Tighter Generalization Bounds for Iterative Privacy-Preserving Algorithms (Supplementary material)

Fengxiang He[*1, 2] **Bohan Wang**[*1] **Dacheng Tao**[1, 2]

[1]School of Computer Science, the University of Sydney, Australia
[2]JD Explore Academy, JD.com, China

A PROOFS OF GENERALIZATION BOUNDS VIA DIFFERENTIAL PRIVACY

This appendix collects all the proofs of theorems and technical lemma in Section 3. It is organized as follows: (1) in Appendix A.1, we prove Theorem 2, Theorem 3, and Lemma 1, which are the preparations of the proof of Theorem 1; and (2) in Appendix A.2 we prove Theorem 1.

A.1 PROOF OF THEOREM 2, THEOREM 3, AND LEMMA 1

We first present the proof of Theorem 2.

Proof of Theorem 2. We first rewrite the expectation of the empirical risk of algorithm $\mathcal{B}$ can be rewritten as

$$\begin{aligned}
&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right] = \mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\mathbb{E}_{z\sim S_{i_{\mathcal{B}(S)}}}\left[\ell\left(h_{\mathcal{B}(S)},z\right)\right]\right]\right]\\
\overset{(*)}{=}&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S}\left[\ell\left(h_{\mathcal{B}(S)},z_{i_{\mathcal{B}(S)}}\right)\right]\right]\right] = \mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\ell\left(h_{\mathcal{B}(S)},z_{i_{\mathcal{B}(S)}}\right)\right]\right]\right]\\
=&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\ell\left(h_{\mathcal{B}(S)},z_{i_{\mathcal{B}(S)}}\right)\right]\right]\right]\\
=&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S}\left[\int_0^1\mathbb{P}\left(\ell\left(h_{\mathcal{B}(S)},z_{i_{\mathcal{B}(S)}}\right)\leq t\right)dt\right]\right]\\
=&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S}\left[\sum_{i=1}^k\int_0^1\mathbb{P}\left(\ell\left(h_{\mathcal{B}(S)},z_i\right)\leq t, i_{\mathcal{B}(S)}=i\right)\mathrm{d}t\right]\right],
\end{aligned}$$

where $\vec{z}$ in the right side of eq.$(*)$ is defined as $\{z_1,\cdots,z_k\}$, and z_i is uniformly selected from S_i. Since $\mathcal{B}$ is (ε,δ)-

[*]The authors contributed equally.

differentially private, we further have

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S}\left[\sum_{i=1}^{k}\int_0^1 \mathbb{P}\left(\ell\left(h_{\mathcal{B}(S)},z_i\right)\le t, i_{\mathcal{B}(S)}=i\right)\mathrm{d}t\right]\right]\\
\le&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S,z_0\sim D}\left[\sum_{i=1}^{k}\int_0^1 \left(e^{\varepsilon}\mathbb{P}\left(\ell\left(h_{\mathcal{B}(S^{z_i:z_0})},z_i\right)\le t, i_{\mathcal{B}(S^{z_i:z_0})}=i\right)+\delta\right)\mathrm{d}t\right]\right]\\
=&e^{\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S,z_0\sim D}\left[\sum_{i=1}^{k}\int_0^1 \mathbb{P}\left(\ell\left(h_{\mathcal{B}(S^{z_i:z_0})},z_i\right)\le t, i_{\mathcal{B}(S^{z_i:z_0})}=i\right)\mathrm{d}t\right]\right]+k\delta\\
=&\sum_{i=1}^{k}e^{\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\vec{z}\sim S,z_0\sim D}\left[\int_0^1 \mathbb{P}\left(\ell\left(h_{\mathcal{B}(S^{z_i:z_0})},z_i\right)\le t, i_{\mathcal{B}(S^{z_i:z_0})}=i\right)\mathrm{d}t\right]\right]+k\delta\\
=&\sum_{i=1}^{k}e^{\varepsilon}\mathop{\mathbb{E}}_{S'\sim\mathcal{D}^{kN-1}}\left[\mathop{\mathbb{E}}_{z_i\sim D,z_0\sim D}\left[\int_0^1 \mathbb{P}\left(\ell\left(h_{\mathcal{B}(S'\cup\{z_0\})},z_i\right)\le t, i_{\mathcal{B}(S'\cup\{z_0\})}=i\right)\mathrm{d}t\right]\right]+k\delta.
\end{aligned}
$$

