Privacy Profiles for Private Selection

Antti Koskela[1] **Rachel Redberg**[2] **Yu-Xiang Wang**[3]

Abstract

Private selection mechanisms (e.g., Report Noisy Max, Sparse Vector) are fundamental primitives of differentially private (DP) data analysis with wide applications to private query release, voting, and hyperparameter tuning. Recent work (Liu & Talwar, 2019; Papernot & Steinke, 2022) has made significant progress in both generalizing private selection mechanisms and tightening their privacy analysis using modern numerical privacy accounting tools, e.g., Renyi DP. But Renyi DP is known to be lossy when (ϵ, δ)-DP is ultimately needed, and there is a trend to close the gap by directly handling privacy profiles, i.e., δ as a function of ϵ or its equivalent dual form known as f-DPs. In this paper, we work out an easy-to-use recipe that bounds the privacy profiles of ReportNoisyMax and PrivateTuning using the privacy profiles of the base algorithms they corral. Numerically, our approach improves over the RDP-based accounting in all regimes of interest and leads to substantial benefits in end-to-end private learning experiments. Our general result also allows analysing the case of binomially-distributed number of rounds, which leads to more concentrated distributions compared to the previously considered Poisson distribution.

1. Introduction

Differential privacy (DP) bounds the privacy loss incurred when an algorithm is run on a dataset. While the analysis of this bound is often quite nuanced, the rough tally is that whenever the algorithm accesses the data, it incurs a privacy cost.

By this rule of thumb, data privacy for modern machine learning (ML) applications is in trouble. Most modern ML algorithms are notoriously finicky and require extensive hyperparameter tuning in order to achieve good performance. In the context of data privacy, this means that the evaluation of every additional hyperparameter candidate could have a privacy cost.

Imagine that you have an (ϵ, δ)-DP training algorithm and wish to evaluate it across k candidates. Naïvely it's possible to analyze the privacy of this procedure via *composition*, a property of differential privacy which in its most basic form implies that the privacy parameters ϵ and δ "add up" (Dwork et al., 2006). This naïve analysis would allow an ML practitioner to privately release all k trained models at a cost of $(k\epsilon, k\delta)$-DP — i.e., the privacy cost scales linearly in the number of model evaluations. But for hyperparameter tuning, practitioners are typically most interested in the "best" model (with the highest quality score) and hence might want to output only one out of the k trained models. Thus, a tighter analysis may often be available for this class of *private selection* mechanisms which choose from a set of candidates the item which approximately maximizes a given quality score.

The aim of our work is to improve the privacy analysis of private hyperparameter tuning algorithms for machine learning algorithms. This line of inquiry originates with Chaudhuri & Vinterbo (2013), who highlighted the need for differentially private parameter tuning in order to achieve end-to-end privacy in machine learning applications. Chaudhuri & Vinterbo (2013) proposed a procedure for training and validation which provides differential privacy under the somewhat stringent condition that the learning algorithm must uphold a notion of Lipschitz-like stability on the score function.

The subsequent work of Liu & Talwar (2019) ushered in the age of black-box frameworks for differentially private hyperparameter tuning. Liu & Talwar (2019) relaxed the stability assumption on the score function to the weaker requirement that each individual candidate be differentially private. They proposed black-box tuning methods that use a random stopping strategy; in particular, the number of candidates evaluated by the learning algorithm follows a geometric distribution. Liu & Talwar (2019) showed that if each candidate (or "base mechanism", as we will call it in our work) satisfies ϵ-DP, then their end-to-end algorithm for

[1]Nokia Bell Labs [2]Northeastern University [3]UC San Diego. Correspondence to: Antti Koskela <antti.h.koskela@nokia-bell-labs.com>.

Proceedings of the 41^{st} International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

private selection satifies 3ϵ-DP. They also gave approximate (ϵ, δ)-DP results.

The Liu & Talwar (2019) bounds apply only to (ϵ, δ)-DP. Mohapatra et al. (2022) gave adaptive tuning methods and showed that for a reasonable number of privately chosen candidates, naïve accounting via Rényi differential privacy (RDP) often yields tighter DP bounds than private selection using the methods by Liu & Talwar (2019). One might then ask, what could more sophisticated RDP accounting do?

We base our approach on the quite notable results of Papernot & Steinke (2022), which built on Liu & Talwar (2019)'s work and provided RDP bounds for black-box tuning algorithms which scale logarithmically in relation to the number of model evaluations. Papernot & Steinke (2022) also generalized the random stopping strategy to encompass distributions other than the geometric distribution: the number of repetitions can now be distributed according to a truncated negative binomial distribution, or to a Poisson distribution. In either case the RDP bound for the end-to-end algorithm is constructed from the RDP and DP bounds on the base mechanisms.

What is missing from the private selection and private hyperparameter tuning line of work, are bounds that would directly use the privacy profiles of the base mechanisms as building blocks and would also be capable of taking advantage of numerical accounting methods (e.g., the recent work of Koskela et al., 2021; Zhu et al., 2022; Gopi et al., 2021). In this work we address this shortfall. Our proposed bounds utilize only point-wise information about the (ϵ, δ)-profile of the base mechanism, similarly to bounds given by Liu & Talwar (2019).

As one application of our results we consider Differentially Private Stochastic Gradient Descent (DP-SGD) (Abadi et al., 2016), a popular technique for training machine learning models with DP. DP-SGD introduces extra hyperparameters such as the noise level σ and clipping constant C; factors like the subsampling ratio q and the training duration also affect both the privacy and accuracy of the methods. As demonstrated by Papernot & Steinke (2022), tweaking DP-SGD's hyperparameters often relies on using sensitive data which require privacy protection. The risk of data leakage from hyperparameters is arguably smaller than from model parameters, but developing methods with low privacy costs has proven challenging. The leading algorithms proposed by Papernot & Steinke (2022) still incur a significant privacy cost overhead. Our aim is to further lower this overhead.

As another application of our analysis we consider Generalized Propose-Test-Release (Redberg et al., 2023), which broadens the reach of the Propose-Test-Release (PTR) framework by allowing it to handle queries with unbounded sensitivity, making it applicable to problems such as linear regression. The method proposed by Redberg et al. (2023) gives point-wise (ϵ, δ)-privacy guarantees for generalized PTR. Thus, in order to account for the privacy cost of tuning the hyperparameters of the underlying queries, we can directly use our analysis for which point-wise (ϵ, δ)-guarantees are sufficient. We show that our bounds are considerably tighter than those of Liu & Talwar (2019), the only previous applicable results. We also empirically illustrate that compared to well-established non-adaptive methods, our bounds considerably improve the privacy-utility trade-off for linear regression problems.

2. Preliminaries

We first give the basic definitions. An input dataset containing n data points is denoted as $X = \{x_1, \ldots, x_n\}$. Denote the set of all possible datasets by $\mathcal{X}$. We say X and X' are neighbors if we get one by adding or removing one data element to or from the other, or by replacing one data element in the other (denoted $X \sim X'$). Consider a randomized mechanism $\mathcal{M} : \mathcal{X} \to \mathcal{O}$, where $\mathcal{O}$ denotes the output space. The (ϵ, δ)-definition of DP can be given as follows (Dwork, 2006).

Definition 2.1. Let $\epsilon > 0$ and $\delta \in [0, 1]$. We say that a mechanism $\mathcal{M}$ is (ϵ, δ)-DP, if for all neighboring datasets X and X' and for every measurable set $E \subset \mathcal{O}$ we have:

$$\Pr(\mathcal{M}(X) \in E) \leq \mathrm{e}^{\epsilon} \Pr(\mathcal{M}(X') \in E) + \delta.$$

We state many of our results for general f-divergences. For a convex function $f : [0, \infty) \to \mathbb{R}$, we define the f-divergence between distributions P and Q taking values in $\mathcal{Y}$ as

$$H_f(P||Q) = \int_{\mathcal{Y}} f\left(\frac{P(y)}{Q(y)}\right) Q(y) \, \mathrm{d}y.$$

Notice that we do not require the normalization $f(1) = 0$ often used in the so-called Czsisár divergences (Liese & Vajda, 2006) as it is not necessary and can be obtained simply by scaling. Especially, our aim is to find tight bounds for the hockey stick divergence, i.e., when $f(z) = [z - \mathrm{e}^{\epsilon}]_+$ for some $\epsilon \in \mathbb{R}$. This is due to the fact that tight (ϵ, δ)-bounds can be obtained using the hockey-stick-divergence:

Lemma 2.2 (Balle et al. 2018, Theorem 1)**.** *A mechanism $\mathcal{M}$ satisfies (ϵ, δ)-DP if and only if, $\max_{X \sim X'} H_f(\mathcal{M}(X)||\mathcal{M}(X')) \leq \delta$ for $f(z) = [z - \mathrm{e}^{\epsilon}]_+$.*

We denote the hockey stick divergence determined by $\epsilon \in \mathbb{R}$ by $H_{\mathrm{e}^{\epsilon}}$ throughout the paper, and will refer to

$$\delta_{\mathcal{M}}(\epsilon) := \max_{X \sim X'} H_{e^{\epsilon}}(\mathcal{M}(X)||\mathcal{M}(X'))$$

as the *privacy profile* of mechanism $\mathcal{M}$.

We will also use the Rényi differential privacy (RDP) (Mironov, 2017) which is defined as follows.

Rényi divergence of order $\alpha > 1$ between two distributions P and Q is defined as

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1} \log \int \left(\frac{P(t)}{Q(t)} \right)^\alpha Q(t) \, \mathrm{d}t \qquad (2.1)$$

and we say that a mechanism $\mathcal{M}$ is (α, ϵ)-RDP, if for all neighboring datasets X and X', the output distributions of $\mathcal{M}(X)$ and $\mathcal{M}(X')$ have Rényi divergence of order $\alpha > 1$ at most ϵ, i.e., if

$$\max_{X \sim X'} D_\alpha\big(\mathcal{M}(X)||\mathcal{M}(X')\big) \leq \epsilon.$$

To convert from RDP to (ϵ, δ)-DP, we use the formula given in Appendix (D.1).

Notice that the Rényi divergence of order $\alpha > 1$ is a scaled logarithm of an f-divergence determined by $f(z) = z^\alpha$. Existing RDP analyses for Report Noisy Max and private selection use the fact $f(u, v) = (\frac{u}{v})^\alpha v$ is a jointly convex function for u and v. We formulate many of our results for general f-divergences, for which the following technical result will then be central (see e.g., Ch. 3, Boyd & Vandenberghe, 2004).

Lemma 2.3. *If $f : \mathbb{R}_+ \to \mathbb{R}_+$ is a convex function, then the function $f\left(\frac{x}{y}\right) y$ is jointly convex for $x \geq 0$ and $y > 0$.*

For the analysis of private selection algorithms, the number of times K that the base algorithm is evaluated is a random variable. To analyze different alternatives for choosing K, we will need the concept of probability generating functions.

Definition 2.4. Let K be a random variable taking values in $\mathbb{N} \cup \{0\}$. The probability generating function (PGF) of K, $\varphi : [0, 1] \to [0, 1]$ is defined as

$$\varphi(z) = \sum_{k=0}^{\infty} \mathbb{P}(K = k) \cdot z^k.$$

Our main result Thm. 4.1 is stated for a general PGF φ and we use it obtain method-specific bounds. Throughout the paper, we will denote $m = \mathbb{E}[K]$.

3. Report Noisy Max for Additive Noise Mechanisms

As a first application of the hockey-stick divergence-based analysis, we consider the Report Noisy Max (RNM) of 1-dimensional additive noise mechanisms. This will serve as a segue into the more involved analysis of private selection, where we also obtain bigger gains compared to the previous results. Our analysis here is based on the RDP analysis by Zhu & Wang (2022). We mention that recent applications of RNM include private in-context learning of LLMs (Wu et al., 2024; Tang et al., 2024).

Let $X = \{x_1, x_2, \ldots, x_N\}$, where $x_i \in \mathcal{X}$ for all $i \in [N]$, be a data set and consider the mechanism

$$\mathcal{M}(X) = \arg\max\{\mathcal{M}_1(X), \ldots, \mathcal{M}_m(X)\}, \qquad (3.1)$$

where for every $i \in [m]$,

$$\mathcal{M}_i(X) = f(X) + Z_i,$$

for some function $f_i : \mathcal{X} \to \mathbb{R}$ such that

$$\max_{X \sim X'} |f_i(X) - f_i(X')| \leq 1$$

and where the noises Z_i, $i \in [m]$, are i.i.d. We have the following existing RDP bound.

