S4}$$

Denote $(Z_1,\ldots,Z_n)=M(P_1^n),(W_1,\ldots,W_n)=M(P_2^n)$, then Eq. (S4) gives

$$D_{kl}\left(M(P_1^n)\|M(P_2^n)\right) = \sum_{i=1}^n \mathbb{E}\left[D_{kl}\left(z_i|Z_{1:i-1}\|w_i|Z_{1:i-1}\right)|Z_{1:n}\sim M(P_1^n)\right].$$

The locally private mechanism is f-FLDP if, for given $i-1$ outputs, the mechanism $M(X_i|z_1,\ldots,z_{i-1})$ is f-FDP for X_i, which is also the case of f-FLDP with only one input. As a result, $D_{kl}\left(z_i|Z_{1:i-1}\|w_i|Z_{1:i-1}\right)$ is bounded by

$c_{f,\kappa}\frac{\|P_1-P_2\|_{TV}^{1+\kappa}}{(1-\|P_1-P_2\|_{TV})^{\kappa}}$ based on Theorem 2 since $z_i|Z_{1:i-1}\sim M(P_1)$ and $w_i|Z_{1:i-1}\sim M(P_2)$, which leads to the following inequality:

$$D_{kl}\left(M(P_1^n)||M(P_2^n)\right)\le nc_{f,\kappa}\frac{\|P_1-P_2\|_{TV}^{1+\kappa}}{(1-\|P_1-P_2\|_{TV})^{\kappa}}.$$

Combining it with Eq. (S3), we have

$$\|M(P_1^n)-M(P_2^n)\|_{TV}\le\sqrt{\frac{1}{2}nc_{f,\kappa}\frac{\|P_1-P_2\|_{TV}^{1+\kappa}}{(1-\|P_1-P_2\|_{TV})^{\kappa}}}$$

and

$$\sup_{P\in\mathcal{P}}\mathbb{E}\left[\Phi\left(\rho\left(\hat{\theta},\theta\left(P\right)\right)\right)\right]\ge\frac{\Phi\left(\eta\right)}{2}\left(1-\sqrt{\frac{1}{2}nc_{f,\kappa}\frac{\|P_1-P_2\|_{TV}^{1+\kappa}}{(1-\|P_1-P_2\|_{TV})^{\kappa}}}\right).$$

Consequently, for a given family $\mathcal{M}_f$ of f-FLDP mechanisms, we have

$$\mathcal{R}_n\left(\theta\left(\mathcal{P}\right),\Phi\circ\rho,\mathcal{M}_f\right)\ge\frac{\Phi\left(\eta\right)}{2}\left(1-\sqrt{nc_{f,\kappa}\frac{\|P_1-P_2\|_{TV}^{1+\kappa}}{(1-\|P_1-P_2\|_{TV})^{\kappa}}}\right).$$

5.5 Proof of Corollary 1

For $0<\eta<1$ and $v\in\{-1,+1\}$, define distributions P_v such that

$$\mathbb{P}_v\left(X=\pm\eta^{-\frac{1}{k}}\right)=\frac{1\pm v}{2}\eta,\quad\mathbb{P}_v\left(X=0\right)=1-\eta.$$

Then,

$$\mathbb{E}_v[X]=v\eta^{1-\frac{1}{k}},\quad\mathbb{E}_v[|X|^k]=1,$$

thus $P_v\in\mathcal{P}_k$. Noting $|\mathbb{E}_{+1}[X]-\mathbb{E}_{-1}[X]|=2\eta^{1-\frac{1}{k}}$, we can apply Theorem 1:

$$\mathcal{R}_n\left(\theta\left(\mathcal{P}_k,\|\cdot\|_2^2,\mathcal{M}_f\right)\right)\ge\frac{1}{2}\left(\eta^{1-\frac{1}{k}}\right)^2\left[1-\sqrt{nc_{f,\kappa}\frac{\|P_{+1}-P_{-1}\|_{TV}^{1+\kappa}}{(1-\|P_{+1}-P_{-1}\|_{TV})^{\kappa}}}\right].$$