Let $S=S'\cup\{z_0\}$ and $z=z_i$ (it is without less of generality since all z_i is i.i.d. drawn from $\mathcal{D}$). Since $S'\cup\{z_0\}\sim\mathcal{D}^{kN}$, we have

$$
\begin{aligned}
&\sum_{i=1}^{k}e^{\varepsilon}\mathop{\mathbb{E}}_{S'\sim\mathcal{D}^{kN-1}}\left[\mathop{\mathbb{E}}_{z_i\sim D,z_0\sim D}\left[\int_0^1 \mathbb{P}\left(\ell\left(h_{\mathcal{B}(S'\cup\{z_0\})},z_i\right)\le t, i_{\mathcal{B}(S'\cup\{z_0\})}=i\right)\mathrm{d}t\right]\right]+k\delta\\
=&\sum_{i=1}^{k}e^{\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{z\sim\mathcal{D}}\left[\int_0^1 \mathbb{P}\left(\ell\left(h_{\mathcal{B}(S)},z\right)\le t, i_{\mathcal{B}(S)}=i\right)\mathrm{d}t\right]\right]+k\delta\\
=&e^{\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{z\sim\mathcal{D}}\left[\int_0^1 \mathbb{P}\left(\ell\left(h_{\mathcal{B}(S)},z\right)\le t\right)\mathrm{d}t\right]\right]+k\delta\\
=&e^{\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{z\sim\mathcal{D}}\left[\mathbb{E}_{\mathcal{B}(S)}\left[\ell\left(h_{\mathcal{B}(S)},z\right)\right]\right]\right]+k\delta.
\end{aligned}
$$

Therefore, we have

$$
\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right]\le k\delta+e^{\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}[\mathop{\mathbb{E}}_{\mathcal{B}(S)}[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)]]. \tag{1}
$$

Rearranging eq. (1), we have

$$
\begin{aligned}
e^{-\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right]&\le e^{-\varepsilon}k\delta+\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\right]\right]\\
-\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}[\mathop{\mathbb{E}}_{\mathcal{B}(S)}[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)]]&\le e^{-\varepsilon}k\delta-e^{-\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right],
\end{aligned}
$$

which further leads to

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right]-\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)]\right]\\
\le&e^{-\varepsilon}k\delta-e^{-\varepsilon}\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right]+\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right]\\
\le&1-e^{-\varepsilon}+e^{-\varepsilon}k\delta.
\end{aligned}
$$

The other side of the inequality can be similarly obtained.

The proof is completed. □

We then prove Theorem 3 based on Theorem 2.

Proof of Theorem 3. By Theorem 2, we have that

$$\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\right]\right]\leq e^{-\varepsilon}k\delta+1-e^{-\varepsilon}+\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right].$$

Since $\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\geq 0$, we have that for any $\alpha>0$,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\right]\right]\\
\geq&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\mathbb{I}_{\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\geq\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha}\right]\right]\\
\geq&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[(\alpha+\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right))\mathbb{I}_{\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\geq\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha}\right]\right].
\end{aligned}$$