Theorem 3.1 (Zhu & Wang 2022, Theorem 8)**.** *Let $\alpha > 1$. The mechanism $\mathcal{M}(X)$ is (α, ϵ')-RDP for*

$$\epsilon' = \epsilon + \frac{\log m}{\alpha - 1}, \qquad (3.2)$$

where ϵ denotes the RDP guarantee of order α for an additive noise mechanism with noise Z and sensitivity 2.

Theorem 3.2. *Let $X \sim X'$ and $\epsilon \in \mathbb{R}$. We have:*

$$H_{\mathrm{e}^\epsilon}\big(\mathcal{M}(X)||\mathcal{M}(X')\big) \leq m \cdot \delta(\epsilon), \qquad (3.3)$$

where $\delta(\epsilon)$ is the privacy profile of the additive noise mechanism with sensitivity 2. If we assume monotonicity, i.e., if $f_i(X) \geq f_i(X')$ or $f_i(X) \leq f_i(X')$ for all $i \in [m]$, then the sensitivity of 2 can be replaced with a sensitivity of 1.

Theorem 3.3. *For an adaptive composition of k mechanisms of the form (3.1), we get the privacy profile upper bound $m^k \cdot \delta(\epsilon)$, where $\delta(\epsilon)$ is the privacy profile of an m-wise composition of an additive noise mechanism with noise Z and sensitivity 2.*

The RDP analysis of the private selection algorithm provided by Papernot & Steinke (2022) shows that the RDP guarantees essentially grow as $\log m$, where m is the expected number of candidates for the private selection algorithm. When Z is normally distributed, we directly see from our analysis that for a fixed δ the ϵ-values of the private selection algorithm grow as $\frac{1}{\sigma} \log^{\frac{1}{2}} \frac{m}{\delta}$. This result is obtained for the RNM using a simple tail bound of the Gaussian.

Corollary 3.4. *Consider the mechanism $\mathcal{M}$ defined in Eq. 3.1 and suppose Z is normally distributed with variance σ^2. Let $\delta > 0$. Then $\mathcal{M}$ is (ϵ, δ)-DP for*

$$\epsilon = \frac{2}{\sigma^2} + \frac{2}{\sigma} \sqrt{2 \log \frac{m}{\delta}}$$

As Figure 1 shows, with the bound given in Thm. 3.2 we observe differences one usually observes between the accurate hockey stick and Rényi divergence bounds (for comparisons, see, e.g. Canonne et al., 2020). When considering the private selection problem where the number of candidates is randomized, we obtain larger differences. Also, the ϵ-growth order $\mathcal{O}(\frac{1}{\sigma} \log^{\frac{1}{2}} \frac{m}{\delta})$ is retained. All of this is discussed in the next two sections.

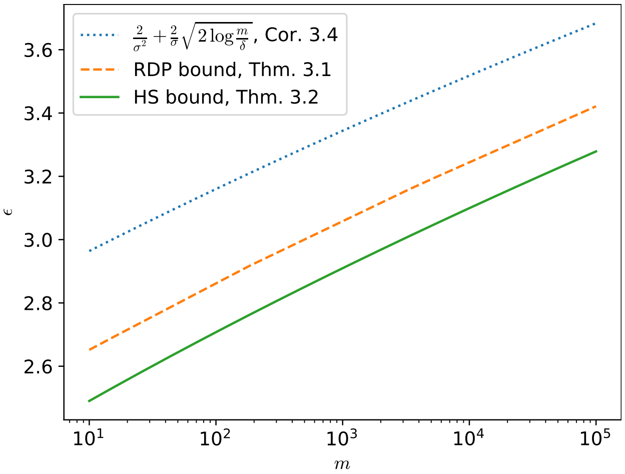

Figure 1. Comparison of the (ϵ, δ)-bounds for the RNM mechanism (3.1) when the base mechanisms $\mathcal{M}_i$, $i \in [m]$, are 1-d Gaussian mechanisms with sensitivity 1 and noise scale $\sigma = 4.0$, and when $\delta = 10^{-6}$. Also plotted is the bound of Corollary 3.4.

4. General Bound for Private Selection

We use the notation and setting of Papernot & Steinke (2022). This means that the tuning algorithm A outputs both the argument of the maximizer (the best hyperparameters or the index) and the output of the base mechanism (e.g., the model trained with the best hyperparameters).

Let $\mathcal{Y}$ denote the finite and ordered output space (the quality score), $Q(y)$ the density function of the quality score of the base mechanism taking values in $\mathcal{Y}$, K the random variable for the number of times the base mechanism is run and $A(y)$ the density function of the tuning algorithm that outputs the best one of the K alternatives. Let A and A' denote the output distributions of the tuning algorithm evaluated on neighboring datasets X and X', respectively. Then, the f-divergence between A and A' can be bounded using the following result. We use the proof technique from Lemma 7 of Papernot & Steinke (2022) and similarly invoke the argument that our proof for finite $\mathcal{Y}$ can be extended to the general case.

Theorem 4.1. *Let $X \sim X'$ and let A and A' be the density functions of the hyperparameter tuning algorithm as defined above, evaluated on X and X', respectively. Let Q and Q' be the density functions of the quality score of the base mechanism, evaluated on X and X', respectively. Let K be random variable for the times the base mechanism is run and $\varphi(z)$ the PGF of K. Let $f : [0, \infty) \to \mathbb{R}$ be a convex function. Then,*

$$H_f(A||A') \leq \sum_{y\in\mathcal{Y}} f\left(\frac{Q(y)\varphi'(q_y)}{Q'(y)\varphi'(q'_y)}\right) \cdot Q'(y)\varphi'(q'_y),$$

where for each $y \in \mathcal{Y}$, q_y and q'_y are obtained by applying the same y-dependent post-processing function to Q and Q', respectively.

Looking at the bound given by Thm. 4.1, we can decompose the right-hand side in case f corresponds to the Rényi divergence and obtain (Lemma 7, Papernot & Steinke, 2022) as a corollary.

Remark 4.2. *Let $f(z) = z^\lambda$ for some $\lambda \geq 1$. Then, by Thm. 4.1,*

$$\begin{aligned} &\mathrm{e}^{(\lambda-1)D_\lambda(A||A')} = H_f(A||A') \\ &\leq \sum_{y\in\mathcal{Y}} \left(\frac{Q(y)}{Q'(y)}\right)^\lambda Q'(y) \cdot \left(\frac{\varphi'(q_y)}{\varphi'(q'_y)}\right)^\lambda \varphi'(q'_y) \\ &\leq \left[\sum_{y\in\mathcal{Y}} \left(\frac{Q(y)}{Q'(y)}\right)^\lambda Q'(y)\right] \left(\frac{\varphi'(q)}{\varphi'(q')}\right)^\lambda \varphi'(q'), \end{aligned} \tag{4.1}$$

for some q and q' that are obtained by applying the same post-processing function to Q and Q', respectively. Taking the logarithm and dividing by $\lambda - 1$, we obtain (Lemma 7, Papernot & Steinke, 2022).

In case of RDP analysis, the bounds for the randomized private selection algorithms (see Theorems 5.1 and 5.6) allow optimizing the bound (4.1) with respect to the privacy profile of Q. For example: as q is a result of post-processing of Q, we have that for all $\epsilon \geq 0$

$$q \leq \mathrm{e}^\epsilon q' + \delta(\epsilon),$$

where $\delta(\epsilon)$ gives the privacy profile of Q, and we can carry out an optimization of the bound (4.1) individually for each RDP order λ w.r.t. ϵ. In case the function f in Thm. 4.1 corresponds to the hockey stick divergence, the best we can have is a uniform bound for the ratio $\frac{\varphi'(q_y)}{\varphi'(q'_y)}$ which uses the bound $q_y \leq \mathrm{e}^\epsilon q'_y + \delta(\epsilon)$ with the same value of ϵ for each $y \in \mathcal{Y}$. As a result there is one degree of freedom less to optimize in the upper bounds we obtain using the hockey stick divergence. Nevertheless, the analysis with the hockey stick divergence becomes much simpler and the resulting (ϵ, δ)-DP bounds for private selection become tighter.

5. Distribution Specific Bounds for Private Selection

We next consider privacy profile bounds for two specific choices of the distribution K, the truncated negative binomial distribution and the binomial distribution. As we show, in both cases the bounds allow evaluating a considerably larger number of private candidates than the state-of-the-art bounds. The bounds for the binomial distribution generalize the bounds for the Poisson distribution, improve the state-of-the-art bounds for the Poisson distribution, and also allow concentrating K further.

5.1. Truncated Negative Binomial Distribution

Suppose the number of trials K is distributed according to the truncated negative binomial distribution $\mathcal{D}_{\eta,\gamma}$ which is determined by $\gamma \in (0,1)$ and $\eta \in (-1,\infty)$ and by the probabilities ($k \in \mathbb{N}$) for $\eta \neq 0$ by

$$\mathbb{P}(K = k) = \frac{(1-\gamma)^k}{\gamma^{-\eta} - 1} \prod_{i=0}^{k-1} \left(\frac{i+\eta}{i+1} \right).$$

and for $\eta = 0$ by

$$\mathbb{P}(K = k) = \frac{(1-\gamma)^k}{k \cdot \log 1/\gamma}.$$

It holds that when $\eta \neq 0$,

$$\mathbb{E}K = \frac{\eta(1-\gamma)}{\gamma(1-\gamma^\eta)}$$

and when $\eta = 0$,

$$\mathbb{E}K = \frac{1/\gamma - 1}{\log 1/\gamma}.$$

The derivative of the corresponding probability generating function is given by

$$\varphi'(z) = \big(1 - (1-\gamma)z\big)^{-\eta-1} \cdot \gamma^{\eta+1} \cdot \mathbb{E}K. \qquad (5.1)$$

As a baseline, we consider the following RDP bound.

Theorem 5.1 (Papernot & Steinke 2022)**.** *Let Q satisfy $\big(\alpha,\epsilon\big)$-RDP and $\big(\widehat{\alpha},\widehat{\epsilon}\big)$-RDP for some $\alpha \in (1,\infty)$ and $\widehat{\alpha} \in [1,\infty)$. Draw K from a truncated negative binomial distribution distribution $\mathcal{D}_{\eta,\gamma}$, where $\gamma \in (0,1)$ and $\eta \in (-1,\infty)$. Run $Q(X)$ for K times. Then $A(X)$ returns the best value of those K runs (also Q's output). Then A satisfies $\big(\alpha,\epsilon'(\alpha)\big)$-RDP, where*

$$\epsilon'(\alpha) = \epsilon(\alpha) + (\eta+1)\left(1 - \frac{1}{\widehat{\alpha}}\right)\widehat{\epsilon} + \frac{(1+\eta)\cdot\log(1/\gamma)}{\hat{\alpha}} + \frac{\log m}{\alpha - 1}.$$

Using the PGF (5.1) and our general result Thm.4.1, we obtain the following bound using the hockey-stick divergence.

Theorem 5.2. *Let $K \sim \mathcal{D}_{\eta,\gamma}$ and let $\delta(\epsilon_1)$, $\epsilon_1 \in \mathbb{R}$, define the privacy profile of the base mechanism Q. Then, for A and A', the output distributions of the selection algorithm evaluated on neighboring datasets X and X', respectively, and for all $\epsilon_1 \geq 0$,*

$$H_{\mathrm{e}^\epsilon}(A||A') \leq m \cdot \delta(\widehat{\epsilon})$$

where

$$\widehat{\epsilon} = \epsilon - (\eta+1)\log\left(\mathrm{e}^{\epsilon_1} + \frac{1-\gamma}{\gamma} \cdot \delta(\epsilon_1)\right).$$

We directly obtain the following pure ϵ-DP result from Thm. 5.2. Notice that it includes Theorem 1.3 (3ϵ-bound) by Liu & Talwar (2019) as a special case.

Corollary 5.3. *Let $K \sim \mathcal{D}_{\eta,\gamma}$. If the base mechanism Q is ϵ-DP, then the selection algorithm A is $(\eta+2)\epsilon$-DP. For $\eta = 1$ we get Theorem 1.3 of (Liu & Talwar, 2019).*

The following (ϵ,δ)-DP result is also a straightforward corollary of Thm. 5.2.

