PRIVATE TOP-k SELECTION UNDER GUMBEL DIFFERENTIAL PRIVACY GUARANTEES

Anonymous authors
Paper under double-blind review

ABSTRACT

From the perspective of hypothesis testing, f-differential privacy (f-DP) as a relaxation of differential privacy (DP) possesses numerous desirable properties, the most prominent of which is its lossless characterization of the composition of DP mechanisms. Within the f-DP class, Gaussian differential privacy (GDP), as a canonical family introduced to design Gaussian mechanism, has gained widespread acceptance. However, Gaussian mechanism is not the optimal option for all scenarios to ensure DP. As a type of extreme value distribution, Gumbel distribution is naturally considered to design private top-k selection algorithms. In this work, a new family in f-DPs, named Gumbel differential privacy (GumDP), is developed to parameterize Gumbel mechanism as similar to GDP. And the composition of Gumbel mechanisms is studied. In addition, two important composition properties of the Gumbel mechanism are discovered among different private selection problems. Utilizing these, a novel privacy-preserving top-k selection algorithm with Gumbel mechanism, called the peeling algorithm under oneshot RNM, is presented based on the Report Noisy Min (RNM) and peeling algorithms. Simulations demonstrate that the privacy-utility performance of the proposed private selection algorithm is significantly improved compared to the peeling algorithm under RNM with Laplace or Gaussian mechanism.

1 INTRODUCTION

With the rapid advancement of the information era, vast amounts of data are generated and released daily. This has led to heightened awareness of personal privacy and increased focus on privacy protection technologies. Based on these, differential privacy (DP) (Dwork et al., 2006a;b), as an emerging technology for protecting individual user privacy, has received widespread attention from both academia and industry. On the one hand, the definition is used by academics for a wide range of research, e.g., the privatization of deep learning (Abadi et al., 2016; Zhao et al., 2020) and federated learning (Wei et al., 2020; Yazdinejad et al., 2024; Cai et al., 2024), and the protection of models and data in statistics (Alparslan & Yıldırım, 2022; Awan & Wang, 2024; Lin et al., 2024; Acharya et al., 2024). On the other hand, in industry, DP is also the core technology used by Apple (Differential Privacy Team, 2017), Google (Erlingsson et al., 2014), Microsoft (Ding et al., 2017), and the US Census Bureau (Abowd, 2018; Groshen & Goroff, 2022).

Under the theoretical framework of DP, designing the privacy-preserving mechanism to perturb the output by adding noise is the core concept of the DP application where the three major ones are Laplace, Gaussian and exponential mechanisms (Dwork et al., 2006a;b; McSherry & Talwar, 2007). With the goal of privacy and utility maximization, a large body of literature examines and parameterizes these mechanisms. However, as stated in (Brenner & Nissim, 2010), there is no universally optimal DP mechanism for all types of queries. Hence, the design of DP mechanism is one of the hot issues in DP research trying to start from the perspective of different noise distributions (Liu, 2019; Sadeghi & Korki, 2022; Muthukrishnan & Kalyani, 2023). In addition, along with the complexity and modularity of the algorithm in large models, there will be multiple queries to the database implying the composition of DP mechanisms. The composition of these mechanisms will degrade the privacy-utility performance. The naive and advanced composition theorems (Dwork et al., 2006a; 2010) are originally formulated to track these privacy performances which only carve out loose privacy upper bounds. To achieve a tighter privacy upper bound, some important variants of DP that have also been proposed to minimize the privacy loss of the composition process are zero-Concentrated Differential Privacy, Rényi Differential Privacy, truncated Concentrated Differential Privacy (Bun & Steinke, 2016; Mironov, 2017; Bun et al., 2018). According to the above approaches, the composition of Laplace, and Gaussian mechanisms all lead to tighter bounds. Unfortunately, these results are still the relaxing ones of the privacy upper bound. Meanwhile, the statistical perspective of transforming DP into a hypothesis testing problem that cannot distinguish the output of two neighboring datasets under the same mechanism has been proposed to enrich the research perspective (Wasserman & Zhou, 2010). The f-differential privacy (f-DP) is put forward in line with this idea (Dong et al., 2022) where Gaussian differential privacy (GDP) as a family of f-DPs gives a loss-free privacy upper bound carving on the composition of Gaussian mechanisms. And GDP is heavily used because of its lossless after composition and the universality of the Gaussian mechanism (Zheng et al., 2021; Bu et al., 2023; Liu et al., 2024). However, Brenner & Nissim (2010) shows that the Gaussian mechanism is not the optimal one in all application scenarios.

Naturally, the Gumbel distribution, as the most common extreme value distribution, has gained interest in the design of DP mechanisms to protect privacy in selection problems. Among selection problems, the top-k query techniques (Ilyas et al., 2008) is one of the top-mentioned techniques, which are widely used in the Web, medical, government, such as information crawling for search engines, sorting queries for medical data, analyzing and researching demographic data. Obviously, attackers can strike this query process to steal the privacy of individual users. Therefore, this work attempts using Gumbel distribution to privatize the top-k selection algorithm and providing the privacy bound for the k-fold Gumbel mechanism embedded implicit in the algorithm. Unfortunately, under the definition of f-DP, the lack of research on other types of trade-off functions allows for a tighter privacy characterization for DP mechanisms beyond the Gaussian mechanism. Building upon that, a new family of trade-off functions for Gumbel mechanism is proposed.

1.1 Related Works

The private top-k selection algorithm under DP has been extensively studied in fields such as statistics (Dwork et al., 2021; Cai et al., 2021; Xia & Cai, 2023) and machine learning (Cohen & Lyu, 2023; Lebeda & Tetek, 2025; Pagh et al., 2025). However, how to select a suitable DP mechanism and depict the composition of those mechanisms are still two central issues in designing a top-k selection algorithm under DP. For simplicity, private top-1 query, also called private selection algorithm, is prior subjected to research that returns the minimum perturbed query value $\tilde{q}_i$ and its corresponding index i given n queries $\{q_1, \ldots, q_n\}$. Exponential mechanism (EM) (McSherry & Talwar, 2007; McKenna & Sheldon, 2020) and Report Noisy Min (RNM) algorithm (Dwork & Roth, 2013; Durfee & Rogers, 2019; Zhu & Wang, 2022) are two common selection algorithms under DP. In the development of RNM algorithm, it is essentially a matter of perturbing the output index by attempting to apply the Laplace, Gaussian or Gumbel mechanisms to the query value and re-perturbing the corresponding query value by the Laplace or Gaussian mechanism. Furthermore, extending top-1 to top-k, the goal of private top-k selection is to design a DP algorithm that outputs the k smallest perturbed query values $\{\tilde{q}_{i_1}, \ldots, \tilde{q}_{i_k}\}$ and their corresponding indexes $\{i_1, \ldots, i_k\}$. There are two main methods to perform k-term items selection, namely peeling algorithm (Hardt & Roth, 2013; Dwork et al., 2021; Xia & Cai, 2023) and oneshot algorithm (Durfee & Rogers, 2019; Qiao et al., 2021). Considering the complexity of analyzing privacy parameters in the oneshot algorithm, this paper considers only the peeling algorithm.

1.2 Contributions

In this work, we try to design a new private top-k selection algorithm with Gumbel mechanism and utilize f-DP to ensure better privacy-utility performance for this algorithm. The main contributions of this work are summarized as follows:

- The Gumbel mechanism is firstly proposed to directly noise the query value to ensure DP. Gumbel differential privacy (GumDP), as a special family of trade-off function in f-DPs, is designed to precisely characterize the Gumbel mechanism and its composition under the assumption that the query functions are consistent. In addition, two equivalent conversion forms between GumDP and DP are given.
- Two attractive composition properties of the Gumbel distribution in the private selection problem are presented, as seen in Lemma 1 and Lemma 2. Based on these and the RNM

algorithm, a newly validated private selection algorithm with Gumbel mechanism, named oneshot RNM algorithm, is introduced which can simultaneously output the index and query value without re-adding noise. Building upon Gumbel mechanism, Theorem 3 guarantees the privacy of this algorithm.

- Extending to top-k private selection, the peeling algorithm under oneshot RNM with Gumbel mechanism, whose privacy is secured by Theorem 4, is put through the peeling algorithm. And simulations show that there is a significant reduction in the variance of the added noise compared to the peeling algorithm under RNM with Laplace or Gaussian mechanism, which also confirms the increase in data availability and ensures that Gumbel mechanism achieves superior privacy-utility performance in the top-k selection problem.

Notations: Let $\mathcal{N}(0,\sigma^2)$, $\mathrm{Lap}(\lambda)$ and $\chi^2(2k)$ represent Gaussian distribution with location parameter 0 and scale parameter σ, Laplace distribution with mean 0 and scale parameter λ, and chi-square distribution with parameter $2k$ respectively. The $\mathrm{sign}(x)$ denotes the signature function, i.e., $\mathrm{sign}(x) = \begin{cases} -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0, \\ 1, & \text{if } x > 1. \end{cases}$ And $[m]$ and $\mathcal{R}$ represent the set of $\{1, \cdot, m\}$ and the set of real numbers respectively.

Mathematical details: Due to the space limitation, all details of the proofs of lemmas, corollaries and theorems in this paper are provided in the appendices.

2 Gumbel Differential Privacy

For the sake of subsequent discussion, we foresee the definitions of DP and f-DP, along with their equivalence transformation relationship. Let $\mathcal{D} = (d_1, d_2, \ldots, d_l)$, $\mathcal{D}' = (d_1', d_2', \ldots, d_l')$, denoted two neighboring datasets containing l data items, of which l can be interpreted as the number of users in the database, be sampling from $\mathcal{X}^l$ where $\mathcal{X}$ is a sample universe. These two datasets differ in one and only one data item, i.e., only one $j \in [l]$ such that $d_j \neq d_j'$. Dwork et al. (2006a;b) propose DP to protect the individual privacy which is unable to distinguish between $\mathcal{D}$ and $\mathcal{D}'$.