Furthermore, by splitting $\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)]\right]$ into two parts, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\right]\right]-\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\right]\right]\\
\geq&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[(\alpha+\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right))\mathbb{I}_{\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\geq\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha}\right]\right]\\
&-\left(\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\mathbb{I}_{\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\geq\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha}\right]\right]\right.\\
&\left.+\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\mathbb{I}_{\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)<\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha}\right]\right]\right)\\
=&\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\alpha\mathbb{I}_{\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\geq\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha}\right]\right]-\mathop{\mathbb{E}}_{S\sim\mathcal{D}^{kN}}\left[\mathop{\mathbb{E}}_{\mathcal{B}(S)}\left[\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)\mathbb{I}_{\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)<\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha}\right]\right]\\
\geq&\alpha\mathbb{P}\left(\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\geq\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha\right)-\mathbb{P}\left(\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)<\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha\right)\\
=&\alpha\left(1-\mathbb{P}\left(\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)<\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha\right)\right)-\mathbb{P}\left(\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)<\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha\right).
\end{aligned}$$

By simple rearranging, we have

$$\mathbb{P}\left(\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)\leq\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)+\alpha\right)\geq\frac{\alpha-(e^{-\varepsilon}k\delta+1-e^{-\varepsilon})}{1+\alpha}.\tag{2}$$

We then calculate the right side of eq.(2) according to the two cases mentioned in Lemma 3:

(1). If $\varepsilon\leq\frac{17}{50}$, let $\alpha=e^{-\varepsilon}k\delta+1.7\varepsilon$. We then have

$$\frac{\alpha-(e^{-\varepsilon}k\delta+1-e^{-\varepsilon})}{1+\alpha}\geq\frac{\alpha-(e^{-\varepsilon}k\delta+\varepsilon)}{1+\alpha}=\frac{0.7\varepsilon}{1+e^{-\varepsilon}k\delta+1.7\varepsilon}\geq\frac{0.7\varepsilon}{1+(1.7+\frac{1}{1.7})\varepsilon}\geq\frac{50}{127}\varepsilon.$$

(2). If $\varepsilon>\frac{17}{50}$, let $\alpha=e^{-\varepsilon}k\delta+1.1(1-e^{-\varepsilon})$. We then have

$$\frac{\alpha-(e^{-\varepsilon}k\delta+1-e^{-\varepsilon})}{1+\alpha}=\frac{0.1(1-e^{-\varepsilon})}{1+e^{\varepsilon}k\delta+1.1(1-e^{-\varepsilon})}\geq\frac{0.1(1-e^{-\varepsilon})}{1+1.19(1-e^{-\varepsilon})}\geq\frac{1-e^{-\varepsilon}}{219}.$$

The proof is completed. □

Before moving on to the proof of Lemma 1, we present a basic lemma about the distance between the output of a exponential mechanism (defined as Definition 4) and the maximum value of the corresponding utility function:

Lemma 1 (c.f. Corollary 3.12, [Dwork and Roth, 2014])**.** *For any fixed sample set S and an exponential mechanism $\mathcal{E}(S,u,\mathcal{I},\varepsilon)$, we have*

$$\mathbb{P}\left[u\left(S,\mathcal{E}(S,u,\mathcal{I},\varepsilon)\right)\leq\max_{i\in\mathcal{I}}u(S,i)-\frac{2\Delta u}{\varepsilon}(\ln(|\mathcal{I}|)+t)\right]\leq e^{-t}.$$

We are now ready to prove Lemma 1.

Proof of Lemma 1. We prove the lemma respectively for the two cases: (1). $\varepsilon \leq \frac{1}{5}$, and (2). $\varepsilon > \frac{1}{5}$.

Case 1: Construct algorithm $\mathcal{B}$ with input $S = \{S_i\}_{i=1}^k$ (where $S_i \in \mathcal{Z}^N$) as follows:

Step 1. Run $\mathcal{A}$ on S_i, $i = 1, \cdots, k$. Denote the output as $h_i = \mathcal{A}(S_i)$.

Step 2. Let utility function as $u_{h_{[k]}}(S, i) = N\left(\hat{\mathcal{R}}_{S_i}(h_i) - \mathcal{R}_{\mathcal{D}}(h_i)\right)$. Apply the utility $u_{h_{[k]}}$ to an 0.7ε-differential private exponential mechanism $\mathcal{E}(S, u_{h_{[k]}}, [k], \varepsilon)$ and return the output.