Corollary 5.4. *Let $K \sim \mathcal{D}_{\eta,\gamma}$. If the base mechanism Q is (ϵ,δ)-DP, then then the selection algorithm A is $\big((\eta+2)\epsilon + \gamma^{-1}\delta, m\delta\big)$-DP.*

Notice that for the geometric distribution ($\eta = 1$), Cor. 5.4 implies that if Q is (ϵ,δ)-DP, then A is $\big(3\epsilon + m\,\delta, m\,\delta\big)$-DP.

We can show that for a fixed δ the ϵ-value of the private selection algorithm is $\mathcal{O}(\log^{\frac{1}{2}}\frac{m}{\delta})$, in case the base mechanism is Gaussian differentially private (Dong et al., 2022), i.e., if its privacy profile is dominated by the hockey-stick divergence between two Gaussians.

Corollary 5.5 (ϵ-values when Q is GDP)**.** *Let $K \sim \mathcal{D}_{\eta,\gamma}$ with $\eta \geq 1$ and suppose the base mechanism is dominated by the Gaussian mechanism with noise parameter $\sigma > 0$ and L_2-sensitivity 1. Then, for a fixed $\delta > 0$, the private selection algorithm A is (ϵ,δ)-DP for*

$$\epsilon = (\eta+1)\left(\frac{1}{2\sigma^2} + \frac{1}{\sigma}\sqrt{2\log\frac{m}{\delta}}\right) + 2\delta.$$

Figure 2 illustrates the various bounds for the truncated negative binomial distribution with $\eta = 1$ when the base mechanism is the Gaussian mechanism with $\sigma = 4$ and L_2-sensitivity 1, when $m = 30, 300$ and 3000. Figure 3 shows the increase of the ϵ-values for a fixed value of δ and the bound of Lemma 5.5 for comparison.

5.2. Binomial Distribution

The choice of distribution for the number of repetitions K has practical implications on the utility of private selection (Papernot & Steinke, 2022). One issue is that when the distribution of K is less concentrated, the value of K is likely to be small — meaning fewer candidates evaluated — even when its expectation is large. This is especially problematic for smaller numbers of candidates where the expectation of K is not even large to begin with! As a practical alternative to the less-concentrated truncated negative binomial distribution, Papernot & Steinke (2022) consider the Poisson distribution for K.

In our work, by considering $K \sim \mathrm{Bin}(n,p)$ to be binomially distributed, we can still further concentrate the distribution of K by allowing a small additional privacy cost.

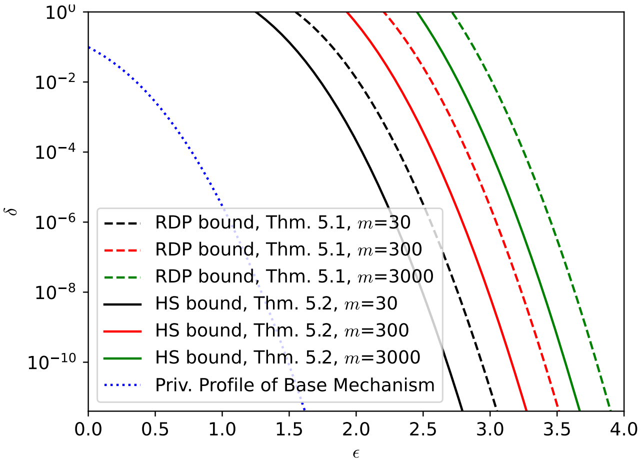

Figure 2. Comparison of various (ϵ, δ)-bounds when $K \sim \mathcal{D}_{\eta,\gamma}$ with $\eta = 1$ (the geometric distribution, $\mathbb{E}[K] = \gamma^{-1}$) and $m = 30, 300, 3000$. The base mechanism is the Gaussian mechanism with L_2-sensitivity 1 and noise parameter $\sigma = 4$.

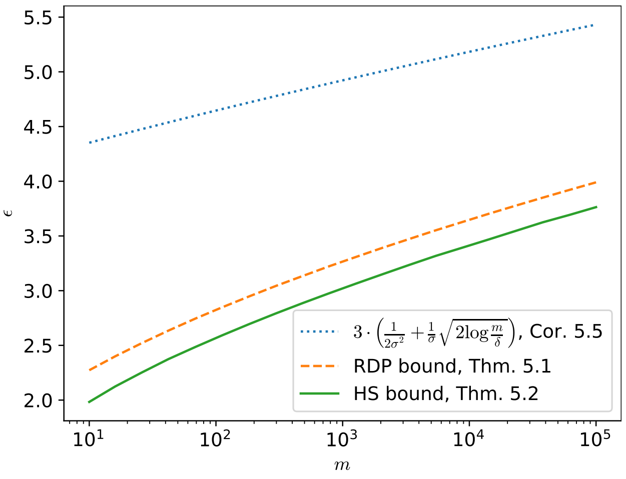

Figure 3. Growth of ϵ-values for $\delta = 10^{-6}$ for the RDP and hockey stick divergence based bounds as a function of m, when $K \sim \mathcal{D}_{\eta,\gamma}$ with $\eta = 1$ and the base mechanism is the Gaussian mechanism with L_2-sensitivity 1 and noise parameter $\sigma = 4$. The privacy profile bound given by Thm. 5.2 retains the $\mathcal{O}(\log^{\frac{1}{2}} \frac{m}{\delta})$ growth of ϵ-values from the DP RNM (See Fig. 1).

The hockey stick-based approach simplifies the analysis such that when compared to the RDP bounds for Poisson-distributed K, we essentially get much more concentrated K for the same privacy cost using the binomial distribution.

Our bound for $K \sim \mathrm{Bin}(n, p)$ is a strict generalization of the Poisson distribution case, as we get the result for the Poisson distribution as a limit when $p \to 0$. The RDP bound to compare is the one given by Papernot & Steinke (2022).

Theorem 5.6 (Papernot & Steinke 2022)**.** *Let Q satisfy (α, ϵ)-RDP and $(\widehat{\epsilon}, \widehat{\delta})$-DP for some $\alpha \in (1, \infty)$ and $\epsilon, \widehat{\epsilon}, \widehat{\delta} \geq 0$. Draw K from a Poisson distribution with mean m. Run $Q(X)$ for K times. Then $A(X)$ returns the best value of those K runs (also Q's output). If $K = 0$, $A(X)$ returns some arbitrary output. If $\mathrm{e}^{\widehat{\epsilon}} \leq 1 + \frac{1}{\alpha - 1}$, then A satisfies $\left(\alpha, \epsilon'(\alpha)\right)$-RDP, where*

$$\epsilon'(\alpha) = \epsilon + m \cdot \widehat{\delta} + \frac{\log m}{\alpha - 1}.$$

For a hockey stick divergence bound, we consider the binomially distributed K which is a strict generalization of the Poisson case and includes its privacy profile bound as a special case. We can derive the following result from Thm. 4.1 when using the PGF of the binomial distribution.

Theorem 5.7. *Let $K \sim \mathrm{Bin}(n, p)$ for some $n \in \mathbb{N}$ and $0 < p < 1$, and let $\delta(\epsilon_1)$, $\epsilon_1 \in \mathbb{R}$, define the privacy profile of the base mechanism Q. Suppose*

$$\epsilon_1 \geq \log\left(1 + \tfrac{p}{1-p} \cdot \delta(\epsilon_1)\right).$$

Then, for A and A', the output distributions of the selection algorithm evaluated on neighboring datasets X and X', respectively, for all $\epsilon > 0$ and for all $\epsilon_1 \geq 0$,

$$H_{\mathrm{e}^{\epsilon}}(A||A') \leq m \cdot \delta(\widehat{\epsilon}), \tag{5.2}$$

where

$$\widehat{\epsilon} = \epsilon - (n-1)\log\left(1 + p \cdot (\mathrm{e}^{\epsilon_1} - 1) + p \cdot \delta(\epsilon_1)\right).$$

We get the hockey stick divergence bound for the case $K \sim \mathrm{Poisson}(m)$ as a corollary of Thm. 5.7.

Corollary 5.8. *Let $K \sim \mathrm{Poisson}(m)$ for some $m \in \mathbb{N}$, and let $\delta(\epsilon_1)$, $\epsilon_1 \in \mathbb{R}$, define the privacy profile of the base mechanism Q. Then, for A and A', the output distributions of the selection algorithm evaluated on neighboring datasets X and X', respectively, and for all $\epsilon > 0$ and for all $\epsilon_1 \geq 0$,*

$$H_{\mathrm{e}^{\epsilon}}(A||A') \leq m \cdot \delta(\widehat{\epsilon}), \tag{5.3}$$

where

$$\widehat{\epsilon} = \epsilon - m \cdot (\mathrm{e}^{\epsilon_1} - 1) - m \cdot \delta(\epsilon_1).$$

The proof of Cor. 5.8 essentially follows from the fact that $\mathrm{Bin}(n, m/n)$ approaches $\mathrm{Poisson}(m)$ in total variation distance as n grows and that the bound (5.2) approaches the bound (5.3) as n grows. Figure 4 illustrates that when compared to the RDP bound of Thm. 5.6 for the Poisson distributed K, we can obtain much smaller probabilities for small values of K for the same privacy cost when using $K \sim \mathrm{Bin}(n, m/n)$ and Thm. 5.7.

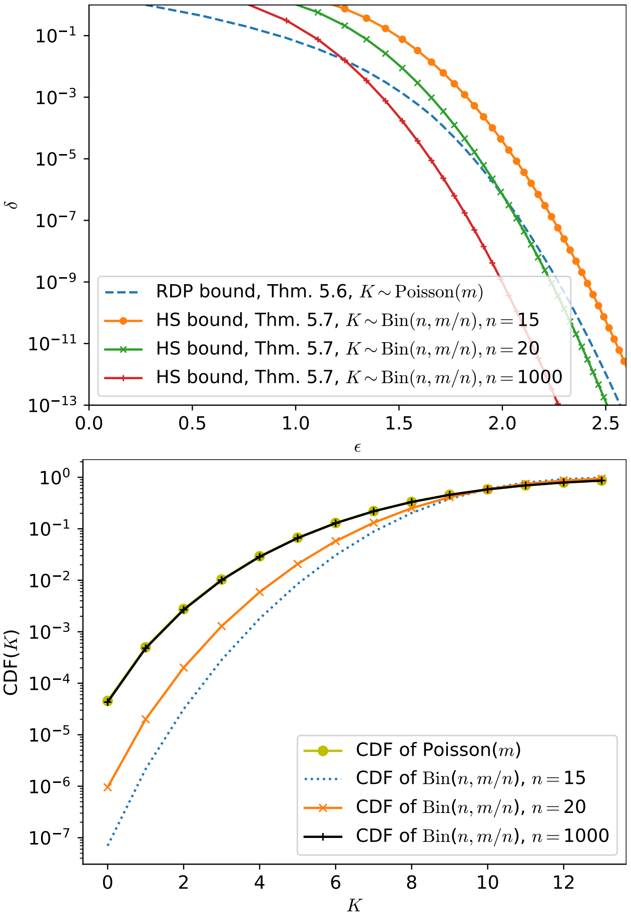

Figure 4. Top: Comparison of the bound of Thm. 5.7 for $K \sim \mathrm{Bin}(n, m/n)$ for different values of n, when $m = 10$, and the RDP bound of Thm. 5.6. The base mechanism is the Gaussian mechanism with sensitivity 1 and $\sigma = 4.0$. Bottom: Comparison of the CDFs for different values of n. When comparing to the RDP bound, we see that at $\delta \approx 10^{-6}$ we get more concentrated K for free by using the binomial distribution and Thm. 5.7.

6. Applications

6.1. Hyperparameter tuning for Propose-Test-Release

Propose-Test-Release (PTR) (Dwork & Lei, 2009) is one of the most versatile recipes for data-adaptive DP mechanism design. In its vanilla form, it involves three steps: (1) propose (or guess) a bound of the local sensitivity; (2) privately test whether the bound is valid; (3) If it passes the test, calibrate noise proportional to the proposed bound; otherwise, refuse to answer. Recently, Redberg et al. (2023) generalized this approach by considering data-adaptive privacy losses instead. The biggest challenge to apply the method is to know which bound to propose. The data-dependent privacy loss will depend on both the dataset and the hyperparameters of the query (as an example, think of the noise scale in additive noise mechanisms). To tune these hyperparameters we consider the private selection algorithm with geometrically distributed K. The tricky issue with PTR is that PTR does not satisfy Rényi DP, and thus disqualifies the approach from Papernot & Steinke (2022). Meanwhile, our methods deal with (ε, δ)-DP and δ-approximate Gaussian DP very naturally. Using our Thm. 5.2 we can select the best threshold to propose in a large number of candidates without resorting to composition. We have the following result for Generalized PTR with an $(\widehat{\epsilon}, \widehat{\delta})$-DP test. We refer to Redberg et al. (2023) for more details.