Definition 1 ((ε,δ)-DP (Dwork & Roth, 2013))**.** *For any $\varepsilon > 0$ and $\delta > 0$, a mechanism $\mathcal{M}$ is (ε,δ)-DP if for all adjacent databases $\mathcal{D}, \mathcal{D}'$ and any measurable event $S \subset \mathcal{R}$,*

$$P(\mathcal{M}(\mathcal{D}) \in S) \leq \mathrm{e}^{\varepsilon} P(\mathcal{M}(\mathcal{D}') \in S) + \delta.$$

From the definition of DP, it is evident that the smaller the privacy parameters ε and δ, the higher the level of privacy protection provided by the corresponding DP mechanism $\mathcal{M}$. In f-DP, it is natural to extend it to the problem of hypothesis testing where the distribution of the null hypothesis follows $\mathcal{M}(\mathcal{D})$ and the alternative one follows $\mathcal{M}(\mathcal{D}')$ making it is difficult to distinguish them. Let ϕ denote the rejection rule. The trade-off function as a tool to characterize the degree of difference between two hypotheses is

$$T(\mathcal{M}(\mathcal{D}), \mathcal{M}(\mathcal{D}'))(\alpha) = \inf_{\phi} \{\beta_\phi : \alpha_\phi \leq \alpha\},$$

where α_ϕ and β_ϕ are its corresponding type I and II errors respectively defined as

$$\alpha_\phi = \mathbb{E}_{\mathcal{M}(\mathcal{D})}[\phi], \quad \beta_\phi = 1 - \mathbb{E}_{\mathcal{M}(\mathcal{D}')}[\phi].$$

Definition 2 (f-DP(Dong et al., 2022))**.** *Given a trade-off function $f : [0,1] \to [0,1]$ satisfies convexity, continuity, and $f(x) \leq 1 - x$ for $x \in [0,1]$. A mechanism $\mathcal{M}$ is said to be f-DP if*

$$T(\mathcal{M}(\mathcal{D}), \mathcal{M}(\mathcal{D}')) \geq f,$$

for all neighbouring datasets $\mathcal{D}$ and $\mathcal{D}'$.

The closer f in Definition 2 approaches $g(x) = 1 - x$ with $x \in [0,1]$, the higher the level of privacy protection provided by the DP mechanism $\mathcal{M}$. Besides, the equivalent conversion between f-DP and DP is also given by Dong et al. (2022) through the concept of convex conjugate. For a function f with $f(x) = \infty$ for $x < 0$ or $x > 1$, its convex conjugate is defined as $f^*(y) = \sup_{-\infty<x<\infty}(yx - f(x))$. For a symmetric trade-off function f, a mechanism is f-DP if and only if it is $(\varepsilon, \delta(\varepsilon))$-DP for all $\varepsilon > 0$ with

$$\delta(\varepsilon) = 1 + f^*\left(-e^{\varepsilon}\right). \tag{1}$$

2.1 μ-GUMBEL DIFFERENTIAL PRIVACY

For a random variable X distributed from the Gumbel (minimum) distribution with location parameter μ and scale parameter $\gamma > 0$, denoted as $X \sim \mathrm{Gum}(\mu, \gamma)$, its variance is $\pi^2\gamma^2/6$, and its cumulative distribution function (CDF) and probability density function (PDF) respectively are

$$F(x; \mu, \gamma) = 1 - e^{-e^{\frac{x-\mu}{\gamma}}}, \; p(x; \mu, \gamma) = \frac{1}{\gamma} e^{\frac{x-\mu}{\gamma} - e^{\frac{x-\mu}{\gamma}}}.$$

Definition 3 (Gumbel Mechanism)**.** *Given a database $\mathcal{D}$ and a query function h, the Gumbel mechanism $\mathcal{M}_{\mathrm{Gum}}$ is defined as*

$$\mathcal{M}_{\mathrm{Gum}}(\mathcal{D}) = h(\mathcal{D}) + \eta, \qquad \eta \sim \mathrm{Gum}(0, \gamma).$$

Analogously to GDP (Dong et al., 2022), from the Gumbel distribution aspect, we design the μ-GumDP as a special family of the trade-off function in f-DP. Consider the following hypothesis testing problem:

$$H_0 : y \sim \mathcal{M}_{\mathrm{Gum}}(\mathcal{D}) \quad \text{versus} \quad H_1 : y \sim \mathcal{M}_{\mathrm{Gum}}(\mathcal{D}'). \tag{2}$$

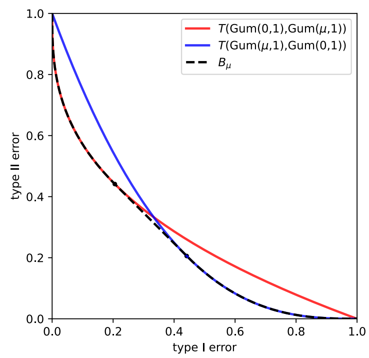


Figure 1: The trade-off functions $T(\mathrm{Gum}(0,1), \mathrm{Gum}(\mu,1)), T(\mathrm{Gum}(\mu,1), \mathrm{Gum}(0,1))$ and B_μ with $\mu = 1$.

Unlike the Gaussian distribution, the Gumbel distribution is asymmetric. From the hypothesis testing problem in (2), as shown in the Fig. 1, for any $\mu \geq 0$,

$$T(\mathrm{Gum}(0,1), \mathrm{Gum}(\mu,1)) \neq T(\mathrm{Gum}(\mu,1), \mathrm{Gum}(0,1)).$$

To facilitate subsequent conversion to DP, we perform a two-step operation in the definition of trade-off function B_μ: symmetrization and convexification. Symmetrization is taking the minimum of $T(\mathrm{Gum}(0,1), \mathrm{Gum}(\mu,1))$ and $T(\mathrm{Gum}(\mu,1), \mathrm{Gum}(0,1))$; and convexification is taking the bi-conjugate one, i.e., the largest convex lower envelope, of

$$\min\{T(\mathrm{Gum}(0,1), \mathrm{Gum}(\mu,1)), T(\mathrm{Gum}(\mu,1), \mathrm{Gum}(0,1))\}.$$

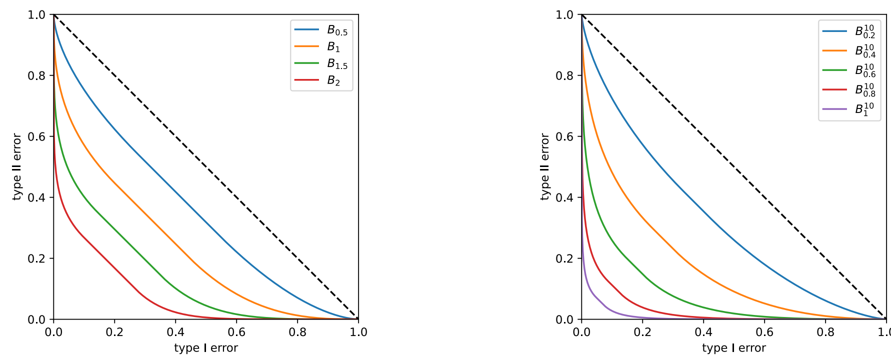
Definition 4. *For $\mu \geq 0$, the trade-off function B_μ is defined as*

$$B_\mu = \min\{T(\mathrm{Gum}(0,1), \mathrm{Gum}(\mu,1)), T(\mathrm{Gum}(\mu,1), \mathrm{Gum}(0,1))\}^{**}. \tag{3}$$

The B_μ for any $\mu \geq 0$ satisfies the requirements for the trade-off function as defined in Definition 2. Fig. 1 also presents the curve of B_μ with $\mu = 1$. And the explicit expression for the trade-off function B_μ in (3) reads

$$B_\mu(\alpha) = \begin{cases} 1 - \alpha^{e^{-\mu}}, & \alpha \in [0, \alpha_1), \\ -\alpha + e^{\frac{\mu}{e^{-\mu}-1}} + 1 - e^{\frac{\mu e^{-\mu}}{e^{-\mu}-1}}, & \alpha \in [\alpha_1, \alpha_2), \\ (1-\alpha)^{e^{\mu}}, & \alpha \in [\alpha_2, 1], \end{cases} \tag{4}$$

where $\alpha_1 = e^{\frac{\mu}{e^{-\mu}-1}}$ and $\alpha_2 = 1 - e^{\frac{\mu e^{-\mu}}{e^{-\mu}-1}}$. The proof details of (4) are provided in Appendix A. From (4), this trade-off function is decreasing in μ that $B_\mu > B_{\mu_0}$ if $\mu < \mu_0$, as shown in Fig. 2(a).



(a) (b)

Figure 2: (a) Changes in B_μ curves for different values of μ. (b) Changes in B_μ^{10} curves for different values of μ.

Definition 5 (Gumbel Differential Privacy)**.** *A mechanism $\mathcal{M}$ is said to satisfy μ-Gumbel differential privacy, denoted as μ-GumDP, if*

$$T(\mathcal{M}(\mathcal{D}), \mathcal{M}(\mathcal{D}')) \geq B_\mu,$$

for all neighboring datasets $\mathcal{D}$ and $\mathcal{D}'$.