Putting $\|P_{+1}-P_{-1}\|_{TV}=\eta$, we have

$$\mathcal{R}_n\left(\theta\left(\mathcal{P}_k,\|\cdot\|_2^2,\mathcal{M}_f\right)\right)\ge\frac{1}{2}\left(\eta^{1-\frac{1}{k}}\right)^2\left[1-\sqrt{nc_{f,\kappa}\frac{\eta^{1+\kappa}}{(1-\eta)^{\kappa}}}\right].$$

Set $\eta=(2nc_{f,\kappa})^{-\frac{1}{1+\kappa}}$, then

$$(1-\eta)^{-\kappa}=\left(1-(2nc_{f,\kappa})^{-\frac{1}{1+\kappa}}\right)^{-\kappa}\le\left(1-\frac{1}{\sqrt{2nc_{f,\kappa}}}\right)^{-1}\le\frac{2\sqrt{2}}{2\sqrt{2}-1}$$

if $n > \frac{4}{c_{f,\kappa}}$. As a result,

$$\mathcal{R}_n\left(\theta\left(\mathcal{P}_k, \|\cdot\|_2^2, \mathcal{M}_f\right)\right) \geq \frac{1}{2}\left(2nc_{f,\kappa}\right)^{-\frac{1}{1+\kappa}\frac{2(k-1)}{k}}\left[1-\sqrt{\frac{\sqrt{2}}{2\sqrt{2}-1}}\right].$$

Therefore,

$$\mathcal{R}_n\left(\theta\left(\mathcal{P}_k, \|\cdot\|_2^2, \mathcal{M}_f\right)\right) \geq c_0\left(nc_{f,\kappa}\right)^{-\frac{1}{1+\kappa}\frac{2(k-1)}{k}}$$

if $n > c_1$ where $c_0 > 0$ depends on k and κ and c_1 depends on $c_{f,\kappa}$.

5.6 Proof of Corollary 2

The lower bound is evident from the fact that G_μ satisfies the conditions stated in Lemma 1 and Corollary 1. The coefficient in the lower bound is given as follows:

$$\begin{aligned}
c_{G_\mu,1} &= 4\int_0^\infty \delta_{G_\mu}(t)dt \\
&= 4\int_0^\infty \Phi\left(-\frac{\log t}{\mu}+\frac{\mu}{2}\right) - t\Phi\left(-\frac{\log t}{\mu}-\frac{\mu}{2}\right)dt \\
&= 4\int_0^\infty \int_{-\infty}^{-\frac{\log t}{\mu}+\frac{\mu}{2}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx - t\int_{-\infty}^{-\frac{\log t}{\mu}-\frac{\mu}{2}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dxdt \\
&= 4\int_0^\infty \int_{-\infty}^{-\frac{\log t}{\mu}+\frac{\mu}{2}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\left(1-te^{\mu x-\frac{1}{2}\mu^2}\right)dxdt \\
&= 4\int_{-\infty}^\infty \int_0^{e^{-\mu x+\frac{1}{2}\mu^2}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\left(1-te^{\mu x-\frac{1}{2}\mu^2}\right)dtdx \\
&= 4\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\frac{1}{2}e^{-\mu x+\frac{1}{2}\mu^2}dx \\
&= 2\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x+\mu)^2}e^{\frac{1}{2}\mu^2}dx \\
&= 2e^{\frac{1}{2}\mu^2}
\end{aligned}$$