Theorem 6.1 (Redberg et al. 2023)**.** *Consider a proposal ϕ and a data-dependent function $\epsilon_\phi(X)$ w.r.t. $\delta > 0$. Suppose that we have an $(\widehat{\epsilon}, \widehat{\delta})$-DP test $\mathcal{T} : \mathcal{X} \to \{0, 1\}$ such that when $\epsilon_\phi(X) > \epsilon$, $\mathcal{T}(X) = 1$ with probability δ' and 0 with probability $1 - \delta'$. Then the Generalized PTR algorithm (Redberg et al., 2023, Alg. 2) is $(\epsilon + \widehat{\epsilon}, \delta + \widehat{\delta} + \delta')$-DP.*

Our approach is to wrap the private selection algorithm around generalized PTR and tune the parameter ϕ. We use the point-wise guarantees given by Thm. 6.1 for Gen. PTR and our Thm. 5.1 for the tuning algorithm. To illustrate the effectiveness of this approach, we consider a linear regression problem on two UCI benchmark data sets (Bache & Lichman, 2013), see Fig. 5. We apply the Generalized PTR to the one-posterior sample (OPS) algorithm described in (Redberg et al., 2023) which includes privately releasing the L_2-norm of the non-private solution and also the smallest eigenvalue of the feature covariance matrix. The parameter to tune in the method is the regularization strength λ (see Alg. 7, Redberg et al., 2023) and we carry out a random search on a pre-defined logarithmically equidistant grid meaning that we pick a random value from the grid at each of the K rounds. Notice that we could draw the candidates from any fixed probability distribution; the only requirement is that each candidate mechanism has the same privacy profile. As baselines we have the same approach using the privacy bounds of Liu & Talwar (Thm. 3.5, 2019), the output perturbation method (Chaudhuri et al., 2011) and the non-adaptive method OPS-Balanced by Wang (2018).

6.2. Private Hyperparameter Tuning of DP-SGD

Our results enable using numerical accountants for computing the privacy profile $\delta(\epsilon)$ for the subsampled Gaussian mechanism (see, e.g., Koskela et al., 2021; Gopi et al., 2021; Zhu et al., 2022). We consider the simplest case, i.e., the Poisson subsampling and the add/remove neighborhood relation of datasets. We apply the numerical method proposed by Koskela et al. (2021) to the dominating pairs given in (Thm. 11, Zhu et al., 2022) to obtain accurate privacy profiles for the base mechanism. We remark that Zhu et al. (2022) give privacy profiles for various subsampling schemes under both add/remove and substitute neighborhood relations of datasets. With these results, one can numer-

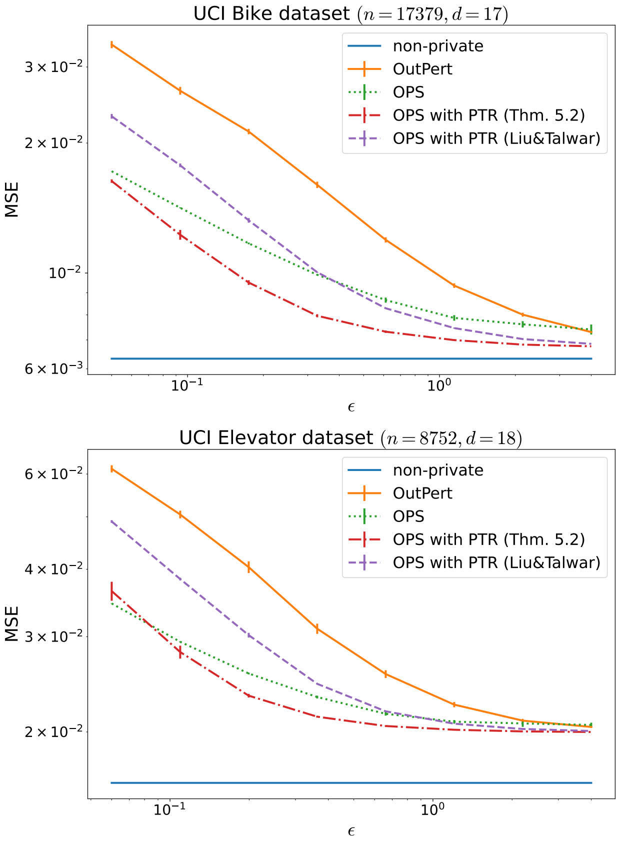

Figure 5. Linear regression problem on two UCI benchmark datasets. Tuning OPS-PTR (i.e., generalized PTR applied to the one-posterior sample algorithm) via the private selection algorithm outperforms baseline methods when the privacy cost of the tuning procedure is calculated using our Thm. 5.2.

ically construct the dominating pair of distributions using methods of Doroshenko et al. (2022) and also obtain upper bounds for compositions.

We find that our bounds are tighter than the RDP bounds across a variety of parameter combinations. Figure 6 shows comparisons with parameters taken from an example of (Papernot & Steinke, 2022). RDP parameters are evaluated using the Opacus library (Yousefpour et al., 2021). Often, using larger batch sizes and noise ratios leads to increased privacy-utility tradeoff (De et al., 2022; Ponomareva et al., 2023). Figure 7 shows comparisons in a setting of such a high-accuracy experiment (De et al., 2022).

Adjusting Hyperparameters. One important question is, how to adjust the hyperparameters in case we are optimizing parameters that affect the privacy guarantees themselves. For example, one may consider tuning the noise parameter σ in DP-SGD in which case one may also consider adjusting the length of the training.

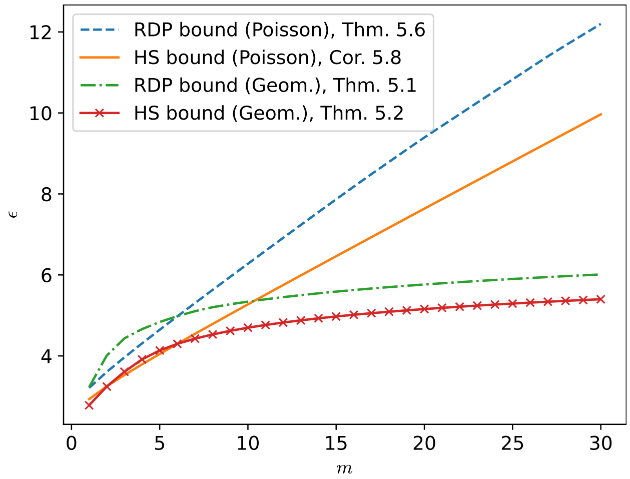

Figure 6. Comparison of the hockey stick divergence bounds of Thm. 5.2 and Cor. 5.8 and the RDP bounds of Thm. 5.1 and 5.6 for the private selection for a Poisson subsampled Gaussian mechanism with subsampling ratio $q = 256/60000$, noise parameter $\sigma = 1.1$ and number of steps $T = \lceil 60/q \rceil$.

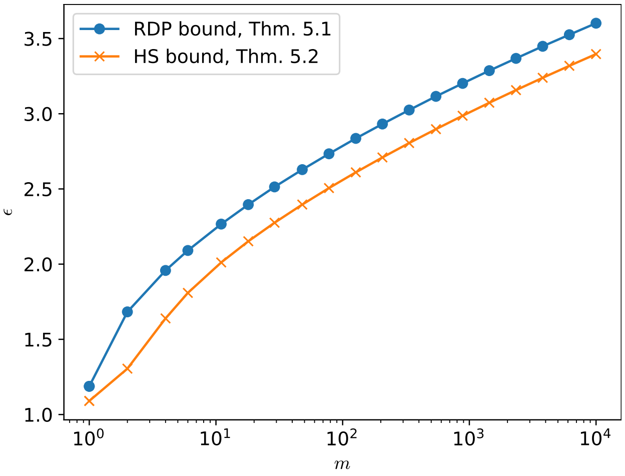

Figure 7. Comparison of the bounds, when the base mechanism is a Poisson subsampled Gaussian mechanism with the parameters $q = 16384/50000$, $\sigma = 21.1$ and $T = 250$ (see Table 18 in De et al., 2022). We see that the bound of Thm. 5.2 allows evaluating approximately 3 times as many models as the RDP bound.

The bound of Thm. 5.2 essentially requires only point-wise information about the privacy profile of the base mechanism Q. However, the bound can be optimized w.r.t. the privacy profile of Q which may affect it considerably. For finding ideal point-wise DP thresholds, we consider the following procedure. Suppose the base mechanisms all satisfy (ϵ_Q, δ) for some $\epsilon_Q > 0$ and $\delta > 0$. As the privacy profiles start to resemble those of the Gaussian mechanism for large numbers of compositions (Dong et al., 2022), we carry out the optimization of the upper bound of Thm. 5.2 w.r.t. to the privacy profile of a Gaussian mechanism (GM) that is adjusted to be (ϵ_Q, δ)-DP. More specifically, we first adjust the noise scale σ of the GM such that it is (ϵ_Q, δ)-DP, to obtain a privacy profile $\delta_Q(\epsilon)$. Then, in case we are using $K \sim \mathcal{D}_{\eta,\gamma}$, we carry out the minimization

$$\epsilon_1 = \arg\min_{\epsilon} \log\left(\mathrm{e}^{\epsilon} + \tfrac{1-\gamma}{\gamma}\delta_Q(\epsilon)\right).$$

and set $\delta_1 = \delta_Q(\epsilon_1)$. This results in the first threshold (ϵ_1, δ_1). Using the same privacy profile we find an $\widehat{\epsilon}$-value where this GM is $(\widehat{\epsilon}, \delta/m)$-DP. The following corollary result of Thm. 5.2 then gives the ϵ-value for which the private selection algorithm is (ϵ, δ)-DP.

Corollary 6.2. *Suppose the base mechanism Q is (ϵ_1, δ_1)-DP and $(\widehat{\epsilon}, \delta/m)$-DP for some $\epsilon_1 \geq 0$ and $\widehat{\epsilon} \geq 0$. Then, the private selection algorithm with $K \sim \mathcal{D}_{\eta,\gamma}$ is (ϵ, δ)-DP for*

$$\epsilon = \widehat{\epsilon} + (\eta + 1)\log\left(\mathrm{e}^{\epsilon_1} + \tfrac{1-\gamma}{\gamma}\delta_1\right).$$

Figure 8 shows the upper bound obtained using this procedure, when we are tuning the σ-parameter for the Poisson subsampled Gaussian mechanism. We fix $q = 0.01$ and set as a threshold $\epsilon_Q = 1.5$ and $\delta = 10^{-6}$. We consider three σ-candidates: 2.0, 3.0 and 4.0 and for each of them the number of iterations T is determined to be the maximum such that the privacy profile of the candidate is below (ϵ_1, δ_1)- and $(\widehat{\epsilon}, \delta/m)$-thresholds. As a result we can run the candidate models for ≈ 4000, 10000 and 18000 iterations, respectively. We compare the (ϵ, δ)-DP bound given by Cor. 6.2 and the bounds we would obtain by optimizing Thm. 5.2 individually for each candidate mechanism and see that there is a very small gap.

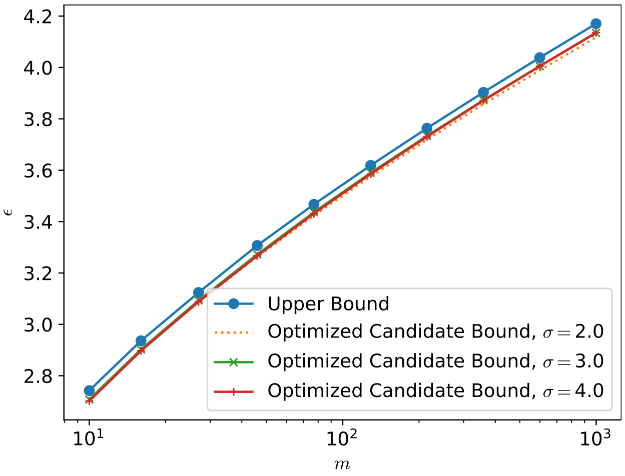

Figure 8. Tuning the σ-parameter for the Poisson subsampled Gaussian mechanism with the geometrically distributed K. We fix $q = 0.01$ and consider three σ-candidates: 2.0, 3.0 and 4.0. Shown are the (ϵ, δ)-bound given by Cor. 6.2 and the bounds obtained by optimizing Thm. 5.2 individually for each candidate mechanism.