Theorem 1. *If a Gumbel mechanism operates on a real statistic h as $\mathcal{M}_{\mathrm{Gum}}(\mathcal{D}) = h(\mathcal{D}) + \eta$, where $\eta \sim \mathrm{Gum}(0, \gamma)$ and $\Delta h = \max_{\mathcal{D},\mathcal{D}'} |h(\mathcal{D}) - h(\mathcal{D}')|$, then $\mathcal{M}_{\mathrm{Gum}}$ is μ-GumDP with $\gamma\mu \geq \Delta h$.*

From Definition 5, μ-GumDP, on the one hand, facilitates the privacy analysis and comparison as a one-parameter privacy definition; on the other hand, achieves a good degree of privacy at $\mu < 0.5$ as shown in Fig. 2(a). Besides, it has a tight privacy carving for the Gumbel mechanism by Theorem 1. Next, we provide an equivalent transformation between μ-GumDP and (ε, δ)-DP to conveniently compare the privacy-utility performance of different DP mechanisms.

Corollary 1. *A mechanism is satisfied μ-GumDP if and only if it is $(\varepsilon, \delta(\varepsilon))$-DP for all $\varepsilon \geq 0$, where $\delta(\varepsilon) = (e^{\mu+\varepsilon} - e^{\varepsilon})e^{\frac{\mu+\varepsilon}{e^{-\mu}-1}}$.*

2.2 (k, μ)-Gumbel Differential Privacy

In practical applications, the composition of DP mechanisms is often involved. Therefore, this section proposes (k, μ)-GumDP to characterize the composition of Gumbel mechanisms.

Definition 6 (k-fold Composed Mechanism)**.** *When $k = 2$, with the first mechanism $\mathcal{M}_1 : \mathcal{X}^l \to \mathcal{R}$ and the second mechanism $\mathcal{M}_2 : \mathcal{X}^l \times \mathcal{R} \to \mathcal{R}$, the 2-fold mechanism $\mathcal{M} : \mathcal{X}^l \to \mathcal{R} \times \mathcal{R}$ is given by $\mathcal{M}(\mathcal{D}) = (y_1, \mathcal{M}_2(\mathcal{D}, y_1))$ with $\mathcal{M}_1(\mathcal{D}) = y_1$ and $\mathcal{D} \in \mathcal{X}^l$. Let $\mathcal{M}_i : \mathcal{X}^l \times \mathcal{R}^{i-1} \to \mathcal{R}$, $i \in [k]$. Extension to the case of $k \geq 2$, the k-fold composed mechanism $\mathcal{M}$ of $\mathcal{M}_i$, $i \in [k]$, is defined as*

$$\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_k) : \mathcal{X}^l \to \mathcal{R}^k.$$

Based on Definition 6, considering a new hypothesis testing problem for the k-fold composed mechanism $\mathcal{M}$:

$$H_0 : (y_1, y_2, \ldots, y_k) \sim \mathcal{M}(\mathcal{D}) \quad \text{versus} \quad H_1 : (y_1, y_2, \ldots, y_k) \sim \mathcal{M}(\mathcal{D}'). \tag{5}$$

Assuming that $\mathcal{M}_i$ is an independent Gumbel mechanism given $\{y_j\}_{j=1}^{i-1}$, the above hypothesis test can be viewed as a discussion of independent composition of k Gumbel mechanisms. For ease of analysis, the Gumbel mechanism corresponding to each $\mathcal{M}_i$ is based on the identical Gumbel distribution, denoted as $\mathcal{M}_{\mathrm{Gum}}$. Under the above assumptions, $y_1, y_2, \ldots, y_k$ are independently and identically distributed (i.i.d.) from $\mathcal{M}_{\mathrm{Gum}}(\mathcal{D})$ given database $\mathcal{D}$. Then, the hypothesis test problem (5) can be converted to

$$H_0 : \{y_j\}_{j=1}^k \overset{\text{i.i.d.}}{\sim} \mathcal{M}_{\mathrm{Gum}}(\mathcal{D}) \quad \text{versus} \quad H_1 : \{y_j\}_{j=1}^k \overset{\text{i.i.d.}}{\sim} \mathcal{M}_{\mathrm{Gum}}(\mathcal{D}'). \tag{6}$$

Similar to B_μ in (3), we propose the following trade-off function B_μ^k.

Definition 7. *For $\mu \geq 0$, the trade-off function is defined as*

$$B_\mu^k = \min\{T(\text{Gum}(0,1)^k, \text{Gum}(\mu,1)^k), T(\text{Gum}(\mu,1)^k, \text{Gum}(0,1)^k)\}^{**}.$$

And the explicit expression for the trade-off function B_μ^k reads

$$B_\mu^k = \begin{cases} F_Y(F_Y^{-1}(1-\alpha)e^{-\mu}), & \alpha \in [0, \alpha_1), \\ \alpha_1 - \alpha + F_Y(F_Y^{-1}(1-\alpha_1)e^{-\mu}), & \alpha \in [\alpha_1, \alpha_2), \\ 1 - F_Y(F_Y^{-1}(\alpha)e^{\mu}), & \alpha \in [\alpha_2, 1], \end{cases} \tag{7}$$

where $\alpha_1 = 1 - F_Y\left(\frac{2k\mu}{1-e^{-\mu}}\right), \alpha_2 = F_Y\left(\frac{2k\mu e^{-\mu}}{1-e^{-\mu}}\right)$ and $Y \sim \chi^2(2k)$ with CDF F_Y. The proof details of (7) are provided in Appendix D. The trade-off function B_μ^k is also decreasing in μ that $B_\mu^k > B_{\mu_0}^k$ if $\mu < \mu_0$, as seen in Fig. 2b.

Definition 8 ((k,μ)-Gumbel Differential Privacy)**.** *A mechanism $\mathcal{M}$ is said to satisfy (k,μ)-Gumbel differential privacy ((k,μ)-GumDP) if*

$$T(\mathcal{M}(\mathcal{D}), \mathcal{M}(\mathcal{D}')) \geq B_\mu^k$$

for all neighbouring data sets $\mathcal{D}$ and $\mathcal{D}'$.

Let the query functions $\{h_i\}_{i=1}^k$ be consistent before characterizing the k-fold composed mechanism under the Gumbel distribution.

Definition 9. *The k query functions $\{h_i\}_{i=1}^k$ are consistent if if either $\text{sign}(h_j(D') - h_j(D)) \leq 0$ for all $j = 1, \ldots, k$, or $\text{sign}(h_j(D') - h_j(D)) \geq 0$ for all $j = 1, \ldots, k$.*

Theorem 2. *Consider the Gumbel mechanism operating on a statistic h_i as $\mathcal{M}_i(\mathcal{D}) = h_i(\mathcal{D}, y_1, y_2, \ldots, y_{i-1}) + \eta_i$ where $i \in [k]$, $\eta_i \overset{i.i.d.}{\sim} \text{Gum}(0, \Delta/\mu)$, $\Delta = \max_{i\in[k]} \max_{\mathcal{D},\mathcal{D}'} |h_i(\mathcal{D}) - h_i(\mathcal{D}')|$. If $\{h_i\}_{i=1}^k$ are consistent, then the k-fold composed mechanism $\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_k)$ is (k,μ)-GumDP.*

By Theorem 2, it can be seen as that the composition of k Gumbel mechanisms satisfied μ-GumDP is (k,μ)-GumDP under certain conditions. Meanwhile, the equivalence transformation between (k,μ)-GumDP and (ε,δ)-DP is proposed as follows.

Corollary 2. *A mechanism is (k,μ)-GumDP if and only if its k-fold mechanism is $(\varepsilon,\, \delta_k(\varepsilon))$-DP for all $\varepsilon > 0$, where $\delta_k(\varepsilon) = 1 - e^\varepsilon + e^\varepsilon F_Z\left(\frac{2(k\mu+\varepsilon)}{1-e^{-\mu}}\right) - F_Z\left(\frac{2(k\mu+\varepsilon)e^{-\mu}}{1-e^{-\mu}}\right)$ and F_Z denotes the CDF of distribution $\chi^2(2k)$.*

Note that the $\delta_k()$ in Corollary 2 is strictly derived from the equivalent conversion between f-DP and DP, i.e. Equation (1), so that $\delta_k(\varepsilon) \in [0,1]$ for all $\varepsilon > 0$.

3 Private top-k selection under Gumbel mechanism

To ease the study, the top-k selection problem in this paper is to perform m real queries for any database $\mathcal{D}$, i.e., $\{h_1(\mathcal{D}), h_2(\mathcal{D}), \ldots, h_m(\mathcal{D})\}$, sort the m queries, i.e., $h_{i_1}(\mathcal{D}) < h_{i_2}(\mathcal{D}) < \cdots < h_{i_m}(\mathcal{D})$, and finally output the smallest k query values and the corresponding indexes, i.e., $\{(i_1, h_{i_1}(\mathcal{D})), (i_2, h_{i_2}(\mathcal{D})), \ldots, (i_k, h_{i_k}(\mathcal{D}))\}$. The output of indexes and query values suffers from the leakage of individual privacy in $\mathcal{D}$. The peeling algorithm under RNM as a top-k selection algorithm under DP protects both of them (Dwork et al., 2021). In this section, based on the above algorithm and the Gumbel mechanism under GumDP given in the previous section, we design a newly private top-k selection algorithm. Moreover, analyzing from the perspective of adding noise variance, the new algorithm guarantees higher privacy-utility performance.

3.1 The Peeling Algorithm under Oneshot Report Noise Min

Before designing the private selection algorithm, there are two important composition properties about Gumbel mechanism.