To establish the upper bound, we provide an explicit mechanism and estimator that resemble those of ϵ-LDP. For $T > 0$, define

$$M(x) = \max\{-T, \min\{x, T\}\} + \epsilon_i$$

where $\epsilon_i \sim N\left(0, \frac{4T^2}{\mu^2}\right)$. It is μ-GDP since

$$T(x+\epsilon_i, x'+\epsilon_i) = T(x-x'+\epsilon_i, \epsilon_i) = T\left(\frac{\mu}{2T}(x-x') + N(0,1), N(0,1)\right),$$

and $\left|\frac{\mu}{2T}(x-x')\right| \leq \mu$. Thus, $T(M(x), M(x')) \geq T(N(\mu,1), N(0,1)) = G_\mu$ and the mechanism is μ-GDP. For given data $X_1, \ldots, X_n$, define a locally private mechanism such that $Z_i = M(X_i)$ for all i, which is μ-GLDP. Under the μ-GLDP mechanism, we can evaluate the estimator $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Z_i$. Then, we have

$$\mathbb{E}\left[\left(\hat{\theta} - \theta\right)^2\right] = Var(\hat{\theta}) + \left(\mathbb{E}[Z_1] - \mathbb{E}\left[X_1\right]\right)^2.$$

We can bound the first term in the right-hand side as follows:

$$\begin{aligned} Var(\hat{\theta}) &= \frac{4T^2}{n\mu^2} + \frac{1}{n}\mathbb{E}\left[\max\{-T, \min\{x, T\}\}^2\right] - \frac{1}{n}\mathbb{E}\left[\max\{-T, \min\{x, T\}\}\right]^2 \\ &\leq \frac{4T^2}{n\mu^2} + \frac{1}{n}T^2. \end{aligned}$$

For the second term, we use the fact that the bounded kth moment implies a tail bounded by the kth order, and we have

$$\begin{aligned} |\mathbb{E}[Z_1] - \mathbb{E}\left[X_1\right]| &= |\mathbb{E}\left[\max\{-T, \min\{X_1, T\}\} - X_1\right]| \\ &\leq \mathbb{E}\left[|X_1|\, 1\{|X_1| > T\}\right] \\ &\leq \mathbb{E}[|X_1|^k]^{\frac{1}{k}}\,\mathbb{E}\left[1\{|X_1| > T\}\right]^{1-\frac{1}{k}} \end{aligned}$$

by Hölder's inequality. Thus,

$$|\mathbb{E}[Z_1] - \mathbb{E}\left[X_1\right]| \leq \mathbb{P}\left(|X_1| > T\right)^{1-\frac{1}{k}} = \mathbb{P}\left(|X_1|^k > T^k\right)^{1-\frac{1}{k}} \leq \frac{1}{T^{k-1}}$$

by Markov's inequality since $\mathbb{E}\left[|X_1|^k\right] \leq 1$. As a result,

$$\mathbb{E}\left[\left(\hat{\theta} - \theta\right)^2\right] \leq \left(1 + \frac{4}{\mu^2}\right)\frac{1}{n}T^2 + \frac{1}{T^{2k-2}}$$

and setting $T = \left[n\left(1 + \frac{4}{\mu^2}\right)^{-1}\right]^{\frac{1}{2k}}$, we have

$$\mathbb{E}\left[\left(\hat{\theta} - \theta\right)^2\right] \leq \left(1 + \frac{4}{\mu^2}\right)^{1-\frac{1}{k}} n^{-\left(1-\frac{1}{k}\right)}.$$

Therefore,

$$\mathcal{R}_n\left(\theta\left(\mathcal{P}_k\right), \Phi \circ \rho, M_\mu\right) \leq \left(n\mu^2(4 + \mu^2)^{-1}\right)^{-\left(1-\frac{1}{k}\right)}.$$

5.7 Proof of Theorem 3

For $\{P_v\}_{v\in\{-1,+1\}^d}$ and $V : \Theta \longrightarrow \{-1, +1\}^d$ satisfying 2η-Hamming separation for $\Phi \circ \rho$, we have

$$\mathcal{R}_n\left(\theta\left(\mathcal{P}\right), \Phi \circ \rho, M\right) \geq \eta \sum_{i=1}^{d} \left(1 - \|M(P_{+j}^n) - M(P_{-j}^n)\|_{TV}\right)$$

by Proposition S2. The Pinsker's inequality in (S2) gives

$$\|M(P_{+j}^n) - M(P_{-j}^n)\|_{TV} \leq \sqrt{\frac{1}{2} D_{kl}\left(M(P_{+j}^n)||M(P_{-j}^n)\right)}.$$