7. Conclusions and Future Work

We have filled a gap in the private selection literature by providing (ϵ, δ)-privacy analysis for various selection algorithms that is able to accurately use the privacy profiles of the candidate mechanisms. Our bounds can be used as a drop-in replacement of the existing RDP bounds. When compared to existing RDP bounds, in DP-SGD tuning, for example, the new bounds allow evaluating approximately 3 times as many candidate models. The bounds also improve existing point-wise (ϵ, δ)-bounds which translates to improved utility in data-adaptive analyses using the generalized PTR framework. We have also shown how to use the bounds to adjust parameters of the candidate models when tuning hyperparameters that affect the privacy guarantees of the candidate mechanisms.

Related to the results of Section 3, it is an open problem how to find tight bounds for a given noise-adding mechanism in case of unbounded queries. It is fairly straightforward to find accurate numerical bounds for a fixed pair of datasets and a given noise-adding mechanism, but it is unclear what would generally be the worst-case pairs of query values in case of unbounded queries (for results in case of bounded queries, see Lebensold et al., 2024). Regarding the results of Section 4 and 5, our impression is that when the number of candidates is randomized, our bounds cannot be much tightened. However, one could consider different randomized selection algorithms such as those by Cohen et al. (2023) where the number of candidates is Beta-binomially distributed. It is an interesting question whether our general result of Thm. 4.1 could be then used to derive tighter privacy bounds. Note that all our results are non-data adaptive in the sense that they do not benefit, for example, from the case that one of the candidates is clearly better than the rest. As an alternative, one could consider data-adaptive mechanisms such as Bayesian optimization methods (Wang et al., 2023) or data-adaptive top-k selection mechanisms (Zhu & Wang, 2022).

Impact Statement

This paper presents work whose goal is to advance the field of Differentially Private Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

Acknowledgments

The work was partially supported by NSF Award #2048091. The authors thank Yuqing Zhu for sharing code from Redberg et al. (2023).

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., and Zhang, L. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 308–318, 2016.

Bache, K. and Lichman, M. UCI machine learning repository, 2013. URL http://archive.ics.uci.edu/ml.

Balle, B. and Wang, Y.-X. Improving the Gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. In *International Conference on Machine Learning*, pp. 394–403. PMLR, 2018.

Balle, B., Barthe, G., and Gaboardi, M. Privacy amplification by subsampling: Tight analyses via couplings and divergences. *Advances in Neural Information Processing Systems*, pp. 6277–6287, 2018.

Boyd, S. P. and Vandenberghe, L. *Convex optimization*. Cambridge university press, 2004.

Canonne, C. L., Kamath, G., and Steinke, T. The discrete gaussian for differential privacy. *Advances in Neural Information Processing Systems*, 33:15676–15688, 2020.

Chaudhuri, K. and Vinterbo, S. A. A stability-based validation procedure for differentially private machine learning. *Advances in Neural Information Processing Systems*, 26, 2013.

Chaudhuri, K., Monteleoni, C., and Sarwate, A. D. Differentially private empirical risk minimization. *J. Mach. Learn. Res.*, 12:1069–1109, 2011.

Cohen, E., Lyu, X., Nelson, J., Sarlós, T., and Stemmer, U. Generalized private selection and testing with high confidence. In *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik, 2023.

De, S., Berrada, L., Hayes, J., Smith, S. L., and Balle, B. Unlocking high-accuracy differentially private image classification through scale. *arXiv preprint arXiv:2204.13650*, 2022.

Dong, J., Roth, A., and Su, W. J. Gaussian differential privacy. *Journal of the Royal Statistical Society Series B*, 84(1):3–37, 2022.

Doroshenko, V., Ghazi, B., Kamath, P., Kumar, R., and Manurangsi, P. Connect the dots: Tighter discrete approximations of privacy loss distributions. *arXiv preprint arXiv:2207.04380*, 2022.

Dwork, C. Differential privacy. In *Proc. 33rd Int. Colloq. on Automata, Languages and Prog. (ICALP 2006), Part II*, pp. 1–12, 2006.

Dwork, C. and Lei, J. Differential privacy and robust statistics. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pp. 371–380, 2009.

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

Gopi, S., Lee, Y. T., and Wutschitz, L. Numerical composition of differential privacy. *Advances in Neural Information Processing Systems*, 34:11631–11642, 2021.

Koskela, A., Jälkö, J., Prediger, L., and Honkela, A. Tight differential privacy for discrete-valued mechanisms and for the subsampled Gaussian mechanism using FFT. In *International Conference on Artificial Intelligence and Statistics*, pp. 3358–3366. PMLR, 2021.

Lebensold, J., Precup, D., and Balle, B. On the privacy of selection mechanisms with gaussian noise. pp. 1495–1503, 2024.

Liese, F. and Vajda, I. On divergences and informations in statistics and information theory. *IEEE Transactions on Information Theory*, 52(10):4394–4412, 2006.

Liu, J. and Talwar, K. Private selection from private candidates. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pp. 298–309, 2019.

Mironov, I. Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pp. 263–275, Aug 2017. doi: 10.1109/CSF.2017.11.

Mohapatra, S., Sasy, S., He, X., Kamath, G., and Thakkar, O. The role of adaptive optimizers for honest private hyperparameter selection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 7806–7813, 2022.

Papernot, N. and Steinke, T. Hyperparameter tuning with renyi differential privacy. In *International Conference on Learning Representations*, 2022.

Ponomareva, N., Hazimeh, H., Kurakin, A., Xu, Z., Denison, C., McMahan, H. B., Vassilvitskii, S., Chien, S., and Thakurta, A. G. How to dp-fy ml: A practical guide to machine learning with differential privacy. *Journal of Artificial Intelligence Research*, 77:1113–1201, 2023.

Redberg, R., Zhu, Y., and Wang, Y.-X. Generalized PTR: User-friendly recipes for data-adaptive algorithms with differential privacy. In *International Conference on Artificial Intelligence and Statistics*, pp. 3977–4005. PMLR, 2023.

Tang, X., Shin, R., Inan, H. A., Manoel, A., Mireshghallah, F., Lin, Z., Gopi, S., Kulkarni, J., and Sim, R. Privacy-preserving in-context learning with differentially private few-shot generation. In *International Conference on Learning Representations*, 2024.

Wang, H., Gao, S., Zhang, H., Su, W. J., and Shen, M. DP-HyPo: An adaptive private framework for hyperparameter optimization. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Wang, Y.-X. Revisiting differentially private linear regression: optimal and adaptive prediction & estimation in unbounded domain. *Uncertainty in Artificial Intelligence (UAI-18)*, 2018.

Wu, T., Panda, A., Wang, J. T., and Mittal, P. Privacy-preserving in-context learning for large language models. In *International Conference on Learning Representations*, 2024.

Yousefpour, A., Shilov, I., Sablayrolles, A., Testuggine, D., Prasad, K., Malek, M., Nguyen, J., Ghosh, S., Bharadwaj, A., Zhao, J., et al. Opacus: User-friendly differential privacy library in pytorch. In *NeurIPS 2021 Workshop Privacy in Machine Learning*, 2021.

Zhu, Y. and Wang, Y.-X. Adaptive private-k-selection with adaptive k and application to multi-label pate. In *International Conference on Artificial Intelligence and Statistics*, pp. 5622–5635. PMLR, 2022.

Zhu, Y., Dong, J., and Wang, Y.-X. Optimal accounting of differential privacy via characteristic function. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, 2022.

A. Proofs for Section 3

A.1. Proof of Theorem 3.2 (Privacy Profile of Additive Noise RNM)

Theorem A.1. *Let $X \sim X'$ and $\epsilon \in \mathbb{R}$. We have*

$$H_{e^\epsilon}\big(\mathcal{M}(X) || \mathcal{M}(X')\big) \leq m \cdot \delta(\epsilon),$$

where $\delta(\epsilon)$ is the privacy profile of the additive noise mechanism with sensitivity 2.

Proof. For each $i \in [m]$, denoting the density function of Z_i by $p(r_i)$, we have that

$$\begin{aligned}
&H_f\big(\mathcal{M}(X)||\mathcal{M}(X')\big) \\
&= \sum_{i=1}^m f\left(\frac{\mathbf{P}(\mathcal{M}(X) = i)}{\mathbf{P}(\mathcal{M}(X') = i)}\right) \mathbf{P}(\mathcal{M}(X') = i) \\
&= \sum_{i=1}^m f\left(\frac{\int_{-\infty}^{\infty} p(r_i - 2) \cdot \mathbf{P}(f_i(X) + r_i - 2 > \max_{j \in [m], j \neq i}\{f_j(X) + r_j\}) \, \mathrm{d}r_i}{\int_{-\infty}^{\infty} p(r_i) \cdot \mathbf{P}(f_i(X') + r_i > \max_{j \in [m], j \neq i}\{f_j(X') + r_j\}) \, \mathrm{d}r_i}\right) \\
&\qquad \cdot \int_{-\infty}^{\infty} p(r_i) \cdot \mathbf{P}(f_i(X') + r_i > \max_{j \in [m], j \neq i}\{f_j(X') + r_j\}) \, \mathrm{d}r_i,
\end{aligned} \tag{A.1}$$

where $p(r_i)$'s are the density functions of Z_i's, respectively, and where the randomness in $\mathbf{P}(f_i(X) + r_i > \max_{j \in [m], j \neq i}\{f_j(X) + r_j\})$ is w.r.t. r_j's.

From the Lipschitz property it follows that for all $i \in [m]$,

$$\frac{\mathbf{P}(f_i(X) + r_i - 2 > \max_{j \in [m], j \neq i}\{f_j(X) + r_j\})}{\mathbf{P}(f_i(X') + r_i > \max_{j \in [m], j \neq i}\{f_j(X') + r_j\})} \leq 1. \tag{A.2}$$

For an f divergence determined by a convex f of the RNM of additive noise mechanisms evaluated at X and X', we have that

$$\begin{aligned}
H_f\big(\mathcal{M}(X)||\mathcal{M}(X')\big) &= \sum_{i=1}^m f\left(\frac{\mathbf{P}(\mathcal{M}(X) = i)}{\mathbf{P}(\mathcal{M}(X') = i)}\right) \mathbf{P}(\mathcal{M}(X') = i) \\
&= \sum_{i=1}^m f\left(\frac{\int_{-\infty}^{\infty} p(r_i - 2) \cdot \mathbf{P}(f_i(X) + r_i - 2 > \max_{j \in [m], j \neq i}\{f_j(X) + r_j\}) \, \mathrm{d}r_i}{\int_{-\infty}^{\infty} p(r_i) \cdot \mathbf{P}(f_i(X') + r_i > \max_{j \in [m], j \neq i}\{f_j(X') + r_j\}) \, \mathrm{d}r_i}\right) \\
&\qquad \cdot \int_{-\infty}^{\infty} p(r_i) \cdot \mathbf{P}(f_i(X') + r_i > \max_{j \in [m], j \neq i}\{f_j(X') + r_j\}) \, \mathrm{d}r_i \\
&\leq \sum_{i=1}^m \int_{-\infty}^{\infty} f\left(\frac{p(r_i - 2) \cdot \mathbf{P}(f_i(X) + r_i - 2 > \max_{j \in [m], j \neq i}\{f_j(X) + r_j\})}{p(r_i) \cdot \mathbf{P}(f_i(X') + r_i > \max_{j \in [m], j \neq i}\{f_j(X') + r_j\})}\right) \\
&\qquad \cdot p(r_i) \cdot \mathbf{P}(f_i(X') + r_i > \max_{j \in [m], j \neq i}\{f_j(X') + r_j\}) \, \mathrm{d}r_i \\
&\leq \sum_{i=1}^m \int_{-\infty}^{\infty} f\left(\frac{p(r_i - 2)}{p(r_i)}\right) \cdot p(r_i) \cdot \mathbf{P}(f_i(X') + r_i > \max_{j \in [m], j \neq i}\{f_j(X') + r_j\}) \, \mathrm{d}r_i \\
&\leq \sum_{i=1}^m \int_{-\infty}^{\infty} f\left(\frac{p(r_i - 2)}{p(r_i)}\right) \cdot p(r_i) \, \mathrm{d}r_i \\
&= \sum_{i=1}^m H_f\big(p(r_i)||p(r_i - 2)\big) \\
&= m \cdot \delta(\epsilon),
\end{aligned} \tag{A.3}$$

in case f corresponds to the hockey-stick-divergence and where $\delta(\epsilon)$ is the privacy profile of the additive noise mechanism with sensitivity 2. In the first inequality we have used Lemma 2.3 and Jensen's inequality term-wise and in the second inequality we have used the Lipschitz property and the fact that $f(z)$ is a non-decreasing function of z. In case of monotonicity, i.e., if $f_i(X) \geq f_i(X')$ for all $i \in [m]$, the condition (A.2) holds with $r_i - 2$ replaced by $r_i - 1$ and we have the result with $r_i - 2$ replaced by $r_i - 1$. □