Lemma 1. *Let $\{\mathcal{M}_{\text{Gum}}^{(i)}(\mathcal{D}) = h_i(\mathcal{D}) + \eta_i\}_{i\in[m]}$ be μ-GumDP where $\{\eta_i\}_{i\in[m]} \overset{i.i.d.}{\sim} \text{Gum}(0, \frac{\Delta}{\mu})$. The minimum Gumbel mechanism $\mathcal{M}_{\text{minGum}}^{m}$ is defined as, for $m \in \mathbb{N}^+$ and any database $\mathcal{D}$,*

$$\mathcal{M}_{\text{minGum}}^{m}(\mathcal{D}) = \min_{i\in[m]} \{\mathcal{M}_{\text{Gum}}^{(i)}(\mathcal{D})\}.$$

The $\mathcal{M}_{\text{minGum}}^{m}$ can be also seen as a Gumbel mechanism which satisfies μ-GumDP.

Lemma 1 illustrates that the minimum output among the noisy query values perturbed by Gumbel noises still satisfies GumDP. However, the private selection problem considered in this paper requires not only the minimum query value but also its corresponding index. To find the best index $j \in [m]$ and output the corresponding query value $h_j(\mathcal{D})$, the RNM algorithm (Dwork & Roth, 2013) with Laplace mechanism satisfied $(\varepsilon, 0)$-DP: Add the independent Laplace noise ω from $\text{Lap}(2\Delta/\varepsilon)$ to each query $\{h_j(\mathcal{D})\}_{j=1}^{m}$, return the index j^* of the smallest noisy one $\tilde{h}_{j^*}(\mathcal{D}) = h_{j^*}(\mathcal{D}) + \omega$ and draw a fresh noise $\text{Lap}(2\Delta/\varepsilon)$ added to $h_{j^*}(\mathcal{D})$ to output the noise one. It is evident that the original RNM algorithm suffers from a privacy allocation issue concerning both the index and the query value. Meanwhile, EM in Dwork & Roth (2013) as another common privacy selection algorithm only outputs the best index $i \in [m]$. The EM $\mathcal{M}_{\text{E}}$ satisfies $(\varepsilon, 0)$-DP which outputs the index i with probability

$$P\big(\mathcal{M}_{\text{E}}(\mathcal{D}, \{h_j\}_{j\in[m]}, \varepsilon) = i\big) = \frac{e^{-\frac{\varepsilon h_i(\mathcal{D})}{\Delta}}}{\sum_{j\in[m]} e^{-\frac{\varepsilon h_j(\mathcal{D})}{\Delta}}}.$$

Fortunately, Durfee & Rogers (2019) demonstrates that $\mathcal{M}_{\text{E}}\big(\cdot, \{h_j\}_{j\in[m]}, \varepsilon\big)$ is equivalent to the RNM with $\text{Gum}(0, \Delta/\varepsilon)$ when both of them output only the index. Building upon Lemma 1 and the relation the EM and the Gumbel mechanism, Lemma 2 gives a natural way to assign privacy-preserving parameters to the output query value and its corresponding index in the private selection problem utilizing Gumbel mechanism .

Lemma 2. *For any database $\mathcal{D}$ and a batch of query values $\{h_j(\mathcal{D})\}_{j\in[m]}$ added independent noise perturbations from $\text{Gum}(0, \frac{\Delta}{\varepsilon})$, output the minimum noise query value and its index concurrently, denoted as $\mathcal{M}_{\text{Gum}}^{*}(\mathcal{D})$, is equal to the independent composition of the Gumbel mechanism $\mathcal{M}_{\text{minGum}}^{m}(\mathcal{D})$ satisfied ε-GumDP and the EM $\mathcal{M}_E(\mathcal{D}, \{h_j\}_{j\in[m]}, \varepsilon)$, i.e., for any $S \subset \mathbb{R}$,*

$$\begin{aligned} &P\left(M_{\text{Gum}}^{*}(D) = (i, h_i(D) + \eta_i) \in [m] \times S\right) \\ &= P\left(\mathcal{M}_{\text{minGum}}^{m}(\mathcal{D}) \in S\right) P\left(\mathcal{M}_E(\mathcal{D}, \{h_j\}_{j\in[m]}, \varepsilon) = i\right). \end{aligned}$$

Actually, Lemma 2 also gives a composition of the Gumbel mechanism and the EM. Based on this, the oneshot RNM algorithm which outputs the query and its index at the same time is formulated and presented in Algorithm 1.

Algorithm 1 Oneshot Report Noisy Min

Input: The database $\mathcal{D}$, functions $h_1, \ldots, h_m$ with sensitivity Δ and scale parameter γ
Output: The index j^* and approximation to $h_{j^*}(\mathcal{D})$

for $j = 1$ to m **do**
  Set $\tilde{h}_j = h_j(\mathcal{D}) + Z_j$, where Z_j is independently sampled from $\text{Gum}(0, \gamma)$;
end for
Solve $j^\star = \arg\min_{j\in[m]} \tilde{h}_j$ and compute $\tilde{h}_{j^\star}$.

It is evident that Algorithm 1 is more efficient than the RNM algorithm. Theorem 3 below provides the privacy assurance for this algorithm.

Theorem 3. *The oneshot RNM algorithm given in Algorithm 1 is $\left(\frac{\Delta}{\gamma} + \varepsilon, \delta(\varepsilon)\right)$-DP where $\delta(\varepsilon) = \left(e^{\frac{\Delta}{\gamma}+\varepsilon} - e^{\varepsilon}\right) e^{\frac{\frac{\Delta}{\gamma}+\varepsilon}{e^{-\frac{\Delta}{\gamma}}-1}}$ for any $\varepsilon > 0$.*

It is natural to design a new private top-k selection algorithm using the peeling algorithm under oneshot RNM proposed and shown in Algorithm 2. This top-k selection algorithm can be seen as the independent composition of k Gumbel mechanisms satisfied $\frac{\Delta}{\gamma}$-GumDP and k EMs satisfied $\left(\frac{\Delta}{\gamma}, 0\right)$-DP.

Algorithm 2 Peeling Algorithm under Oneshot Report Noisy Min

Input: database $\mathcal{D}$, functions $h_1, \ldots, h_m$ with sensitivity Δ, number of invocations k and scale parameter γ

for $j = 1$ to k **do**
  Let $\left(i_j, \tilde{h}_{i_j}\right)$ be returned by oneshot Report Noisy Min applied to $(\mathcal{D}, h_1, \ldots, h_m)$.
  Set $h_{i_j} \equiv +\infty$.
end for

Output: indices $i_1, \ldots, i_k$ and approximations to $h_{i_1}(\mathcal{D}), \ldots, h_{i_k}(\mathcal{D})$

Therefore, for characterizing the degree of privacy preservation of Algorithm 2, the optimal DP composition theorem of EMs are required and stated in the following Lemma 3.

Lemma 3. *(Dong et al., 2020) If $\mathcal{M}$ is a k-fold non-adaptive composition of ε-BR mechanisms, then it is $\left(\varepsilon_g, \delta_k^{EM}(\varepsilon_g)\right)$-DP with*

$$\delta_k^{EM}(\varepsilon_g) = \max_{0 \le \ell \le k} \sum_{i=0}^{k} \binom{k}{i} p_{t_\ell^*}^{k-i} \left(1 - p_{t_\ell^*}\right)^i \left(\mathrm{e}^{k t_\ell^* - i\varepsilon} - \mathrm{e}^{\varepsilon_g}\right)_+,$$

where $(a)_+$is defined as $\max\{a, 0\}$, $p_t = \frac{e^{-t} - e^{-\varepsilon}}{1 - e^{-\varepsilon}}$ and $t_\ell^* = \frac{\varepsilon_g + (\ell+1)\varepsilon}{k+1}$ where if $t_\ell^* \notin [0, \varepsilon]$, then we round it to the closest point in $[0, \varepsilon]$.

Theorem 4. *If $\{h_i\}_{i=1}^k$ are consistent, then Algorithm 2 ensures ($\varepsilon_1 + \varepsilon_2$, $\delta_k(\varepsilon_1) + \delta_k^{EM}(\varepsilon_2)$)-DP for all $\varepsilon_1, \varepsilon_2 > 0$, where the expressions for $\delta_k(\varepsilon_1)$ and $\delta_k^{EM}(\varepsilon_2)$ are respectively given in Theorem 2 and Lemma 3 in which $\mu = \varepsilon = \frac{\Delta}{\gamma}$.*

3.2 Privacy-Utility Performance Comparison

The most intuitive way to analyze the privacy-utility performance of the private top-k algorithm is to compare the variance of the added noise. Under the same privacy guarantee, a smaller noise variance indicates that the output values are closer to the true values, and also signifies the higher privay-utility performance. Therefore, in this subsection, for the peeling algorithm under RNM with Laplace or Gaussian mechanim and the peeling algorithm under oneshot RNM with Gumbel mechanism, we compare the corresponding noise variances of Laplace, Gaussian and Gumbel mechanisms in these algorithms.