Again, we can decompose the K-L divergence by Eq. (S4). Denote $m_{\pm j}(z_1, \cdots, z_n)$ for the distribution of $(Z_1, \ldots, Z_n) = M\left(P_{\pm j}^n\right)$, then the chain rule gives

$$D_{kl}\left(m_{+j}||m_{-j}\right) = \sum_{i=1}^{n} \mathbb{E}\left[D_{kl}\left(m_{+j}(x_i|Z_{1:i-1})||m_{-j}(x_i|Z_{1:i-1})\right)\right]$$

where $m_{\pm}(x_i|z_{1:i-1})$ is the distribution of $Z_i = M_i(x_i, z_{1:i-1})$ for given $z_{1:i-1}$.

Now we show that $m_{\pm j}(x_i|Z_{1:i-1}) = M(P_{\pm j})$, implying that X_i and $Z_{1:i-1}$ are independent. It was previously noted by Duchi et al. (2018) but we show it for the sake of completeness in the proof. Note that, if X_is are i.i.d. drawn from P_v, X_i and $Z_{1:i-1}$ are independent due to the sequential assumption of the locally private mechanism. Without loss of generality, we only need to show $m_{+j}(x_i|Z_{1:i-1}) = M(P_{+j})$. Let $(X_1, \ldots, X_n)$ be a random vector following P_{+j}^n and $(Z_1, \ldots, Z_n) = M(X_1, \ldots, X_n)$. In what follows, we use the induction on i. When $i = 1$, the statement is trivial. For $i > 1$, we have

$$\begin{aligned}
m_{+j}(x_i, z_{i-1}|z_{1:i-2}) &= \int_{\mathcal{X}} m_{+j}(x_i, z_{i-1}|x_{i-1}, z_{1:i-2}) dP_{+j}(x_{i-1}|z_{1:i-2}) \\
&= \int_{\mathcal{X}} m_{+j}(x_i, z_{i-1}|x_{i-1}, z_{1:i-2}) dP_{+j}(x_{i-1}) \\
&= \int_{\mathcal{X}} m_{+j}(x_i, z_{i-1}|x_{i-1}, z_{1:i-2}) p_{+j}(x_{i-1}) dx_{i-1}
\end{aligned}$$

by the induction hypothesis and the fact that the ith marginal of P_{+j}^n is P_{+j}, which leads to

$$m_{+j}(x_i, z_{i-1}|z_{1:i-2}) = \int_{\mathcal{X}} m_{+j}(x_i|x_{i-1}, z_{1:i-2}) q(z_{i-1}|x_{i-1}, z_{1:i-2}) dP_{+j}(x_{i-1}).$$

Also, X_i and X_{i-1} are independent since

$$\begin{aligned}
\mathbb{P}\left(X_i \in A, X_{i-1} \in B\right) &= \int_A \int_B \int_{\mathcal{X}^{n-2}} dP_{+j}^n \\
&= \frac{1}{2^{d-1}} \sum_{v:v_j=+1} \int_A \int_B \int_{\mathcal{X}^{n-2}} \prod_{i=1}^{n} p_v(x_i) d\mathbf{x} \\
&= \frac{1}{2^{d-1}} \sum_{v:v_j=+1} \mathbb{P}_v\left(X_i \in A\right) \mathbb{P}_v\left(X_{i-1} \in B\right) \\
&= \mathbb{P}_{+j}\left(X_i \in A\right) \mathbb{P}_{+j}\left(X_{i-1} \in B\right) \\
&= \mathbb{P}\left(X_i \in A | X_{1:n} \sim P_{+j}^n\right) \mathbb{P}\left(X_{i-1} \in B | X_{1:n} \sim P_{+j}^n\right).
\end{aligned}$$