By the classical composition property of differential privacy, one can easily obtain that $\mathcal{B}$ is $(1.7\varepsilon, \delta)$ differentially private. We then prove that $\mathcal{B}$ satisfies

$$\mathbb{P}\left[\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right) \leq \hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right) + ke^{-2\varepsilon}\delta + 6\varepsilon\right] < \varepsilon.$$

Since eq.(4) holds, we have that

$$\mathbb{P}\left(\forall i, \mathcal{R}_{\mathcal{D}}(\mathcal{A}(S_i)) \leq 4\varepsilon + \hat{\mathcal{R}}_{S_i}(\mathcal{A}(S_i))\right) \leq \left(1 - \frac{e^{-1.7\varepsilon}\delta}{\varepsilon}\ln\left(\frac{2}{\varepsilon}\right)\right)^k,$$

which leads to

$$\mathbb{P}\left(\exists i, \mathcal{R}_{\mathcal{D}}(\mathcal{A}(S_i)) > 4\varepsilon + \hat{\mathcal{R}}_{S_i}(\mathcal{A}(S_i))\right) > 1 - \left(1 - \frac{1}{k}\ln\left(\frac{2}{\varepsilon}\right)\right)^k \overset{(*)}{\geq} 1 - \frac{\varepsilon}{2}, \tag{3}$$

where eq.$(*)$ is due to $1 + x \leq e^x$.

On the other hand, the sensitivity of $u_{h_{[k]}}$ for any fixed $h_{[k]}$ can then be bounded as

$$\begin{aligned}
&\max_i \max_{S \text{ and } S' \text{ adjacent}} \left|N\left(\hat{\mathcal{R}}_{S_i}(h_i) - \mathcal{R}_{\mathcal{D}}(h_i)\right) - N\left(\hat{\mathcal{R}}_{S'_i}(h_i) - \mathcal{R}_{\mathcal{D}}(h_i)\right)\right| \\
=&\max_i \max_{S \text{ and } S' \text{ adjacent}} \left|N\left(\hat{\mathcal{R}}_{S_i}(h_i) - \hat{\mathcal{R}}_{S'_i}(h_i)\right)\right| \\
=&\max_i \max_{S \text{ and } S' \text{ adjacent}} \left|N\left(\hat{\mathcal{R}}_{S_i}(h_i) - \hat{\mathcal{R}}_{S'_i}(h_i)\right)\right| \leq 1.
\end{aligned}$$

Therefore, due to Lemma 1, for any fixed $h_{[k]}$ and any $\alpha \in \mathbb{R}$, we have

$$\mathbb{P}\left[u_{h_{[k]}}\left(S, \mathcal{E}(S, u_{h_{[k]}}, \mathcal{I}, \varepsilon)\right) \leq \max_{i \in \mathcal{I}} u_{h_{[k]}}(S, i) - \alpha\right] \leq |\mathcal{I}|e^{-\frac{0.7\alpha\varepsilon}{2}}.$$

We further set $\alpha = 0.11N\varepsilon$ and have

$$\mathbb{P}\left[\mathrm{Gen}_{S_{\mathcal{E}\left(S, u_{h_{[k]}}, \mathcal{I}, \varepsilon\right)}, h_{\mathcal{E}\left(S, u_{h_{[k]}}, \mathcal{I}, \varepsilon\right)}} \leq \max_{i \in \mathcal{I}} \mathrm{Gen}_{S_i, h_i} - \varepsilon\right] \leq |\mathcal{I}|e^{-\frac{0.077N\varepsilon^2}{2}} \leq \varepsilon\left(\frac{85}{127} - \frac{1}{2}\right). \tag{4}$$

Combing eq.(3) and eq.(4), we have

$$\mathbb{P}\left(\mathrm{Gen}_{S_{i_{\mathcal{B}(S)}}, h_{\mathcal{B}(S)}} > 1.7 \times 1.7\varepsilon + \varepsilon\right) > 1 - \frac{85}{127}\varepsilon,$$

which completes the proof by $\varepsilon = ke^{-1.7\varepsilon}\delta$.