To ensure fairness in comparison, let the peeling algorithm under RNM with Laplace, Gaussian mechanism and the peeling algorithm under oneshot RNM satisfy (ε, δ)-DP separately in private top-k selection. By formulating the following optimization problems, we obtain the minimum noise variance corresponding to several algorithms. The peeling algorithm under RNM with $\mathrm{Lap}\left(0, \frac{\Delta\sqrt{10k\ln(1/\delta)}}{\varepsilon}\right)$ is (ε, δ)-DP (Dwork et al., 2021). The variance of Laplace distribution in the peeling algorithm under RNM is $\frac{\varepsilon^2}{5k\Delta^2\ln(1/\delta)}$. Meanwhile, combined with the result in Cai et al. (2024), the peeling algorithm under RNM with $\mathcal{N}(0, \sigma^2)$ is (ε, δ)-DP where the variance of Gaussian distribution σ^2 satisfies

$$\begin{aligned} \min_{0 < \varepsilon_0 < \varepsilon} \quad & \sigma^2 \\ \text{s.t.} \quad & \delta_{\text{Gauss}}(\varepsilon_0) \le \delta, \end{aligned}$$

where $\delta_{\text{Gauss}}(\varepsilon_0)$ is provided in Corollary 1 of Dong et al. (2022) with $\mu = \frac{\sqrt{8k}\Delta}{\sigma}$. Lastly, for Gumbel mechanism, Algorithm 2 is (ε, δ)-DP utilizing Theorem 4 if the variance of Gumbel distribution

$\frac{\pi^2}{6}\gamma^2$ is satisfied that

$$\begin{aligned} \min_{\varepsilon_1,\varepsilon_2>0} \quad & \frac{\pi^2}{6}\gamma^2 \\ \text{s.t.} \quad & \varepsilon_1+\varepsilon_2 \le \varepsilon, \\ & \delta_k(\varepsilon_1)+\delta_k^{\text{EM}}(\varepsilon_2) \le \delta, \end{aligned}$$

where $\delta_k(\varepsilon_1)$ and $\delta_k^{\text{EM}}(\varepsilon_2)$ are respectively illustrated in Corollary 2 and Lemma 3 in which $\mu = \varepsilon = \frac{\Delta}{\gamma}$. Based on the above results, as shown in Fig. 3, by comparing at the same level of privacy protection, i.e., (ε, δ)-DP, the noise variances of Gumbel mechanism are smaller than those of both Laplace and Gaussian mechanisms which also implies that the application of the Gumbel mechanism offers superior privacy-utility performance.

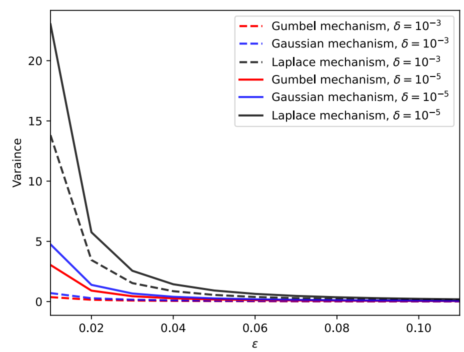


Figure 3: Noise variances comparison of Gaussian, Laplace mechanisms in the peeling algorithms under RNM and Gumbel mechanism in the peeling algorithm under oneshot RNM with ε varying, $k = 10$ and $\delta = 10^{-3}, 10^{-5}$.

4 Conclusion

In this paper, we provide a different privacy-preserving top-k selection algorithm with Gumbel mechanism, i.e., the peeling algorithm under oneshot RNM. Exploiting two special composition properties of the Gumbel mechanism, the oneshot RNM algorithm is designed, which is more efficient than the previous one as an algorithm that outputs the index and its query value without re-noising the query. To better characterize the privacy upper bound for the composition of k Gumbel mechanisms hidden within the peeling algorithm, GumDP is presented in this work as a novel family of f-DPs. The μ-GumDP analytically and tightly characterize the privacy of a single Gumbel mechanism, while the (k, μ)-GumDP is presented as an extension to characterize the composition of k Gumbel mechanisms under the assumption of consistency. To fairly compare different private top-k selection algorithms, two equivalent transformation relationships between GumDP and DP are provided. Based on the above equivalence relations, the variance-based comparison shows that the new Gumbel-based algorithm outperforms the original Laplace- and Gaussian-based algorithms under the same privacy guarantees. It is evident that the Gumbel mechanism holds advantages as compared to the Gaussian and Laplace mechanisms in privacy-preserving selection algorithms.

Due to the wide range of practical applications involving top-k selection algorithms, conducting in-depth and comprehensive research on private selection algorithms holds significant value. However, the topic of privacy selection still receives many challenges. In the process of extending the top-1 selection algorithm to the top-k selection algorithm, only the peeling algorithm is studied in this work. The oneshot algorithm can be subsequently used to further improve the performance. Moreover, the composition of k Gumbel mechanisms is taken under the strong assumption of consistency. Additionally, the practical application of this new private top-k selection algorithm remains to be explored.

REFERENCES

Martin Abadi, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS'16)*, pp. 308–318, Vienna, AU, 24-28 October 2016. ACM. doi: 10.1145/2976749.2978318.

John M. Abowd. The U.S. Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD'18)*, pp. 2867, London, UK, 19-23 August 2018. ACM. doi: 10.1145/3219819.3226070.

Krishna Acharya, Franziska Boenisch, Rakshit Naidu, and Juba Ziani. Personalized differential privacy for ridge regression, 2024. URL https://arxiv.org/abs/2401.17127.

Barış Alparslan and Sinan Yıldırım. Statistic selection and MCMC for differentially private Bayesian estimation. *Statistics and Computing*, 32(66), 2022. ISSN 1573-1375. doi: 10.1007/s11222-022-10129-8.

Jordan Awan and Zhanyu Wang. Simulation-based, finite-sample inference for privatized data. *Journal of the American Statistical Association*, pp. 1–14, 2024. ISSN 1537-274X. doi: 10.1080/01621459.2024.2427436.

Hai Brenner and Kobbi Nissim. Impossibility of differentially private universally optimal mechanisms. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pp. 71–80, Las Vegas, NV, USA, 23-26 October 2010. IEEE. doi: 10.1137/110846671.

Zhiqi Bu, Yu-Xiang Wang, Sheng Zha, and George Karypis. Automatic clipping: Differentially private deep learning made easier and stronger. In *the Thirty-Seventh Annual Conference on Neural Information Processing Systems*, volume 36, pp. 41727–41764, New Orleans, LA, USA, 10–16 December 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/8249b30d877c91611fd8c7aa6ac2b5fe-Paper-Conference.pdf.

Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In Martin Hirt and Adam Smith (eds.), *Theory of Cryptography*, volume 9985, chapter Lecture Notes in Computer Science, pp. 635–658. Springer, Berlin Heidelberg, 2016. ISBN 9783662536414. doi: 10.1007/978-3-662-53641-4_24.

Mark Bun, Cynthia Dwork, Guy N. Rothblum, and Thomas Steinke. Composable and versatile privacy via truncated CDP. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC'18)*, pp. 74–86, Los Angeles, CA, USA, 25-29 June 2018. ACM. doi: 10.1145/3188745.3188946.

Jianping Cai, Qingqing Ye, Haibo Hu, Ximeng Liu, and Yanggeng Fu. Boosting accuracy of differentially private continuous data release for federated learning. *IEEE Transactions on Information Forensics and Security*, 19:10287–10301, 2024. ISSN 1556-6021. doi: 10.1109/tifs.2024.3477325.

T. Tony Cai, Yichen Wang, and Linjun Zhang. The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *The Annals of Statistics*, 49(5):2825 – 2850, 2021. doi: 10.1214/21-AOS2058.

Edith Cohen and Xin Lyu. The target-charging technique for privacy analysis across interactive computations. In *Advances in Neural Information Processing Systems 36*, volume 36, pp. 62139–62168, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/c3fe2a07ec47b89c50e89706d2e23358-Paper-Conference.pdf.

Apple. Differential Privacy Team. Learning with privacy at scale., May 2017. URL https://machinelearning.apple.com/2017/12/06/learning-with-privacyat-scale.html.

Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*, pp. 3574–3583, Long Beach, CA, USA, 4-9 December 2017. Curran Associates, Inc.

Jinshuo Dong, David Durfee, and Ryan Rogers. Optimal differential privacy composition for exponential mechanisms. In *Proceedings of the 37th International Conference on Machine Learning (ICML'20)*, number 243, pp. 2597–2606, Vienna, AUT, 13-18 July 2020. PMLR.

Jinshuo Dong, Aaron Roth, and Weijie J. Su. Gaussian differential privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(1):3–37, 2022. ISSN 1467-9868. doi: 10.1111/rssb.12454.

David Durfee and Ryan Rogers. Practical differentially private top-k selection with pay-what-you-get composition. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems (NIPS'19)*, number 317, pp. 3532 – 3542, Vancouver, BC, CA, 8-14 December 2019. Curran Associates Inc.

Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2013. ISSN 1551-3068. doi: 10.1561/0400000042.

Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In Serge Vaudenay (ed.), *Advances in Cryptology - EUROCRYPT 2006*, volume 4004, chapter Lecture Notes in Computer Science, pp. 486–503. Springer Berlin Heidelberg, Petersburg, Russia, May 2006a. ISBN 9783540345473. doi: 10.1007/11761679_29.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the Third conference on Theory of Cryptography (TCC'6)*, pp. 265–284, New York, NY, USA, 4-7 March 2006b. Springer-Verlag. ISBN 3540327312. doi: 10.1007/11681878_14.

Cynthia Dwork, Guy N. Rothblum, and Salil Vadhan. Boosting and differential privacy. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pp. 51–60, Las Vegas, NV, USA, 23-26 October 2010. IEEE. doi: 10.1109/focs.2010.12.

Cynthia Dwork, Weijie Su, and Li Zhang. Differentially private false discovery rate control. *Journal of Privacy and Confidentiality*, 11(2), 2021. ISSN 2575-8527. doi: 10.29012/jpc.755.

Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. RAPPOR: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS'14)*, pp. 1054–1067, Scottsdale, AZ, USA, 3-7 November 2014. ACM. doi: 10.1145/2660267.2660348.

Erica L. Groshen and Daniel Goroff. Disclosure avoidance and the 2020 Census: What do researchers need to know? *Harvard Data Science Review*, (Special Issue 2), 2022. doi: 10.1162/99608f92.aed7f34f.

Moritz Hardt and Aaron Roth. Beyond worst-case analysis in private singular vector computation. In *Proceedings of the forty-fifth annual ACM symposium on Theory of Computing (STOC'13)*, pp. 331–340, Palo Alto, CA, USA, 1-4 June 2013. ACM. doi: 10.1145/2488608.2488650.