Consequently, we have

$$\begin{aligned} m_{+j}(x_i, z_{i-1}|z_{1:i-2}) &= \int_{\mathcal{X}} m_{+j}(x_i|z_{1:i-2}) q(z_{i-1}|x_{i-1}, z_{1:i-2}) p_{+j}(x_{i-1}) dx_{i-1} \\ &= m_{+j}(x_i|z_{1:i-2}) m_{+j}(z_{i-1}|z_{1:i-2}). \end{aligned}$$

Repeating the procedure inductively, we arrive at the independency between X_i and $Z_{1:i-1}$. Thus, we have

$$D_{kl}\left(m_{+j}(x_i|Z_{1:i-1})||m_{-j}(x_i|Z_{1:i-1})\right) \leq c_{f,\kappa} \frac{\|P_{+j} - P_{-j}\|_{TV}^{1+\kappa}}{(1 - \|P_{+j} - P_{-j}\|_{TV})^{\kappa}}.$$

Therefore,

$$\|M(P_{+j}^n) - M(P_{-j}^n)\|_{TV} \leq \sqrt{\frac{nc_{f,\kappa}}{2} \frac{\|P_{+j} - P_{-j}\|_{TV}^{1+\kappa}}{(1 - \|P_{+j} - P_{-j}\|_{TV})^{\kappa}}}$$

for all $j = 1, \ldots, d$, which yields

$$\begin{aligned} \mathcal{R}_n\left(\theta\left(\mathcal{P}, \Phi \circ \rho, f\right)\right) &\geq d\eta - \sum_{j=1}^{d} \sqrt{\frac{nc_{f,\kappa}}{2} \frac{\|P_{+j} - P_{-j}\|_{TV}^{1+\kappa}}{(1 - \|P_{+j} - P_{-j}\|_{TV})^{\kappa}}} \\ &\geq d\eta - \sqrt{d \sum_{j=1}^{d} \frac{nc_{f,\kappa}}{2} \frac{\|P_{+j} - P_{-j}\|_{TV}^{1+\kappa}}{(1 - \|P_{+j} - P_{-j}\|_{TV})^{\kappa}}} \\ &\geq d\eta \left[1 - \sqrt{\frac{nc_{f,\kappa}}{d} \sum_{j=1}^{d} \frac{\|P_{+j} - P_{-j}\|_{TV}^{1+\kappa}}{(1 - \|P_{+j} - P_{-j}\|_{TV})^{\kappa}}}\right] \end{aligned}$$

by the Cauchy-Schwarz inquality.

5.8 Proof of Corollary 3

Duchi et al. (2018) showed that there exists a function $g_\beta : [0, 1] \longrightarrow \mathbb{R}$ such that g_β is β-times differentiable, non-negative on $[0, 1/2]$ and non-positive on $[1/2, 1]$ with $g^{(i)}(0) = g^{(i)}(1)$ for $i \leq \beta - 1$, $\int_0^1 g_\beta(x)dx = 0$, $|g_\beta^{(\beta)}(x)|$, and $|g_\beta(x)| \leq 1$. Finally,

$$\int_0^{1/2} g_\beta(x)dx = c_{1/2}, \quad \int_0^1 \left(g_\beta^{(i)}(x)\right)^2 dx \geq c$$

for some $c_{1/2}, c > 0$ for $i < \beta$. For $d \geq 1$ if we define $f_v = 1 + \sum_{j=1}^{d} r v_j g_{\beta,j}$ where $g_{\beta,j} = \frac{1}{d^\beta} g_\beta\left(d\left(x - \frac{j-1}{d}\right)\right) 1\left\{x \in \left[\frac{j-1}{d}, \frac{j}{d}\right]\right\}$ then,

$$f_v^{(j)}(0) = f_v^{(j)}(1) \quad \int_0^1 \left|f_v^{(\beta)}(x)\right|^2 dx \leq r^2$$