Case 2: We construct $\mathcal{B}$ the same as **Case 1** except setting the privacy parameter as $-\ln(0.9)$. Similar to **Case 1**, we have

$$\mathbb{P}\left(\exists i, \mathcal{R}_{\mathcal{D}}(\mathcal{A}(S_i)) > 1.2(1 - 0.9e^{-\varepsilon}) + \hat{\mathcal{R}}_{S_i}(\mathcal{A}(S_i))\right) > 1 - \left(1 - \frac{1}{k}\ln\left(\frac{220}{1 - 0.9e^{-\varepsilon}}\right)\right)^k \overset{(*)}{\geq} 1 - \frac{1 - 0.9e^{-\varepsilon}}{220}, \tag{5}$$

and

$$\mathbb{P}\left[\mathrm{Gen}_{S_{\mathcal{E}\left(S, u_{h_{[k]}}, \mathcal{I}, \varepsilon\right)}, h_{\mathcal{E}\left(S, u_{h_{[k]}}, \mathcal{I}, \varepsilon\right)}} \leq \max_{i \in \mathcal{I}} \mathrm{Gen}_{S_i, h_i} - 0.01(1 - 0.9e^{-\varepsilon})\right] \leq |\mathcal{I}|e^{0.01\frac{N\ln(0.9)(1 - 0.9e^{-\varepsilon})}{2}} \leq \frac{1 - 0.9e^{-\varepsilon}}{220 \times 219}. \tag{6}$$

Combining eq.(5) and eq.(6) completes the proof of **Case 2**.

□

A.2 PROOFS OF THEOREM 1

We use Theorem 3 and Lemma 1 to derive the proof of Theorem 1.

Proof of Theorem 1. We prove Theorem 1 by reduction to absurdity. We only prove $\mathbb{P}\left(\mathrm{Gen}_{S,\mathcal{A}(S)} > a\right) < \frac{b}{2}$, while $\mathbb{P}\left(\mathrm{Gen}_{S,\mathcal{A}(S)} < -a\right) < \frac{b}{2}$, can be proven following the same routine by reversing the order of $\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right)$ and $\hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right)$ in Lemma 1.

Suppose there exists an algorithm $\mathcal{A}$ satisfies (ε, δ) differential privacy such that either the bound for $\varepsilon \leq \frac{1}{5}$ or $\varepsilon > \frac{1}{5}$ in Theorem 1 is false.

Then, by Theorem 3, for a multi-sample-set $(1.7\varepsilon, \delta)$-DP preserving algorithm, if $1.7\varepsilon \leq \frac{17}{50}$, by letting $k = \frac{1.7\varepsilon}{1.7e^{-1.7\varepsilon}\delta} = \frac{\varepsilon}{e^{-1.7\varepsilon}\delta}$, we have

$$\mathbb{P}\left(\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right) \leq \hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right) + \varepsilon + 1.7 \times 1.7\varepsilon\right) \geq \frac{85\varepsilon}{127}, \tag{7}$$

which contradicts the conclusion of Lemma 1 for $\varepsilon \leq \frac{1}{5}$. On the other hand, if $\varepsilon > \frac{1}{5}$, we have

$$\mathbb{P}\left(\mathcal{R}_{\mathcal{D}}\left(h_{\mathcal{B}(S)}\right) \leq \hat{\mathcal{R}}_{S_{i_{\mathcal{B}(S)}}}\left(h_{\mathcal{B}(S)}\right) + 1.2(1 - 0.9e^{-\varepsilon})\right) \geq \frac{1 - 0.9e^{-\varepsilon}}{219}, \tag{8}$$

which contradicts the conclusion of Lemma 1 for $\varepsilon > \frac{1}{5}$.