Ihab F. Ilyas, George Beskales, and Mohamed A. Soliman. A survey of top-k query processing techniques in relational database systems. *ACM Computing Surveys*, 40(4):1–58, 2008. ISSN 1557-7341. doi: 10.1145/1391729.1391730.

Christian Janos Lebeda and Jakub Tetek. Better differentially private approximate histograms and heavy hitters using the misra-gries sketch. *ACM Transactions on Database Systems*, 50(3), May 2025. ISSN 0362-5915. doi: 10.1145/3716375.

Shurong Lin, Mark Bun, Marco Gaboardi, Eric D. Kolaczyk, and Adam Smith. Differentially private confidence intervals for proportions under stratified random sampling. *Electronic Journal of Statistics*, 18(1):1455–1494, 2024. ISSN 1935-7524. doi: 10.1214/24-ejs2234.

Fang Liu. Generalized Gaussian mechanism for differential privacy. *IEEE Transactions on Knowledge and Data Engineering*, 31(4):747–756, 2019. ISSN 2326-3865. doi: 10.1109/tkde.2018.2845388.

Hanyang Liu, Yong Wang, Zhiqiang Zhang, Jiangzhou Deng, Chao Chen, and Leo Yu Zhang. Matrix factorization recommender based on adaptive Gaussian differential privacy for implicit feedback. *Information Processing Management*, 61(4):103720, 2024. ISSN 0306-4573. URL `https://www.sciencedirect.com/science/article/pii/S0306457324000803`.

Ryan McKenna and Daniel R Sheldon. Permute-and-flip: A new mechanism for differentially private selection. In *Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS '20)*, volume 1, pp. 193–203, Vancouver, BC, CA, 6-12 December 2020. Curran Associates Inc.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pp. 94–103, Providence, RI, USA, 20-23 October 2007. doi: 10.1109/focs.2007.66.

Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pp. 263–275, Santa Barbara, CA, USA, 21-25 August 2017. doi: 10.1109/csf.2017.11.

Gokularam Muthukrishnan and Sheetal Kalyani. Grafting Laplace and Gaussian distributions: A new noise mechanism for differential privacy. *IEEE Transactions on Information Forensics and Security*, 18:5359–5374, 2023. ISSN 1556-6021. doi: 10.1109/tifs.2023.3306159.

Rasmus Pagh, Lukas Retschmeier, Hao Wu, and Hanwen Zhang. Optimal bounds for private minimum spanning trees via input perturbation. *Proceedings of the ACM on Management of Data*, 3 (2), 2025. doi: 10.1145/3725240.

Gang Qiao, Weijie Su, and Li Zhang. Oneshot differentially private top-k selection. In *Proceedings of the 38th International Conference on Machine Learning (ICML'21)*, volume 139, pp. 8672–8681. PMLR, 18–24 July 2021.

Parastoo Sadeghi and Mehdi Korki. Offset-symmetric Gaussians for differential privacy. *IEEE Transactions on Information Forensics and Security*, 17:2394–2409, 2022. ISSN 1556-6021. doi: 10.1109/tifs.2022.3185770.

Larry Wasserman and Shuheng Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389, 2010. ISSN 1537-274X. doi: 10.1198/jasa.2009.tm08651.

Kang Wei, Jun Li, Ming Ding, Chuan Ma, Howard H. Yang, Farhad Farokhi, Shi Jin, Tony Q. S. Quek, and H. Vincent Poor. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15:3454–3469, 2020. ISSN 1556-6021. doi: 10.1109/tifs.2020.2988575.

Xintao Xia and Zhanrui Cai. Adaptive false discovery rate control with privacy guarantee. *Journal of Machine Learning Research*, 24(252):1–35, 2023.

Abbas Yazdinejad, Ali Dehghantanha, Hadis Karimipour, Gautam Srivastava, and Reza M. Parizi. A robust privacy-preserving federated learning model against model poisoning attacks. *IEEE Transactions on Information Forensics and Security*, 19:6693–6708, 2024. ISSN 1556-6021. doi: 10.1109/tifs.2024.3420126.

Lingchen Zhao, Qian Wang, Qin Zou, Yan Zhang, and Yanjiao Chen. Privacy-preserving collaborative deep learning with unreliable participants. *IEEE Transactions on Information Forensics and Security*, 15:1486–1500, 2020. ISSN 1556-6021. doi: 10.1109/tifs.2019.2939713.

Qinqing Zheng, Shuxiao Chen, Qi Long, and Weijie Su. Federated f-differential privacy. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 2251–2259. PMLR, 13–15 April 2021. URL `https://proceedings.mlr.press/v130/zheng21a.html`.

Yuqing Zhu and Yu-Xiang Wang. Adaptive private-k-selection with adaptive k and application to multi-label pate. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 151, pp. 5622–5635, Valencia, ES, 2022. PMLR.

A Proof of Equation (4)

Let $\text{Gum}(0,1)$ and $\text{Gum}(\mu,1)$ be the distributions of $\mathcal{M}_{\text{Gum}}(\mathcal{D})$ and $\mathcal{M}_{\text{Gum}}(\mathcal{D}')$ in (2) respectively, and p_0 and p_1 be the PDFs of $\text{Gum}(0,1)$ and $\text{Gum}(\mu,1)$ respectively. For the hypothesis testing problem (2), the likelihood ratio is

$$\frac{p_1(x)}{p_0(x)} = \frac{e^{x-\mu-e^{x-\mu}}}{e^{x-e^x}} = e^{-\mu+e^x(1-e^{-\mu})},$$

which is a monotone increasing function in x. Thus, the rejection domain in (2) is $W = \{X > t\}$ where X is random sample from Gumbel distribution and $t \in \mathcal{R}$. The corresponding type I and type II errors are

$$\alpha(t) = P(X > t | X \sim \text{Gum}(0,1)) = e^{-e^t},$$
$$\beta(t) = P(X < t | X \sim \text{Gum}(\mu,1)) = 1 - e^{-e^{t-\mu}}.$$

Solving $\alpha(t) = \alpha$ yields $t = \ln(-\ln \alpha)$. So

$$T(\text{Gum}(0,1), \text{Gum}(\mu,1))(\alpha) = 1 - e^{-e^{-\mu}(-\ln \alpha)} = 1 - (\alpha)^{e^{-\mu}}.$$

And

$$T(\text{Gum}(\mu,1), \text{Gum}(0,1))(\alpha) = T(\text{Gum}(0,1), \text{Gum}(\mu,1))^{-1}(\alpha) = (1-\alpha)^{e^{\mu}}.$$

Let α_1 be unique solution of $T(\text{Gum}(0,1), \text{Gum}(\mu,1))'(\alpha) = -1$ and $\alpha_2 = T(\text{Gum}(0,1), \text{Gum}(\mu,1))(\alpha_1)$. Then, $\alpha_1 = e^{\frac{\mu}{e^{-\mu}-1}}, \alpha_2 = 1 - e^{\frac{\mu e^{-\mu}}{e^{-\mu}-1}}$. Similar to Eq.(13) in Dong et al. (2022),

$$\begin{aligned} B_\mu(\alpha) &= \min\{T(\text{Gum}(0,1), \text{Gum}(\mu,1)), T(\text{Gum}(\mu,1), \text{Gum}(0,1))\}^{**} \\ &= \begin{cases} T(\text{Gum}(0,1), \text{Gum}(\mu,1))(\alpha), & \alpha \in [0, \alpha_1), \\ \alpha_1 - \alpha + T(\text{Gum}(0,1), \text{Gum}(\mu,1))(\alpha_1), & \alpha \in [\alpha_1, \alpha_2), \\ T(\text{Gum}(\mu,1), \text{Gum}(0,1))(\alpha), & \alpha \in [\alpha_2, 1]. \end{cases} \end{aligned}$$

B Proof of Theorem 1

For any two neighboring databases $\mathcal{D}$ and $\mathcal{D}'$ and $\gamma \geq \Delta h/\mu$, we get

$$\begin{aligned} T(\mathcal{M}_{\text{Gum}}(\mathcal{D}), \mathcal{M}_{\text{Gum}}(\mathcal{D}')) &= T(\text{Gum}(h(\mathcal{D}),\gamma), \text{Gum}(h(\mathcal{D}'),\gamma)) \\ &\geq \min\big\{T(\text{Gum}(0,1), \text{Gum}(|h(\mathcal{D}) - h(\mathcal{D}')|/\gamma, 1)), \\ &\qquad T(\text{Gum}(|h(\mathcal{D}) - h(\mathcal{D}')|/\gamma, 1), \text{Gum}(0,1))\big\} \\ &\geq B_{\frac{|h(\mathcal{D})-h(\mathcal{D}')|}{\gamma}}. \end{aligned}$$

By the definition of sensitivity, $|h(\mathcal{D}) - h(\mathcal{D}')| \leq \Delta h \leq \gamma\mu$. Therefore, we get

$$T(\mathcal{M}_{\text{Gum}}(\mathcal{D}), \mathcal{M}_{\text{Gum}}(\mathcal{D}')) \geq B_{\frac{|h(\mathcal{D})-h(\mathcal{D}')|}{\gamma}} \geq B_\mu.$$

C Proof of Corollary 1

Based on the equivalent conversion of f-DP and DP and the symmetry of the function B_μ, μ-GumDP is equal to $(\varepsilon, 1 + B_\mu^*(-e^\varepsilon))$-DP. Therefore, we only need to compute the $B_\mu^*(-e^\varepsilon)$. From the definition of convex conjugate function, $B_\mu^*(y) = \sup_{x\in[0,1]}(yx - B_\mu(x))$. And, from the shape of B_μ, the supremum is obtained only at the unique critical point when $y \in (-\infty, -1)$. From

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}x}(yx - B_\mu(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(yx - 1 + x^{e^{-\mu}}) \\ &= y + e^{-\mu}x^{e^{-\mu}-1}, \end{aligned}$$

we have $x = (-e^\mu y)^{\frac{1}{e^{-\mu}-1}}$. Then,

$$B_\mu^*(y) = y(-e^\mu y)^{\frac{1}{e^{-\mu}-1}} + (-e^\mu y)^{\frac{e^{-\mu}}{e^{-\mu}-1}} - 1, \qquad y \in (-\infty, -1).$$

Setting $y = -e^\varepsilon$ implies $B_\mu^*(-e^\varepsilon) = (e^{\mu+\varepsilon} - e^\varepsilon)e^{\frac{\mu+\varepsilon}{e^{-\mu}-1}} - 1$. Thus, this corollary holds.