A.2. Proof of Theorem 3.3

Lemma A.2. *In case of an adaptive composition of k mechanisms of the form (3.1), we get the privacy profile upper bound $m^k \cdot \delta(\epsilon)$, where $\delta(\epsilon)$ is the privacy profile of a m-wise composition of the additive noise mechanism with noise Z and sensitivity 2.*

Proof. We use the proof technique used in (Thm. 27 Zhu et al., 2022) and consider an adaptive composition of two mechanisms. The general case follows from the proof. Let $X \sim X'$. Denote the density functions of $\mathcal{M}^1(X)$ and $\mathcal{M}^2\big(\mathcal{M}^1(X), X\big)$ by $f_1(t)$ and $f_2(t,s)$, respectively, and the density functions of $\mathcal{M}^1(X')$ and $\mathcal{M}^2\big(\mathcal{M}^1(X'), X'\big)$ by $f_1'(t)$ and $f_2'(t,s)$, respectively. Denote the density function of Z by $p(t)$. Using the bound given by Thm. 3.2, we have:

$$\begin{aligned}
H_{\mathrm{e}^\epsilon}\big((\mathcal{M}(X),\mathcal{M}(X))||(\mathcal{M}(X),\mathcal{M}(X))\big) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \max\{f_1(t)f_2(t,s) - \mathrm{e}^\epsilon f_1'(t)f_2'(t,s), 0\}\,\mathrm{d}s\mathrm{d}t \\
&= \int_{\mathbb{R}} f_1(t)\left(\int_{\mathbb{R}} \max\{f_2(t,s) - \mathrm{e}^{\epsilon - \log\frac{f_1(t)}{f_1'(t)}} f_2'(t,s), 0\}\,\mathrm{d}s\right)\mathrm{d}t \\
&\leq m \cdot \int_{\mathbb{R}} f_1(t)\left(\int_{\mathbb{R}} \max\{p(s-2) - \mathrm{e}^{\epsilon - \log\frac{f_1(t)}{f_1'(t)}} p(s), 0\}\,\mathrm{d}s\right)\mathrm{d}t \\
&= m \cdot \int_{\mathbb{R}} p(s-2)\left(\int_{\mathbb{R}} \max\{f_1(t) - \mathrm{e}^{\epsilon - \log\frac{p(s-2)}{p(s)}} f_1'(t), 0\}\,\mathrm{d}t\right)\mathrm{d}s \\
&\leq m^2 \cdot \int_{\mathbb{R}} p(s-2)\left(\int_{\mathbb{R}} \max\{p(t-2) - \mathrm{e}^{\epsilon - \log\frac{p(s-2)}{p(s)}} p(t), 0\}\,\mathrm{d}t\right)\mathrm{d}s \\
&= m^2 \cdot \int_{\mathbb{R}}\int_{\mathbb{R}} \max\{p(s-2)p(t-2) - \mathrm{e}^\epsilon p(s)p(t), 0\}\,\mathrm{d}t\mathrm{d}s
\end{aligned}$$

which shows the claim for $k = 2$. The general case follows by induction. □

A.3. Proof of Corollary 3.4

Lemma A.3. *Consider the mechanism $\mathcal{M}$ defined in Eq. 3.1 and suppose Z is normally distributed with variance σ^2. Let $\delta > 0$. Then $\mathcal{M}$ is (ϵ, δ)-DP for*

$$\epsilon = \frac{2}{\sigma^2} + \frac{2}{\sigma}\sqrt{2\log\frac{m}{\delta}}$$

Proof. Let $\delta > 0$. By (Lemma 3, Balle & Wang, 2018) we know that the Gaussian mechanism with L_2-sensitivity Δ and noise scale σ is (ϵ, δ)-DP if

$$\mathbb{P}(\omega \geq \epsilon) \leq \delta,$$

where

$$\omega \sim \mathcal{N}\left(\frac{\Delta^2}{2\sigma^2}, \frac{\Delta^2}{\sigma^2}\right).$$

Using a simple Chernoff bound for the Gaussian, we see that for any $\epsilon \geq \frac{\Delta^2}{2\sigma^2}$,

$$\mathbb{P}(\omega \geq \epsilon) \leq \mathrm{e}^{-\frac{\widetilde{\epsilon}^2\sigma^2}{2}},$$

where $\widetilde{\epsilon} = \epsilon - \frac{\Delta^2}{2\sigma^2}$. Setting

$$\epsilon = \tfrac{\Delta^2}{2\sigma^2} + \tfrac{\Delta}{\sigma}\sqrt{2\log\tfrac{m}{\delta}},$$

we see that

$$\mathbb{P}(\omega \geq \epsilon) \leq \frac{1}{m}.$$

The claim follows setting $\Delta = 2$ and using Thm. 3.2.

□

B. Proof of Theorem 4.1 (General Bound for the Privacy Selection)

Similarly to (Lemma 7, Papernot & Steinke, 2022) we denote $Q(\leq y) = \sum_{y' \leq y} Q(y')$ and $Q(< y) = \sum_{y' < y} Q(y')$ and formulate our main result for the case $\mathbb{P}(K = 0) = 0$. The case $\mathbb{P}(K = 0) > 0$ could be included in the upper bound as a small additional term of the form $f(0) \cdot \mathbb{P}(K = 0)$. E.g., in case of the hockey stick divergence, it does not account for the divergence in case $\epsilon > 0$ so we neglect it.

Also, φ denotes the probability generating function of K, i.e.,

$$\varphi(z) = \sum_{k=1}^{\infty} \mathbb{P}(K = k) \cdot z^k.$$

As shown in (Lemma 7, Papernot & Steinke, 2022),

$$\begin{aligned} A(y) &= \sum_{k=1}^{\infty} \mathbf{P}[K = k] \cdot \big(Q(\leq y)^k - Q(< y)^k\big) \\ &= \varphi\big(Q(\leq y)\big) - \varphi\big(Q(< y)\big) \\ &= \int_{Q(<y)}^{Q(\leq y)} \varphi'(z) \, \mathrm{d}z \\ &= Q(y) \cdot \mathop{\mathbb{E}}_{X \leftarrow [Q(<y), Q(\leq y)]} [\varphi'(X)], \end{aligned} \tag{B.1}$$

where $X \leftarrow [Q(< y), Q(\leq y)]$ denotes uniformly distributed r.v. on the interval $[Q(< y), Q(\leq y)]$.

Theorem B.1. *Let $X \sim X'$ and let A and A' be the density functions of the hyperparameter tuning algorithm, evaluated on X and X', respectively. Let Q and Q' be the density functions of the quality score of the base mechanism, evaluated on X and X', respectively. Let K be random variable for the times the base mechanism is run and $\varphi(z)$ the PGF of K. Let $f : [0, \infty) \to \mathbb{R}$ be a convex function. Then,*

$$H_f(A||A') \leq \sum_{y \in \mathcal{Y}} f\left(\frac{Q(y)\varphi'(q_y)}{Q'(y)\varphi'(q'_y)}\right) \cdot Q'(y)\varphi'(q'_y),$$

where for each $y \in \mathcal{Y}$, q_y and q'_y are obtained by applying the same y-dependent post-processing function to Q and Q', respectively.

Proof. For the mechanism A defined in (B.1), we can bound the HS divergence as follows:

$$\begin{aligned} H_f(A||A') &= \sum_{y \in \mathcal{Y}} f\left(\frac{A(y)}{A'(y)}\right) \cdot A(y) \\ &= \sum_{y \in \mathcal{Y}} f\left(\frac{Q(y)\,\mathbb{E}_{X \leftarrow [Q(<y),Q(\leq y)]}[\varphi'(X)]}{Q'(y)\,\mathbb{E}_{X' \leftarrow [Q'(<y),Q'(\leq y)]}[\varphi'(X')]}\right) \cdot Q'(y) \mathop{\mathbb{E}}_{X' \leftarrow [Q'(<y),Q'(\leq y)]} [\varphi'(X')] \\ &\leq \sum_{y \in \mathcal{Y}} \mathop{\mathbb{E}}_{X \leftarrow [Q(<y),Q(\leq y)],\; X' \leftarrow [Q'(<y),Q'(\leq y)]} f\left(\frac{Q(y)\varphi'(X)}{Q'(y)\varphi'(X')}\right) \cdot Q'(y)\varphi'(X') \\ &\leq \sum_{y \in \mathcal{Y}} \max_{y'} \mathop{\mathbb{E}}_{X \leftarrow [Q(<y'),Q(\leq y')],\; X' \leftarrow [Q'(<y'),Q'(\leq y')]} f\left(\frac{Q(y)\varphi'(X)}{Q'(y)\varphi'(X')}\right) \cdot Q'(y)\varphi'(X'), \end{aligned} \tag{B.2}$$

where in the first inequality we use Lemma 2.3 and Jensen's inequality. Notice that in the second inequality the maximum is taken only over the arguments in the expectation over X and X'.

Jensen's inequality applies in case X and X' are arbitrarily coupled and we use the same coupling between X and X' as in (Lemma 7, Papernot & Steinke, 2022), i.e., we couple X and X' such that

$$\frac{X - Q(< y)}{Q(y)} = \frac{X' - Q'(< y)}{Q'(y)}. \tag{B.3}$$

We see that for $X \leftarrow [Q(< y'), Q(\leq y')]$ and $X' \leftarrow [Q'(< y'), Q'(\leq y')]$, the expressions (B.3) are between 0 and 1. Thus, continuing from (B.2), we find that

$$\begin{aligned} H_f(A||A') &\leq \sum_{y \in \mathcal{Y}} \max_{y'} \underset{X \leftarrow [Q(<y'),Q(\leq y')],\ X' \leftarrow [Q'(<y'),Q'(\leq y')]}{\mathbb{E}} f\left(\frac{Q(y)\varphi'(X)}{Q'(y)\varphi'(X')}\right) \cdot Q'(y)\varphi'(X') \\ &\leq \sum_{y \in \mathcal{Y}} \max_{y'} \max_{(X,X')} f\left(\frac{Q(y)\varphi'(X)}{Q'(y)\varphi'(X')}\right) \cdot Q'(y)\varphi'(X') \\ &= \sum_{y \in \mathcal{Y}} \max_{y'} f\left(\frac{Q(y)\varphi'(Q(< y') + t_y \cdot Q(y'))}{Q'(y)\varphi'(Q'(< y') + t_y \cdot Q'(y'))}\right) \cdot Q'(y)\varphi'(Q'(< y') + t_y \cdot Q'(y')) \end{aligned}$$

for some $\{t_y\}_{y \in \mathcal{Y}}$, where $t_y \in [0, 1]$ for all $y \in \mathcal{Y}$. Furthermore, taking the maximum over t_y's, we get

$$\begin{aligned} H_f(A||A') &\leq \sum_{y \in \mathcal{Y}} \max_{y', t_y} f\left(\frac{Q(y)\varphi'(Q(< y') + t_y \cdot Q(y'))}{Q'(y)\varphi'(Q'(< y') + t_y \cdot Q'(y'))}\right) \cdot Q'(y)\varphi'(Q'(< y') + t_y \cdot Q'(y')) \\ &\leq \sum_{y \in \mathcal{Y}} f\left(\frac{Q(y)\varphi'(q_y)}{Q'(y)\varphi'(q'_y)}\right) \cdot Q'(y)\varphi'(q'_y), \end{aligned}$$

where for each $y \in \mathcal{Y}$, q_y and q'_y are obtained by applying the same y-dependent post-processing function to Q and Q', respectively. This can be seen using a similar reasoning as in (Lemma 7, Papernot & Steinke, 2022). It follows from the fact that for all $y \in \mathcal{Y}$, there clearly exist y_*, t_* such that

$$(y_*, t_*) = \underset{y', t_y}{\operatorname{argmax}} f\left(\frac{Q(y)\varphi'(Q(< y') + t_y \cdot Q(y'))}{Q'(y)\varphi'(Q'(< y') + t_y \cdot Q'(y'))}\right) \cdot Q'(y)\varphi'(Q'(< y') + t_y \cdot Q'(y')).$$

The kernel of the post-processing function is then given by

$$g(z) = \begin{cases} 1, & \text{if } z < y_*, \\ t_*, & \text{if } z = y_*, \\ 0, & \text{else,} \end{cases}$$

i.e., $q_y = \sum_{z \in \mathcal{Y}} g(z)Q(z)$ and $q'_y = \sum_{z \in \mathcal{Y}} g(z)Q'(z)$. □

B.1. The Case of Continuous Output Score Function

In case the ordered output space $\mathcal{Y}$ of the base mechanism is continuous, the proof simplifies considerably.