for $j \leq \beta - 1$. It is shown that such f_v is a member of $\mathcal{F}_\beta\left[\frac{r}{\pi^\beta}\right] \subset \mathcal{F}_\beta[r]$ in Tsybakov (2009). Define

$$V_j(f) := \arg\min_{v \in \{-1,+1\}} \int_{\frac{j-1}{d}}^{\frac{j}{d}} \left(f(x) - rvg_{\beta,j}(x)\right)^2 dx$$

for $j = 1, 2, \ldots, d$. Then, we have

$$\int_{\frac{j-1}{d}}^{\frac{j}{d}} \left(f(x) + rV_j(f)g_{\beta,j}(x)\right)^2 dx \geq \int_{\frac{j-1}{d}}^{\frac{j}{d}} \left(f(x) - rV_j(f)g_{\beta,j}(x)\right)^2 dx$$

which gives

$$\int_{\frac{j-1}{d}}^{\frac{j}{d}} f(x)rV_j(f)g_{\beta,j}(x)dx \geq 0,$$

and

$$\int_{\frac{j-1}{d}}^{\frac{j}{d}} \left(f(x) + rV_j(f)g_{\beta,j}(x)\right)^2 dx \geq r^2 \int_{\frac{j-1}{d}}^{\frac{j}{d}} g_{\beta,j}^2(x)dx = \frac{r^2 c}{d^{2\beta+1}}.$$

As a result,

$$\begin{aligned}
\|f - f_v\|_2^2 &= \sum_{j=1}^{d} \int_{\frac{j-1}{d}}^{\frac{j}{d}} \left(f(x) - rv_j g_{\beta,j}(x)\right)^2 dx \\
&\geq \sum_{j=1}^{d} 1\{V_j(f) \neq v_j\} \frac{r^2 c}{d^{2\beta+1}} = \frac{r^2 c}{d^{2\beta+1}} \sum_{j=1}^{d} 1\{V_j(f) \neq v_j\}
\end{aligned}$$

for all $f \in \mathcal{F}_\beta[r]$. Setting $V : \mathcal{F}_\beta[r] \longrightarrow \{-1,+1\}^d$, we can apply Theorem 3 to derive

$$\mathcal{R}_n\left(\mathcal{F}_\beta[r^2], \|\cdot\|_2^2, f\right) \geq \frac{r^2 c}{2d^{2\beta}} \left[1 - \sqrt{\frac{nc_{f,\kappa}}{d} \sum_{j=1}^{d} \frac{\|P_{+j} - P_{-j}\|_{TV}^{1+\kappa}}{\left(1 - \|P_{+j} - P_{-j}\|_{TV}\right)^\kappa}}\right]$$

where $f_{\pm j} = \frac{1}{2^{d-1}} \sum_{v_j := \pm 1} f_v = 1 \pm rg_{\beta,j}$ for $j \in 1, \ldots, d$. Consequently,

$$\begin{aligned}
\|f_{+j} - f_{-j}\|_{TV} &= \int_0^1 (2rg_{\beta,j}(x))_+ dx = \int_{\frac{j-1}{d}}^{\frac{j-0.5}{d}} 2rg_{\beta,j}(x)dx \\
&= \int_{\frac{j-1}{d}}^{\frac{j-0.5}{d}} \frac{2r}{d^\beta} g_\beta\left(d\left(x - \frac{j-1}{d}\right)\right) dx \\
&= \int_0^{\frac{1}{2}} \frac{2r}{d^{\beta+1}} g_\beta(u)\, du \\
&= 2\frac{rc_{1/2}}{d^{\beta+1}}.
\end{aligned}$$