D Proof of Equation (7)

Let $\mathrm{Gum}(0,1)$ and $\mathrm{Gum}(\mu,1)$ be the distributions of $\mathcal{M}_{\mathrm{Gum}}(\mathcal{D})$ and $\mathcal{M}_{\mathrm{Gum}}(\mathcal{D}')$ in (6) respectively, and p_0 and p_1 be the PDFs of $(y_1, y_2, \ldots, y_k)$ under H_0 and H_1 respectively. For the hypothesis testing problem (6), the likelihood ratio is

$$\frac{p_1(x_1, x_2, \ldots, x_k)}{p_0(x_1, x_2, \ldots, x_k)} = \prod_{i=1}^{k} \frac{e^{(x_i-\mu)-e^{(x_i-\mu)}}}{e^{x_i - e^{x_i}}} = e^{-k\mu} e^{(1-e^{-\mu})\sum_{i=1}^{k} e^{x_i}}.$$

It is a monotonically increasing function in $\sum_{i=1}^{k} e^{\frac{x_i}{\beta}}$. Thus, the rejection domain in (6) is $W = \{\sum_{i=1}^{k} e^{\frac{X_i}{\beta}} > t\}$ where X_i is a random sample and $t > 0$. The corresponding type I and type II errors respectively are

$$\alpha(t) = P\left(\sum_{i=1}^{k} e^{X_i} > t \middle| \{X_i\}_{i=1}^{k} \overset{\text{i.i.d.}}{\sim} \mathrm{Gum}(0,1)\right), \tag{8}$$

$$\beta(t) = P\left(\sum_{i=1}^{k} e^{X_i} < t \middle| \{X_i\}_{i=1}^{k} \overset{\text{i.i.d.}}{\sim} \mathrm{Gum}(\mu,1)\right). \tag{9}$$

To facilitate the analysis, let $y_i = e^{x_i}, i = 1, 2, \ldots, k$. When $X_i \sim \mathrm{Gum}(0,1)$, $P_{Y_i}(y_i \le t) = P_{X_i}(e^{x_i} \le t) = P_{X_i}(x_i \le \ln t) = 1 - e^{-t}$, the distribution of Y_i is the exponential distribution with parameter 1, denoted as $\mathrm{Exp}(1)$. Since $\{X_i\}_{i=1}^{k}$ are distributed independently and identically, so are $\{Y_i\}_{i=1}^{k}$. Based on the nature of the exponential distribution, $\sum_{i=1}^{k} Y_i \sim \Gamma(k,1)$ where $\Gamma(k,1)$ denotes the Gamma distribution with shape parameter k and inverse scale parameter 1. Similarly, when $X_i \sim \mathrm{Gum}(\mu,1)$, $P_{Y_i}(y_i \le t) = 1 - e^{-e^{\ln t - \mu}} = 1 - e^{-e^{-\mu}t}$, so $Y_i \sim \mathrm{Exp}(e^{-\mu})$ and $\sum_{i=1}^{k} Y_i \sim \Gamma(k, e^{-\mu})$. Then, (8) and (9) respectively become

$$\begin{aligned}
\alpha(t) &= P\left(\sum_{i=1}^{k} Y_i > t \middle| \{Y_i\}_{i=1}^{k} \overset{\text{i.i.d.}}{\sim} \mathrm{Exp}(1)\right) \\
&= P(\xi > t | \xi \sim \Gamma(k,1)) \\
&= e^{-t}\left(1 + \sum_{i=1}^{k-1} \frac{t^i}{i!}\right), \\
\beta(t) &= P\left(\sum_{i=1}^{k} Y_i < t | \{Y_i\}_{i=1}^{k} \overset{\text{i.i.d.}}{\sim} \mathrm{Exp}(e^{-\mu})\right) \\
&= P(\xi < t | \xi \sim \Gamma(k, e^{-\mu})) \\
&= 1 - e^{-te^{-\mu}}\left(1 + \sum_{i=1}^{k-1} \frac{(te^{-\mu})^i}{i!}\right).
\end{aligned}$$

Due to $Y \sim \chi^2(2k)$, $F_Y(x) = 1 - e^{-\frac{x}{2}}\left(1 + \sum_{i=1}^{k-1} \frac{(\frac{x}{2})^i}{i!}\right)$. So, $\alpha(t) = 1 - F_Y(2t)$ and $\beta(t) = F_Y\left(\frac{1}{2}F_Y^{-1}(1-\alpha(t))2e^{-\mu}\right) = F_Y\left(F_Y^{-1}(1-\alpha(t))\, e^{-\mu}\right)$, which yields

$$T(\mathrm{Gum}(0,1)^k, \mathrm{Gum}(\mu,1)^k)(\alpha) = F_Y\left(F_Y^{-1}(1-\alpha)\, e^{-\mu}\right).$$

Analogously, under the original hypothesis obeying $\mathrm{Gum}(\mu,1)^k$ and the alternative hypothesis obeying $\mathrm{Gum}(0,1)^k$, the type I and type II errors are respectively $\alpha(t) = 1 - e^{-te^{-\mu}}\left(1 + \sum_{i=1}^{k-1} \frac{(te^{-\mu})^i}{i!}\right)$ and $\beta(t) = e^{-t}\left(1 + \sum_{i=1}^{k-1} \frac{t^i}{i!}\right)$. Easily obtained, $\alpha(t) = F_Y(2e^{-\mu}t)$ and $\beta(t) = 1 - F_Y\left(2F_Y^{-1}(\alpha(t))\frac{1}{2}e^{\mu}\right) = 1 - F_Y\left(F_Y^{-1}(\alpha(t))\, e^{\mu}\right)$, which yields

$$T(\mathrm{Gum}(\mu,1)^k, \mathrm{Gum}(0,1)^k)(\alpha) = 1 - F_Y\left(F_Y^{-1}(\alpha(t))\, e^{\mu}\right).$$

Being identical to the proof of Equation (4), let α_1 be unique solution of $T(\mathrm{Gum}(0,1)^k, \mathrm{Gum}(\mu,1)^k)'(\alpha) = -1$ and $\alpha_2 = T(\mathrm{Gum}(0,1)^k, \mathrm{Gum}(\mu,1)^k)(\alpha_1)$. Taking the derivative of $T(\mathrm{Gum}(0,1)^k, \mathrm{Gum}(\mu,1)^k)(\alpha)$ and setting it to -1 yields

$$\begin{aligned}
-1 &= \frac{\mathrm{d}}{\mathrm{d}\alpha} T(\mathrm{Gum}(0,1)^k, \mathrm{Gum}(\mu,1)^k)(\alpha) = \frac{\mathrm{d}}{\mathrm{d}\alpha} F_Y\left(F_Y^{-1}(1-\alpha)\, e^{-\mu}\right) \\
&= \frac{-p_Y\left(F_Y^{-1}(1-\alpha)\, e^{-\mu}\right) e^{-\mu}}{p_Y\left(F_Y^{-1}(1-\alpha)\right)} \\
&= -e^{\frac{1}{2} F_Y^{-1}(1-\alpha)\left(1-e^{-\mu}\right) - k\mu},
\end{aligned}$$

where p_Y is the PDF of $\chi^2(2k)$. Then, $\alpha_1 = 1 - F_Y\left(\frac{2k\mu}{1-e^{-\mu}}\right)$ and $\alpha_2 = F_Y(F_Y^{-1}(1-\alpha_1)e^{-\mu}) = F_Y\left(\frac{2k\mu e^{-\mu}}{1-e^{-\mu}}\right)$. This proof is finished.

E Proof of Theorem 2

Because of the consistence of $\{h_i\}_{i=1}^k$,

$$\begin{aligned}
& T(\mathcal{M}(\mathcal{D}), \mathcal{M}(\mathcal{D}')) \\
&= T(\mathrm{Gum}(h_1(\mathcal{D}),\gamma) \times \cdots \times \mathrm{Gum}(h_k(\mathcal{D}),\gamma), \mathrm{Gum}(h_1(\mathcal{D}'),\gamma) \times \cdots \times \mathrm{Gum}(h_k(\mathcal{D}'),\gamma)) \\
&= T(\mathrm{Gum}(0,1)^k, \mathrm{Gum}((h_1(\mathcal{D}') - h_1(\mathcal{D}))/\gamma, 1) \times \cdots \times \mathrm{Gum}((h_k(\mathcal{D}') - h_k(\mathcal{D}))/\gamma, 1)) \\
&\geq T(\mathrm{Gum}(0,1)^k, \mathrm{Gum}(\mathrm{sign}(h_1(\mathcal{D}') - h_1(\mathcal{D}))\mu, 1)) \times \cdots \times \mathrm{Gum}(\mathrm{sign}(h_k(\mathcal{D}') - h_k(\mathcal{D}))\mu, 1)) \\
&\geq \min\{T(\mathrm{Gum}(0,1)^k, \mathrm{Gum}(\mu,1)^k), T(\mathrm{Gum}(\mu,1)^k, \mathrm{Gum}(0,1)^k)\} \\
&\geq B_\mu^k.
\end{aligned}$$

The proof is completed.