Theorem B.2. *For continuous output space $\mathcal{Y}$,*

$$H_f(A||A') = \int_{\mathcal{Y}} f\left(\frac{Q(y) \cdot \varphi'\big(Q(\leq y)\big)}{Q'(y) \cdot \varphi'\big(Q'(\leq y)\big)}\right) \cdot Q'(y) \cdot \varphi'\big(Q'(\leq y)\big)\, \mathrm{d}y, \tag{B.4}$$

where where $Q(y)$ and $Q(\leq y)$ denote the density function the CDF of the base mechanism Q, respectively.

Proof. The CDF of the private selection algorithm at $y \in \mathcal{Y}$ is given by

$$A(\leq y) = \sum_{k=1}^{\infty} \mathbf{P}[K = k] \cdot Q(\leq y)^k,$$

Therefore, differentiating, we see that the density function of A is given by

$$\begin{aligned} A(y) &= \sum_{k=1}^{\infty} k \cdot \mathbf{P}[K = k] \cdot Q(y) \cdot Q(\leq y)^{k-1} \\ &= Q(y) \sum_{k=1}^{\infty} k \cdot \mathbf{P}[K = k] \cdot Q(\leq y)^{k-1} \\ &= Q(y) \cdot \varphi'\big(Q(\leq y)\big). \end{aligned} \tag{B.5}$$

We have a similar representation for the density $A'(y)$, and the claim follows. □

C. Proofs for Section 5

C.1. Proof of Theorem 5.2

Theorem C.1. *Let $K \sim \mathcal{D}_{\eta,\gamma}$ and let $\delta(\epsilon_1)$, $\epsilon_1 \in \mathbb{R}$, define the privacy profile of the base mechanism Q. Then, for A and A', the output distributions of the selection algorithm evaluated on neighboring datasets X and X', respectively, and for all $\epsilon_1 \geq 0$,*

$$H_{\mathrm{e}^\epsilon}(A||A') \leq m \cdot \delta(\widehat{\epsilon})$$

where

$$\widehat{\epsilon} = \epsilon - (\eta+1)\log\left(\mathrm{e}^{\epsilon_1} + \frac{1-\gamma}{\gamma} \cdot \delta(\epsilon_1)\right).$$

Proof. Let q and q' be results of applying some post-processing function to Q and Q', respectively. Then,

$$1 - q' \leq \mathrm{e}^{\epsilon_1}(1-q) + \delta(\epsilon_1) \tag{C.1}$$

for all $\epsilon_1 \in \mathbb{R}$. Thus, we see that for all $\epsilon_1 \geq 0$,

$$\begin{aligned}
\frac{\varphi'(q)}{\varphi'(q')} &= \left(\frac{(\gamma-1)q'+1}{(\gamma-1)q+1}\right)^{\eta+1} \\
&= \left(\frac{(1-\gamma)(1-q')+\gamma}{(1-\gamma)(1-q)+\gamma}\right)^{\eta+1} \\
&\leq \left(\frac{(1-\gamma)(1-q)\mathrm{e}^{\epsilon_1} + \delta(\epsilon_1)(1-\gamma)+\gamma}{(1-\gamma)(1-q)+\gamma}\right)^{\eta+1} \\
&\leq \left(\frac{(1-\gamma)(1-q)\,\mathrm{e}^{\epsilon_1} + \delta(\epsilon_1)(1-\gamma)+\gamma\,\mathrm{e}^{\epsilon_1}}{(1-\gamma)(1-q)+\gamma}\right)^{\eta+1} \\
&= \left(\mathrm{e}^{\epsilon_1} + \frac{\delta(\epsilon_1)(1-\gamma)}{(1-\gamma)(1-q)+\gamma}\right)^{\eta+1} \\
&\leq \left(\mathrm{e}^{\epsilon_1} + \frac{1-\gamma}{\gamma}\delta(\epsilon_1)\right)^{\eta+1},
\end{aligned} \tag{C.2}$$

where we have used the inequality (C.1) in the first inequality and the inequality $\gamma \leq \mathrm{e}^{\epsilon_1}\gamma$ in the second inequality. Using Thm. 4.1 for the hockey-stick divergence $f(z) = [z - \mathrm{e}^\epsilon]_+$, we have that for all $\epsilon_1 \geq 0$,

$$\begin{aligned}
H_f(A||A') &\leq \sum_{y\in\mathcal{Y}} \left[\frac{Q(y)\varphi'(q_y)}{Q'(y)\varphi'(q'_y)} - \mathrm{e}^\epsilon\right]_+ \cdot Q'(y) \cdot \varphi'(q'_y) \\
&= \sum_{y\in\mathcal{Y}} \left[1 - \mathrm{e}^{\epsilon - \log\frac{Q(y)}{Q'(y)} - \log\frac{\varphi'(q_y)}{\varphi'(q'_y)}}\right]_+ \cdot Q(y) \cdot \frac{\mathbb{E}[K]\cdot\gamma^{\eta+1}}{\big(1-q_y(1-\gamma)\big)^{\eta+1}} \\
&\leq \sum_{y\in\mathcal{Y}} \left[1 - \mathrm{e}^{\epsilon - \log\frac{Q(y)}{Q'(y)} - \log\frac{\varphi'(q_y)}{\varphi'(q'_y)}}\right]_+ \cdot Q(y) \cdot \mathbb{E}[K] \\
&\leq \sum_{y\in\mathcal{Y}} \left[1 - \mathrm{e}^{\epsilon - \log\frac{Q(y)}{Q'(y)} - (\eta+1)\log\left(\mathrm{e}^{\epsilon_1} + \frac{1-\gamma}{\gamma}\delta(\epsilon_1)\right)}\right]_+ \cdot Q(y) \cdot \mathbb{E}[K] \\
&= \mathbb{E}[K] \cdot H_{\mathrm{e}^{\widetilde{\epsilon}}}\big(Q||Q'\big),
\end{aligned} \tag{C.3}$$

where $\widetilde{\epsilon} = \epsilon - (\eta+1)\log\left(\mathrm{e}^{\epsilon_1} + \frac{1-\gamma}{\gamma}\delta(\epsilon_1)\right)$. In the third inequality we have used the inequality (C.2) the fact that $[1 - \mathrm{e}^{\epsilon - s}]_+$ is a non-decreasing function of s for all $\epsilon \in \mathbb{R}$. □

C.2. Proof of Corollaries 5.3 and 5.4

Corollary C.2. *Let $K \sim \mathcal{D}_{\eta,\gamma}$. If the base mechanism Q is ϵ-DP, then the selection algorithm A is $(\eta+2)\epsilon$-DP. For $\eta = 1$ we get Theorem 1.3 of (Liu & Talwar, 2019).*

Proof. Let ϵ_1 be such that $\delta(\epsilon_1) = 0$, where $\big(\epsilon_1, \delta(\epsilon_1)\big)$ gives a privacy profile for the base mechanism Q. Then,

$$\begin{aligned}\widehat{\epsilon} &= \epsilon - (\eta+1)\log\left(\mathrm{e}^{\epsilon_1} + \frac{1-\gamma}{\gamma}\delta(\epsilon_1)\right) \\ &= \epsilon - (\eta+1)\epsilon_1,\end{aligned}$$

and by Theorem 5.2, $H_{\mathrm{e}^\epsilon}(A||A') = 0$ if $\epsilon = (\eta+2)\epsilon_1$. □

Corollary C.3. *Let $K \sim \mathcal{D}_{\eta,\gamma}$. If the base mechanism Q is (ϵ,δ)-DP, then then the selection algorithm A is $\big((\eta+2)\epsilon + \gamma^{-1}\delta, m\delta\big)$-DP.*

Proof. Let $\big(\epsilon_1, \delta(\epsilon_1)\big)$ be a privacy profile for the base mechanism Q and $\epsilon_1 \geq 0$. Then,

$$\begin{aligned}\widehat{\epsilon} &= \epsilon - (\eta+1)\log\left(\mathrm{e}^{\epsilon_1} + \frac{1-\gamma}{\gamma}\delta(\epsilon_1)\right) \\ &\geq \epsilon - (\eta+1)\log\left(\mathrm{e}^{\epsilon_1}\left(1+\gamma^{-1}\delta(\epsilon_1)\right)\right) \\ &\geq \epsilon - (\eta+1)\left(\epsilon_1 + \gamma^{-1}\delta(\epsilon_1)\right)\end{aligned}$$

and by Thm. 5.2, $H_{\mathrm{e}^\epsilon}(A||A') = m\delta$ if $\epsilon = (\eta+2)\epsilon_1 + \gamma^{-1}\delta$. □

C.3. Proof of Corollary 5.5

Corollary C.4 (ϵ-values when Q is GDP)**.** *Let $K \sim \mathcal{D}_{\eta,\gamma}$ with $\eta \geq 1$ and suppose the base mechanism is dominated by the Gaussian mechanism with noise parameter $\sigma > 0$ and L_2-sensitivity 1. Then, for a fixed $\delta > 0$, the private selection algorithm A is (ϵ,δ)-DP for*

$$\epsilon = (\eta+1)\left(\frac{1}{2\sigma^2} + \frac{1}{\sigma}\sqrt{2\log\frac{m}{\delta}}\right) + 2\delta.$$

Proof. By (Lemma 3, Balle & Wang, 2018), we know that for the Gaussian mechanism with sensitivity 1 and noise scale σ the privacy loss random variable ω is distributed as $\omega \sim \mathcal{N}\left(\frac{1}{2\sigma^2}, \frac{1}{\sigma^2}\right)$. Thus, using a simple Chernoff bound for the Gaussian, for its privacy profile $\delta(\epsilon_1)$ we have

$$\delta(\epsilon_1) \leq \mathbb{P}(\omega \geq \epsilon_1) \leq \mathrm{e}^{-\frac{\widetilde{\epsilon}^2\sigma^2}{2}},$$

where $\widetilde{\epsilon} = \epsilon_1 - \frac{1}{2\sigma^2}$. Choosing

$$\epsilon_1 = \frac{1}{2\sigma^2} + \frac{1}{\sigma}\sqrt{2\log\frac{m}{\delta}}, \tag{C.4}$$

we have that

$$\delta(\epsilon_1) \leq \frac{\delta}{m}. \tag{C.5}$$

Furthermore, for any $\eta \geq 1$, for ϵ_1 of Eq. C.4, we have the following bound for the additional term in the bound of Thm. 5.2:

$$\begin{aligned}(\eta+1)\log\left(\mathrm{e}^{\epsilon_1} + \tfrac{1-\gamma}{\gamma}\delta(\epsilon_1)\right) &\leq (\eta+1)\log\left(\mathrm{e}^{\epsilon_1} + \mathrm{e}^{\epsilon_1}\tfrac{1-\gamma}{\gamma}\delta(\epsilon_1)\right) \\ &= (\eta+1)\epsilon_1 + (\eta+1)\log\left(1+\tfrac{1-\gamma}{\gamma}\delta(\epsilon_1)\right) \\ &\leq (\eta+1)\epsilon_1 + (\eta+1)\log\left(1+\tfrac{1-\gamma}{\gamma}\tfrac{\delta}{m}\right) \\ &= (\eta+1)\epsilon_1 + (\eta+1)\log\left(1+\tfrac{1-\gamma}{\gamma}\tfrac{\gamma(1-\gamma^\eta)}{\eta(1-\gamma)}\delta\right) \\ &= (\eta+1)\epsilon_1 + (\eta+1)\log\left(1+\tfrac{1-\gamma^\eta}{\eta}\delta\right) \\ &\leq (\eta+1)\epsilon_1 + (\eta+1)\log\left(1+\tfrac{1}{\eta}\delta\right) \\ &\leq (\eta+1)\epsilon_1 + \frac{\eta+1}{\eta}\delta \\ &\leq (\eta+1)\epsilon_1 + 2\delta.\end{aligned} \tag{C.6}$$

Setting $\epsilon = (\eta+2)\epsilon_1 + 2\delta$, Thm. 5.2 and the inequalities (C.6) and (C.5) show that

$$\begin{aligned} H_{\mathrm{e}^\epsilon}(A||A') \leq & m \cdot \delta\big(\epsilon - (\eta+1)\log\big(1 + \tfrac{1-\gamma}{\gamma}\delta(\epsilon_1)\big)\big) \\ \leq & m \cdot \delta(\epsilon_1) \\ \leq & m \cdot \frac{\delta}{m} = \delta \end{aligned}$$

and the claim follows. □

C.4. Proof of Theorem 5.7 (Private Selection, Binomial Distribution)

First, the following auxiliary lemma is needed.