The proof is completed. □

B PROOFS OF COMPOSITION THEOREMS

This section proves the composition theorems. It is organized as follows: Section B.1 shows several definitions and lemmas which will be used throughout the proof; Section B.2 proves a preparation lemma on the KL divergence $D_{KL}(\mathcal{A}(S)\|\mathcal{A}(S'))$ between the hypotheses $\mathcal{A}(S)$ and $\mathcal{A}(S')$; based on this lemma Section B.3 proves a composition theorem of ε-differential privacy; Section B.4 extends the composition theorem to (ε, δ)-differential privacy; Section B.5 further tightens the estimate of δ' under some assumptions; and Section B.6 analyses the tightness of this estimation.

B.1 PRELIMINARIES

In this section, we define several measures between random variables that is omitted in Lemma 3, and further presents several lemmas that will be used in the proof of composition theorems.

Definition 1 (Max Divergence; cf. [Dwork and Roth, 2014], Definition 3.6)**.** *For any random variables X and Y, the max divergence between X and Y is defined as*

$$D_{\infty}(X\|Y) = \max_{U \subseteq \mathrm{Supp}(X)} \left[\log \frac{\mathbb{P}(X \in U)}{\mathbb{P}(Y \in U)}\right].$$

Definition 2 (δ-Approximate Max Divergence; cf. [Dwork and Roth, 2014], Definition 3.6)**.** *For any random variables X and Y, the δ-approximate max divergence between X to Y is defined as*

$$D_{\infty}^{\delta}(X\|Y) = \max_{U \subseteq \mathrm{Supp}(X): \mathbb{P}(Y \in U) \geq \delta} \left[\log \frac{\mathbb{P}(X \in U) - \delta}{\mathbb{P}(Y \in U)}\right].$$

Definition 3 (Statistical Distance; cf. [Dwork and Roth, 2014])**.** *For any random variables X and Y, the statistical distance between X and Y is defined as*

$$\Delta(X\|Y) = \max_{U} |\mathbb{P}(X \in U) - \mathbb{P}(Y \in U)|.$$

We then recall the following lemma which shows for any two distributions, there exist another two distributions with the same max-divergence and order-invariant KL divergence.

Lemma 2 (cf. [Dwork and Rothblum, 2016], Lemmas 3.9 and 3.10)**.** *For any two distributions P and P', there exist distributions Q and Q' such that*

$$\max\{D_\infty(Q\|Q'), D_\infty(Q'\|Q)\} = \max\{D_\infty(P\|P'), D_\infty(P'\|P)\},$$

and

$$D_{KL}(P\|P') \le D_{KL}(Q\|Q') = D_{KL}(Q'\|Q).$$

Azuma Lemma on concentration inequality of martingales will also be used in the proof of composition theorems.

Lemma 3 (Azuma Lemma; cf. [Mohri et al., 2018], p. 371)**.** *Suppose $\{Y_i\}_{i=1}^T$ is a sequence of random variables, where $Y_i \in [-a_i, a_i]$. Let $\{X_i\}_{i=1}^T$ be a sequence of random variables such that,*

$$\mathbb{E}(Y_i|X_{i-1}, ..., X_1) \le C_i,$$

where $\{C_i\}_{i=1}^T$ is a sequence of constant real numbers. Then, we have the following inequality,

$$\mathbb{P}\left(\sum_{i=1}^T Y_i \ge \sum_{i=1}^T C_i + t\sqrt{\sum_{i=1}^T a_i^2}\right) \le e^{-\frac{t^2}{2}}.$$

When proving the composition bound of Theorem 4 and Theorem 6, we will need to calculate the form of the (maximum) optimal point of function $f(\{\alpha_i\}_{i=1}^T) \stackrel{\triangle}{=} 1 - \prod_{i=1}^T(1-\alpha_i A_i)$ under some constraints. We put the form into the following lemma.

Lemma 4. *The maximum of function $f(\{\alpha_i\}_{i=1}^T) = 1 - \prod_{i=1}^T(1-\alpha_i A_i)$ over $\Omega\{1 \le \alpha_i \le c_i,$ and $\prod_{i=1}^T \alpha_i = c\}$ (c_i, A_i is fixed positive real and $c_i A_i \le 1$) is achieved at $\partial\Omega$.*

The proof can be derived by simple reduction to absurdity, and we omit it here.