F Proof of Corollary 2

Similarly to the proof of Theorem 1, by the symmetry of the function B_μ^k, (k,μ)-GumDP is equal to $(\varepsilon, 1 + B_\mu^{k*}(-e^\varepsilon))$-DP. Therefore, we only need to compute the $B_\mu^{k*}(-e^\varepsilon)$. Before that we need to know that the PDF of Z is

$$p_Z(x) = \begin{cases} \frac{1}{2^k \Gamma(k)} x^{k-1} e^{-\frac{x}{2}}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

It is easy to get $B_\mu^{k*}(y) = \sup_{x \in [0,1]} (yx - B_\mu^k(x))$. And, from the shape of B_μ^k, the supremum is obtained only at the unique critical point when $y \in (-\infty, -1)$. Taking the derivative of the objective function and setting it to zero yields, when $y \in (-\infty, -1)$,

$$\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}x}(yx - B_\mu^k(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(yx - F_Z(F_Z^{-1}(1-x)e^{-\mu}) \\
&= y + \frac{e^{-\mu} p_Z(F_Z^{-1}(1-x)e^{-\mu})}{p_Z(F_Z^{-1}(1-x))} \\
&= y + e^{\frac{F_Z^{-1}(1-x)}{2}(1-e^{-\mu}) - k\mu}.
\end{aligned}$$

Getting $x = 1 - F_Z\left(\frac{2(k\mu + \ln(-y))}{1-e^{-\mu}}\right)$ and taking it into B_μ^{k*} leads to

$$\begin{aligned}
B_\mu^{k*}(x) &= y\left(1 - F_Z\left(\frac{2(k\mu + \ln(-y))}{1-e^{-\mu}}\right)\right) - F_Z\left(F_Z^{-1}\left(F_Z\left(\frac{2(k\mu + \ln(-y))}{1-e^{-\mu}}\right)\right) e^{-\mu}\right) \\
&= y - yF_Z\left(\frac{2(k\mu + \ln(-y))}{1-e^{-\mu}}\right) - F_Z\left(\frac{2(k\mu + \ln(-y))e^{-\mu}}{1-e^{-\mu}}\right), \qquad y \in (-\infty, -1).
\end{aligned}$$

When $y = -e^\varepsilon$, $B_\mu^{k*}(-e^\varepsilon) = -e^\varepsilon + e^\varepsilon F_Z\left(\frac{2(k\mu+\varepsilon)}{1-e^{-\mu}}\right) - F_Z\left(\frac{2(k\mu+\varepsilon)e^{-\mu}}{1-e^{-\mu}}\right)$.

G PROOF OF LEMMA 1

Start by analyzing the distribution of $\mathcal{M}^{m}_{\text{minGum}}(\mathcal{D})$, if $\{\eta_i\}_{i=1}^{m} \overset{\text{i.i.d.}}{\sim} \text{Gum}(0, \frac{\Delta}{\varepsilon})$,

$$\begin{aligned}
P\left(\mathcal{M}^{m}_{\text{minGum}}(\mathcal{D}) > t\right) &= P\left(\min_{i\in[m]} \left\{\mathcal{M}^{(i)}_{\text{Gum}}(\mathcal{D})\right\}\right) \\
&= P\left(\min_{i\in[m]} \{h_i(\mathcal{D}) + \eta_i\} > t\right) \\
&= \prod_{i=1}^{m} P(h_i(\mathcal{D}) + \eta_i > t) \\
&= \prod_{i=1}^{m} e^{-e^{\frac{t-h_i(\mathcal{D})}{\frac{\Delta}{\varepsilon}}}} \\
&= e^{-e^{\frac{t+\frac{\Delta}{\varepsilon}\ln\left(\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D})}{\frac{\Delta}{\varepsilon}}}\right)}{\frac{\Delta}{\varepsilon}}}}.
\end{aligned}$$

It is easy to see that $\mathcal{M}^{m}_{\text{minGum}}$ as a Gumbel mechanism and $\mathcal{M}^{m}_{\text{minGum}}(\mathcal{D})$ distributed from

$$\text{Gum}\left(-\frac{\Delta}{\varepsilon}\ln\left(\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D})}{\frac{\Delta}{\varepsilon}}}\right), \frac{\Delta}{\varepsilon}\right).$$

Let $g(\mathcal{D}) = -\frac{\Delta}{\varepsilon}\ln\left(\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D})}{\frac{\Delta}{\varepsilon}}}\right)$. Then, $\mathcal{M}^{m}_{\text{minGum}}(\mathcal{D}) = g(\mathcal{D}) + \eta$ where $\eta \sim \text{Gum}(0, \frac{\Delta}{\varepsilon})$. Because of $\Delta = \max_{i\in[k]} \max_{\mathcal{D},\mathcal{D}'} |h_i(\mathcal{D}) - h_i(\mathcal{D}')|$, $|h_i(\mathcal{D}) - h_i(\mathcal{D}')| \le \Delta$. So,

$$e^{-\varepsilon}\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D})}{\frac{\Delta}{\varepsilon}}} \le \sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D}')}{\frac{\Delta}{\varepsilon}}} \le e^{\varepsilon}\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D})}{\frac{\Delta}{\varepsilon}}}.$$

Then, for any $i \in [k]$,

$$\begin{aligned}
|g(\mathcal{D}) - g(\mathcal{D}')| &= \left|-\frac{\Delta}{\varepsilon}\ln\left(\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D})}{\frac{\Delta}{\varepsilon}}}\right) + \frac{\Delta}{\varepsilon}\ln\left(\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D}')}{\frac{\Delta}{\varepsilon}}}\right)\right| \\
&= \left|\frac{\Delta}{\varepsilon}\ln\left(\frac{\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D}')}{\frac{\Delta}{\varepsilon}}}}{\sum_{i=1}^{m} e^{-\frac{h_i(\mathcal{D})}{\frac{\Delta}{\varepsilon}}}}\right)\right| \le \left|\frac{\Delta}{\varepsilon}\ln e^{\varepsilon}\right| = \Delta.
\end{aligned}$$

So $|g(\mathcal{D}) - g(\mathcal{D}')| \le \Delta$ holds in the general case. Because of $\max_{\mathcal{D},\mathcal{D}'} |g(\mathcal{D}) - g(\mathcal{D}')| \le \Delta$, $\mathcal{M}^{m}_{\text{minGum}}$ satisfies μ-GumDP using Theorem 2.

H PROOF OF LEMMA 2

Let $\{\eta_j\}_{j=1}^{m}$ be i.i.d.copied from $\text{Gum}(0, \frac{\Delta}{\varepsilon})$ and $\mathcal{M}^{(j)}_{\text{Gum}}(\mathcal{D}) = h_j(\mathcal{D}) + \eta_j$, $j \in [m]$. Then, for any $i \in [m]$ and $t \in \mathcal{R}$,

$$\begin{aligned}
&P\left(M^{*}_{\text{Gum}}(D) = (i, h_i(D) + \eta_i) \in [m] \times (t, +\infty)\right) \\
&= \int_{t}^{\infty} p\left(u_i - h_i(D), 0, \frac{\Delta}{\varepsilon}\right) \prod_{j\in[m]\setminus\{i\}} \left(1 - F\left(u_i - h_j(D), 0, \frac{\Delta}{\varepsilon}\right)\right) du_i \\
&= e^{-\sum_{j=1}^{m} e^{\frac{\varepsilon(t-h_j(D))}{\Delta}}} \cdot \frac{e^{\frac{-\varepsilon h_i(D)}{\Delta}}}{\sum_{j=1}^{m} e^{\frac{-\varepsilon h_j(D)}{\Delta}}} \\
&= P\left(\min_{j\in[m]} \left\{\mathcal{M}^{(j)}_{\text{Gum}}(\mathcal{D})\right\} > t\right) \cdot P\left(\mathcal{M}_{\text{E}}(\mathcal{D}, \{h_j\}_{j\in[m]}, \varepsilon) = i\right) \\
&= P\left(\mathcal{M}^{m}_{\text{minGum}}(\mathcal{D}) > t\right) \cdot P\left(\mathcal{M}_{\text{E}}(\mathcal{D}, \{h_j\}_{j\in[m]}, \varepsilon) = i\right),
\end{aligned}$$

where p and F are the PDF and CDF of $\mathrm{Gum}(0, \frac{\Delta}{\varepsilon})$ respectively. With Lemma 1, the proof is complete.

I PROOF OF THEOREM 3

By Lemma 2, Algorithm 1 is equivalent to the independent composition of a mechanism satisfied $\frac{\Delta}{\gamma}$-GumDP and the EM $\mathcal{M}_{\mathrm{E}}\left(\mathcal{D}, \{h_j\}_{j\in[m]}, \frac{\Delta}{\gamma}\right)$. Because of the monotone, $\mathcal{M}_{\mathrm{E}}$ is $\left(\frac{\Delta}{\gamma}, 0\right)$-DP. And, due to Theorem 1 and the basic composition theorem in Dwork & Roth (2013), the algorithm is $\left(\frac{\Delta}{\gamma} + \varepsilon, \delta(\varepsilon)\right)$-DP where $\delta(\varepsilon) = \left(e^{\frac{\Delta}{\gamma}+\varepsilon} - e^{\varepsilon}\right) e^{\frac{\frac{\Delta}{\gamma}+\varepsilon}{e^{-\frac{\Delta}{\gamma}}-1}}$ for any $\varepsilon > 0$.

J PROOF OF THEOREM 4

Based on the result in Dong et al. (2020), the ε-BR is equal to ε-DP when the functions $\{h_i\}_{i\in[m]}$ are monotone. Similarly to Theorem 3, it can be proved by Theorem 2, Lemma 3 and the basic composition theorem.