Lemma C.5. *Let $a, b, c, d > 0$. Then, for $x \geq 0$, the function*

$$f(x) = \frac{ax+b}{cx+d}$$

is non-decreasing if and only if $\frac{a}{b} \geq \frac{c}{d}$.

Proof. The claim follows from the expression

$$f'(x) = \frac{ad-cb}{(cx+d)^2}.$$

□

Recall that for $K \sim \mathrm{Bin}(n,p)$, the probability generating function is given by

$$\varphi(z) = (1-p+pz)^n. \tag{C.7}$$

Theorem C.6. *Let $K \sim \mathrm{Bin}(n,p)$ for some $n \in \mathbb{N}$ and $0 < p < 1$, and let $\delta(\epsilon_1)$, $\epsilon_1 \in \mathbb{R}$, define the privacy profile of the base mechanism Q. Suppose*

$$\epsilon_1 \geq \log\left(1 + \tfrac{p}{1-p}\delta(\epsilon_1)\right).$$

Then, for A and A', the output distributions of the selection algorithm evaluated on neighboring datasets X and X', respectively, for all $\epsilon > 0$ and for all $\epsilon_1 \geq 0$,

$$H_{\mathrm{e}^\epsilon}(A||A') \leq m \cdot \delta(\widehat{\epsilon}), \tag{C.8}$$

where

$$\widehat{\epsilon} = \epsilon - (n-1)\log\big(1 + p(\mathrm{e}^{\epsilon_1}-1) + p\delta(\epsilon_1)\big).$$

Proof. Using the PGF of the binomial distribution given in Eq. (C.7) by the auxiliary Lemma C.5, for each $y \in \mathcal{Y}$,

$$\begin{aligned} \frac{\varphi'(q_y)}{\varphi'(q_y')} &= \left(\frac{1-p+pq_y}{1-p+pq_y'}\right)^{n-1} \\ &\leq \left(\frac{1-p+p\mathrm{e}^{\epsilon_1}q_y' + p\delta(\epsilon_1)}{1-p+pq_y'}\right)^{n-1} \\ &= \left(1 + \frac{p(\mathrm{e}^{\epsilon_1}-1)q_y' + p\delta(\epsilon_1)}{1-p+pq_y'}\right)^{n-1} \\ &\leq \big(1 + p(\mathrm{e}^{\epsilon_1}-1) + p\delta(\epsilon_1)\big)^{n-1}, \end{aligned} \tag{C.9}$$

in case

$$\frac{p(\mathrm{e}^{\epsilon_1}-1)}{p\delta(\epsilon_1)} \geq \frac{p}{1-p},$$

i.e., if

$$\epsilon_1 \geq \log\left(1 + \frac{p}{1-p}\delta(\epsilon_1))\right).$$

Moreover, we see that

$$\varphi'(q) \leq n \cdot p = m \tag{C.10}$$

for all $0 \leq q \leq 1$. The claim follows from the inequalities (C.9) and (C.10) and from Theorem 4.1. □

C.5. Proof of Corollary 5.8

Corollary C.7. *Let $K \sim \mathrm{Poisson}(m)$ for some $m \in \mathbb{N}$, and let $\delta(\epsilon_1)$, $\epsilon_1 \in \mathbb{R}$, define the privacy profile of the base mechanism Q. Then, for A and A', the output distributions of the selection algorithm evaluated on neighboring datasets X and X', respectively, and for all $\epsilon > 0$, and for all $\epsilon_1 \geq 0$,*

$$H_{\mathrm{e}^\epsilon}(A||A') \leq m \cdot \delta(\widehat{\epsilon}),$$

where

$$\widehat{\epsilon} = \epsilon - m \cdot (\mathrm{e}^{\epsilon_1} - 1) - m \cdot \delta(\epsilon_1).$$

Proof. Let $K \sim \mathrm{Poisson}(m)$ and $K_n \sim \mathrm{Bin}(n, m/n)$ and let A, A' denote the density functions of the private selection algorithm corresponding to K and let A_n, A'_n those corresponding to K_n, evaluated on neighboring datasets X, X', respectively. Looking at the form of A given in Eq. (B.1), we have that

$$\begin{aligned}
H_{\mathrm{e}^\epsilon}\big(A||A'\big) &= \sum_{y \in \mathcal{Y}} \max\{A(y) - \mathrm{e}^\epsilon A'(y), 0\} \\
&= \sum_{k=1}^{\infty} \max\{\mathbf{P}[K=k] \cdot \big(Q(\leq y)^k - Q(< y)^k\big) - \mathrm{e}^\epsilon \mathbf{P}[K=k] \cdot \big(Q'(\leq y)^k - Q'(< y)^k\big), 0\} \\
&\leq \sum_{k=1}^{\infty} \max\{\mathbf{P}[K_n=k] \cdot \big(Q(\leq y)^k - Q(< y)^k\big) - \mathrm{e}^\epsilon \mathbf{P}[K_n=k] \cdot \big(Q'(\leq y)^k - Q'(< y)^k\big), 0\} \\
&\quad + (1+\mathrm{e}^\epsilon) \sum_{k=0}^{\infty} |\mathbf{P}[K=k] - \mathbf{P}[K_n=k]| \\
&= H_{\mathrm{e}^\epsilon}\big(A_n||A'_n\big) + (1+\mathrm{e}^\epsilon) \sum_{k=0}^{\infty} |\mathbf{P}[K=k] - \mathbf{P}[K_n=k]|,
\end{aligned}$$

where the inequality follows from the fact that $\max\{a+b, 0\} \leq |a| + \max\{b, 0\}$ for all $a, b \in \mathbb{R}$. Since by Le Cam's inequality, $\mathrm{Bin}(n, m/n) \to \mathrm{Poisson}(m)$ in total variation distance as $n \to \infty$, we have that

$$H_{\mathrm{e}^\epsilon}\big(A||A'\big) = \lim_{n \to \infty} H_{\mathrm{e}^\epsilon}\big(A_n||A'_n\big).$$

Fixing $n \cdot p = m$ in the bound (5.2) of Thm. 5.7 (bound for the case $K \sim \mathrm{Bin}(n, n/m)$), we see that the bound approaches the bound (5.3) of Cor. 5.8 (bound for the case $K \sim \mathrm{Poisson}(m)$) as $p \to 0$, since then

$$(n-1)\log\big(1 + p(\mathrm{e}^{\epsilon_1} - 1) + p\delta(\epsilon_1)\big) \to m \cdot (\mathrm{e}^{\epsilon_1} - 1) + m \cdot \delta(\epsilon_1).$$

This follows from the fact that $\frac{\log(1+x)}{x} \to 1$ as $x \to 0$. □

Remark C.8. *We can also get Cor. 5.8 directly using the PGF of the Poisson distribution. For $K \sim \mathrm{Poisson}(m)$, the PGF* *is $\varphi(z) = \mathrm{e}^{m(z-1)}$, i.e. $\varphi'(z) = m \cdot \mathrm{e}^{m(z-1)}$. Using Thm. 4.1 for the hockey-stick divergence $f(z) = [a - \mathrm{e}^{\epsilon}]_+$, we get*

$$\begin{aligned} H_f(A||A') &\leq \sum_{y \in \mathcal{Y}} f\left(\frac{Q(y)\varphi'(q_y)}{Q'(y)\varphi'(q'_y)}\right) \cdot Q'(y) \cdot \varphi'(q'_y) \\ &= \sum_{y \in \mathcal{Y}} \left[\frac{Q(y)\varphi'(q_y)}{Q'(y)\varphi'(q'_y)} - \mathrm{e}^{\epsilon}\right]_+ \cdot Q'(y) \cdot \varphi'(q'_y) \\ &= \sum_{y \in \mathcal{Y}} \left[1 - \mathrm{e}^{\epsilon - \log \frac{Q(y)}{Q'(y)} - \log \frac{\varphi'(q_y)}{\varphi'(q'_y)}}\right]_+ \cdot Q(y) \cdot \varphi'(q_y) \\ &= \sum_{y \in \mathcal{Y}} \left[1 - \mathrm{e}^{\epsilon - \log \frac{Q(y)}{Q'(y)} - m \cdot (q_y - q'_y)}\right]_+ \cdot Q(y) \cdot m \cdot \mathrm{e}^{m \cdot (q_y - 1)}. \end{aligned} \tag{C.11}$$

As the probabilities q_y and q'_y are obtained by applying the same post-processing to Q and Q', for all $\epsilon_1 \geq 0$, $q'_y \leq \mathrm{e}^{\epsilon_1} q_y + \delta(\epsilon_1)$. Using also the fact that $[1 - \mathrm{e}^{\epsilon - s}]_+$ is a non-decreasing function of s for all $\epsilon \in \mathbb{R}$, we get from the inequality (C.11) that for all $\epsilon_1 \geq 0$,

$$\begin{aligned} H_f(A||A') &\leq \sum_{y \in \mathcal{Y}} \left[1 - \mathrm{e}^{\epsilon - \log \frac{Q(y)}{Q'(y)} - m \cdot (q_y - q'_y)}\right]_+ \cdot Q(y) \cdot m \cdot \mathrm{e}^{m \cdot (q_y - 1)} \\ &\leq \sum_{y \in \mathcal{Y}} \left[1 - \mathrm{e}^{\epsilon - \log \frac{Q(y)}{Q'(y)} - m \cdot q'_y(1 - \mathrm{e}^{\epsilon_1}) - m \cdot \delta(\epsilon_1)}\right]_+ \cdot Q(y) \cdot m \cdot \mathrm{e}^{m \cdot (q_y - 1)}. \\ &\leq \sum_{y \in \mathcal{Y}} \left[1 - \mathrm{e}^{\epsilon - \log \frac{Q(y)}{Q'(y)} - m \cdot (1 - \mathrm{e}^{\epsilon_1}) - m \cdot \delta(\epsilon_1)}\right]_+ \cdot Q(y) \cdot m \end{aligned} \tag{C.12}$$

which gives the claim of Cor. 5.8.

D. Converting RDP Bounds to (ϵ, δ)-Bounds

To convert from Rényi DP to approximate DP we use following formula.

Lemma D.1 (Canonne et al. 2020)**.** *Suppose the mechanism $\mathcal{M}$ is (α, ϵ')-RDP. Then $\mathcal{M}$ is also $(\epsilon, \delta(\epsilon))$-DP for arbitrary $\epsilon \geq 0$ with*

$$\delta(\epsilon) = \frac{\exp\big((\alpha - 1)(\epsilon' - \epsilon)\big)}{\alpha} \left(1 - \frac{1}{\alpha}\right)^{\alpha - 1}. \tag{D.1}$$