B.2 PROOF OF LEMMA 2

Proof of Lemma 2. By Lemma 2, we have a random variable $\mathcal{M}(S)$ and $\mathcal{M}(S')$, which satisfies

$$D_\infty(\mathcal{M}(S)\|\mathcal{M}(S')) \le \varepsilon,\ D_\infty(\mathcal{M}(S')\|\mathcal{M}(S)) \le \varepsilon,$$

and

$$D_{KL}(\mathcal{A}(S)\|\mathcal{A}(S')) \le D_{KL}(\mathcal{M}(S)\|\mathcal{M}(S')) = D_{KL}(\mathcal{M}(S')\|\mathcal{M}(S)). \tag{9}$$

Therefore, we only need to derive a bound for $D_{KL}(\mathcal{M}(S)\|\mathcal{M}(S'))$.

By direct calculation,

$$\begin{aligned}
&D_{KL}(\mathcal{M}(S)\|\mathcal{M}(S')) \\
\stackrel{(*)}{=}&\frac{1}{2}\left[D_{KL}(\mathcal{M}(S)\|\mathcal{M}(S')) + D_{KL}(\mathcal{M}(S')\|\mathcal{M}(S))\right] \\
=&\frac{1}{2}\int \log\frac{\mathrm{d}\mathbb{P}(\mathcal{M}(S))}{\mathrm{d}\mathbb{P}(\mathcal{M}(S'))}\mathrm{d}\mathbb{P}(\mathcal{M}(S)) + \frac{1}{2}\int \log\frac{\mathrm{d}\mathbb{P}(\mathcal{M}(S'))}{\mathrm{d}\mathbb{P}(\mathcal{M}(S))}\mathrm{d}\mathbb{P}(\mathcal{M}(S')) \\
=&\frac{1}{2}\int \log\frac{\mathrm{d}\mathbb{P}(\mathcal{M}(S))}{\mathrm{d}\mathbb{P}(\mathcal{M}(S'))}\mathrm{d}\left[\mathbb{P}(\mathcal{M}(S)) - \mathbb{P}(\mathcal{M}(S'))\right] \\
&+\frac{1}{2}\int\left(\log\frac{\mathrm{d}\mathbb{P}(\mathcal{M}(S'))}{\mathrm{d}\mathbb{P}(\mathcal{M}(S))} + \log\frac{\mathrm{d}\mathbb{P}(\mathcal{M}(S))}{\mathrm{d}\mathbb{P}(\mathcal{M}(S'))}\right)\mathrm{d}\mathbb{P}(\mathcal{M}(S')) \\
=&\frac{1}{2}\int \log\frac{\mathrm{d}\mathbb{P}(\mathcal{M}(S))}{\mathrm{d}\mathbb{P}(\mathcal{M}(S'))}\mathrm{d}\left[\mathbb{P}(\mathcal{M}(S)) - \mathbb{P}(\mathcal{M}(S'))\right] + \frac{1}{2}\int \log 1\ \mathrm{d}\mathbb{P}(\mathcal{M}(S')) \\
=&\frac{1}{2}\int \log\frac{\mathrm{d}\mathbb{P}(\mathcal{M}(S))}{\mathrm{d}\mathbb{P}(\mathcal{M}(S'))}\mathrm{d}\left[\mathbb{P}(\mathcal{M}(S)) - \mathbb{P}(\mathcal{M}(S'))\right],
\end{aligned} \tag{10}$$

where eq. $(*)$ comes from eq. (9).