Therefore, the lower bound becomes

$$
\begin{aligned}
\mathcal{R}_n\left(\mathcal{F}_\beta[r], \|\cdot\|_2^2, f\right) &\geq \frac{r^2 c}{2d^{2\beta}}\left[1-\sqrt{\frac{nc_{f,\kappa}}{d}\sum_{j=1}^{d}\frac{1}{\left(1-\frac{2rc_{1/2}}{d^{\beta+1}}\right)^{\kappa}}\left(\frac{2rc_{1/2}}{d^{\beta+1}}\right)^{1+\kappa}}\right] \\
&\geq \frac{r^2 c}{d^{2\beta}}\left[1-\sqrt{\frac{nc_{f,\kappa}}{\left(1-\frac{2rc_{1/2}}{d^{\beta+1}}\right)^{\kappa}}\frac{(2rc_{1/2})^{1+\kappa}}{d^{(\beta+1)(1+\kappa)}}}\right].
\end{aligned}
$$

Put $d = (2^{1+\kappa}\cdot 2^{1+\kappa} r^{1+\kappa} nc_{f,\kappa})^{\frac{1}{(\beta+1)(1+\kappa)}}$ to get

$$
\mathcal{R}_n\left(\mathcal{F}_\beta[r], \|\cdot\|_2^2, f\right) \geq 2^{-\frac{2\beta}{\beta+1}} r^{\frac{2}{\beta+1}} c\left(1-\sqrt{\frac{1}{2}}\right)(nc_{f,\kappa})^{-\frac{2\beta}{(\beta+1)(1+\kappa)}}
$$

for $n > \frac{1}{c_{f,\kappa}}\left(c_{1/2}\right)^{1+\kappa}$.

5.9 Proof of Corollary 4

The lower bound can be shown by using Corollary 3 and the fact that G_μ satisfies the condition of Corollary 3 for $\kappa = 1$. Also, the coefficient can be obtained similarly as in the proof of Corollary 2. Now as in the proof of Corollary 2, we show the upper bound with the explicit estimator and mechanism. For $d \in \mathbb{N}$ and given $x \in [0,1]$, define the mechanism M as

$$
M(x) = (\phi_1(x), \cdots, \phi_d(x)) + \epsilon
$$

where $\epsilon \sim N(0, \frac{2d}{\mu^2} I_{d\times d})$. Then, the mechanism $\phi_i(x) \longrightarrow N(\phi_i(x), \frac{2d}{\mu^2})$ is $\frac{\mu}{\sqrt{d}}$-GDP. It is known that composition of $\mu_1-, \cdots, \mu_d$-GDP mechanisms is $\sqrt{\mu_1^2 + \cdots + \mu_d^2}$-GDP (Dong et al., 2022). So, M is μ-GDP since $\sqrt{\left(\frac{\mu}{\sqrt{d}}\right)^2 + \cdots + \left(\frac{\mu}{\sqrt{d}}\right)^2} = \mu$.

Let $X_i \sim f \in \mathbb{F}_\beta[r^2]$ and $(Z_{i1}, \ldots, Z_{id}) = M(X_i)$, and define the estimator as follows:

$$
\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{d} Z_{ij}\phi_j.
$$

Then, the risk under l_2 norm is give as:

$$
\mathbb{E}\left[\|\hat{\theta} - f\|_2^2\right] = \sum_{j=1}^{d}\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} Z_{ij} - \theta_j\right)^2\right] + \sum_{j>d}\theta_j^2.
$$

Note that if $f = \sum \theta_j \phi_j$, $\mathbb{E}\left[\phi_j(X) | X \sim f\right] = \theta_j$, thus $\mathbb{E}\left[Z_{ij}\right] = \mathbb{E}\left[\phi_j(X)\right] = \theta_j$.

Denote $Z_{i,j} = \phi_j(X_i) + \epsilon_{i,j}$ where $\epsilon_{i,j} \sim N(0, \frac{2d}{\mu^2})$. Then, we have

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} Z_{ij} - \theta_j\right)^2\right] &= \frac{1}{n} Var\left(Z_{1j}\right) \\
&= \frac{1}{n}\left[Var(\phi_j(X_1)) + Var(\epsilon_{1,j})\right] \leq \frac{1}{n} + \frac{2d}{n\mu^2},
\end{aligned}$$

since $|\phi_j| \leq 1$. Note that

$$\sum_{j>d} \theta_j^2 \leq \frac{1}{d^{2\beta}} \sum_{j>d} j^{2\beta}\theta_j^2 \leq \frac{r^2}{d^{2\beta}}.$$

As a result,

$$\mathbb{E}\left[\|\hat{\theta} - f\|_2^2\right] \leq \frac{1}{n}\frac{2d^2}{\mu^2} + \frac{r^2}{d^{2\beta}} + \frac{d}{n},$$

and by setting $d = (0.5n\mu^2 r^2)^{\frac{1}{2\beta+2}}$, we have

$$\mathbb{E}\left[\|\hat{\theta} - f\|_2^2\right] \leq r^2(0.5n\mu^2 r^2)^{-\frac{2\beta}{2\beta+2}} + (0.5\mu^2 r^2)^{\frac{1}{2\beta+2}} n^{-\frac{2\beta+1}{2\beta+2}}.$$

Therefore,

$$\mathcal{R}_n\left(\theta\left(\mathcal{F}_\beta[r^2]\right), \|\cdot\|_2^2, \mathcal{M}_\mu\right) \leq c r^{\frac{2}{\beta+1}} n^{-\frac{2\beta}{2\beta+2}}$$

for some $c > 0$ depending on μ and β.

5.10 Proof of Lemma 2

Note that

$$\begin{aligned}
\mathbb{P}_{m_2}\left(\frac{m_1(Z)}{m_2(Z)} > a\right) &= \int_{\mathcal{Z}} 1\left\{\frac{m_1}{m_2} > a\right\} m_2(z)dz \\
&\leq \int_{\mathcal{Z}} \left[\left(\frac{m_1}{m_2} - a + 1\right)_+ - \left(\frac{m_1}{m_2} - a\right)_+\right] m_2(z)dz \\
&= \mathbb{E}_{a-1}(m_1||m_2) - \mathbb{E}_a(m_1||m_2).
\end{aligned}$$

Then, Proposition S4 states that $\mathbb{E}_{a-1}(m_1||m_2) \leq \delta_f(a-1)$ for $a \geq 2$, which leads to

$$\mathbb{P}_{m_2}\left(\frac{m_1(Z)}{m_2(Z)} > a\right) \leq \delta_f(a-1).$$

References

Asoodeh, S., Aliakbarpour, M., and Calmon, F. P. (2021). Local differential privacy is equivalent to contraction of an f-divergence. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pages 545–550.

Assouad, P. (1983). Deux remarques sur l'estimation. *Comptes rendus des séances de l'Académie des sciences. Série 1, Mathématique*, 296(23):1021–1024.

Chen, E. (2023). Laplacian random number generator. *(https://www.mathworks.com/matlabcentral/fileexchange/13705-laplacian-random-number-generator), MATLAB Central File Exchange. Retrieved May 16, 2023.*

Dong, J., Roth, A., and Su, W. J. (2022). Gaussian differential privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84:3–37.

Duchi, J. C., Jordan, M. I., and Wainwright, M. J. (2013). Local privacy and statistical minimax rates. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 429–438.

Duchi, J. C., Jordan, M. I., and Wainwright, M. J. (2018). Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113(521):182–201.

Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., and Naor, M. (2006a). Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology-EUROCRYPT 2006: 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Proceedings 25*. Springer.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006b). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography: Third Theory of Cryptography Conference, Proceedings 3*, pages 265–284. Springer.

LeCam, L. (1973). Convergence of estimates under dimensionality restrictions. *The Annals of Statistics 1*, pages 38–53.

Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation, 2009*. Springer Series in Statistics.

Yu, B. (1997). Assouad, fano, and le cam. *Festschrift for Lucien Le Cam: research papers in probability and statistics*, pages 423–435.