We now analyse the last integration in eq. (10). Define

$$k(y) \triangleq \frac{\mathrm{d}\mathbb{P}(\mathcal{M}(S)=y)}{\mathrm{d}\mathbb{P}(\mathcal{M}(S')=y)} - 1. \tag{11}$$

Therefore,

$$k(y)\mathrm{d}\mathbb{P}(\mathcal{M}(S')=y) = \mathrm{d}\mathbb{P}(\mathcal{M}(S)=y) - \mathrm{d}\mathbb{P}(\mathcal{M}(S')=y). \tag{12}$$

Additionally,

$$\begin{aligned}\mathbb{E}_{\mathcal{M}(S')}k(\mathcal{M}(S')) =& \int_{y\in\mathcal{H}} k(y)\mathrm{d}\mathbb{P}(\mathcal{M}(S')=y)\\ =& \int_{y\in\mathcal{H}} \mathrm{d}\left(\mathbb{P}(\mathcal{M}(S)=y) - \mathrm{d}\mathbb{P}(\mathcal{M}(S')=y)\right)\\ =& 0.\end{aligned} \tag{13}$$

By calculating the integration of the both sides of eq. (12), we have

$$\int k(y)\mathrm{d}\mathbb{P}(\mathcal{M}(S')=y) = 0.$$

Also, combined with the definition of $k(y)$ (see eq. 11), the right-hand side (RHS) of eq. (10) becomes

$$\text{RHS} = \mathbb{E}_{\mathcal{M}(S')}k(\mathcal{M}(S'))\log(k(\mathcal{M}(S'))+1). \tag{14}$$

Since M is ε-differentially private, $k(y)$ is bounded from both sides as follows,

$$e^{-\varepsilon} - 1 \le k(y) \le e^{\varepsilon} - 1. \tag{15}$$

We now calculate the maximum of eq. (14) subject to eqs. (13) and (15).

First, we argue that the maximum is achieved when $k(\mathcal{M}(S')) \in \{e^{-\varepsilon}-1, e^{\varepsilon}-1\}$ with probability 1 (almost surely). When $k(\mathcal{M}(S')) \in \{e^{-\varepsilon}-1, e^{\varepsilon}-1\}$, almost surely, the distribution for $k(\mathcal{M}(S'))$ is as following,

$$\begin{aligned}\mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon}-1) &= \frac{1}{1+e^{\varepsilon}},\\ \mathbb{P}^*(k(\mathcal{M}(S')) = e^{-\varepsilon}-1) &= \frac{e^{\varepsilon}}{1+e^{\varepsilon}}.\end{aligned}$$

We argue that it is the distribution that maximizes $k(\mathcal{M}(S'))$.

For the brevity, we denote the probability measure for a given distribution Q as $\mathbb{P}_Q$. Similarly, $\mathbb{P}^*$ corresponds the distribution Q^*. We prove that Q^* maximizes eq. (14) in the following two cases: (1) $\mathbb{P}_Q(k(\mathcal{M}(S')) \ge 0) \le \mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon}-1)$, and (2) $\mathbb{P}_Q(k(\mathcal{M}(S')) \ge 0) > \mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon}-1)$

Case 1: $\mathbb{P}_Q(k(\mathcal{M}(S')) \ge 0) \le \mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon}-1)$

We have

$$\begin{aligned}&\mathbb{E}_{\mathcal{M}(S')\sim Q^*}(k(\mathcal{M}(S')\log(k(\mathcal{M}(S'))+1))\\ =&\mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon}-1)\cdot\varepsilon(e^{\varepsilon}-1) - \mathbb{P}^*(k(\mathcal{M}(S')) = e^{-\varepsilon}-1)\cdot\varepsilon(e^{-\varepsilon}-1)\\ =&(\mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon}-1) - \mathbb{P}_Q(k(\mathcal{M}(S')) \ge 0))\cdot\varepsilon(e^{\varepsilon}-1)\\ &+ \mathbb{P}_Q(k(\mathcal{M}(S')) \ge 0)\cdot\varepsilon(e^{\varepsilon}-1) - \mathbb{P}^*(k(\mathcal{M}(S')) = e^{-\varepsilon}-1)\cdot\varepsilon(e^{-\varepsilon}-1)\\ \ge&\mathbb{P}_Q(k(\mathcal{M}(S')) \ge 0)\cdot\varepsilon(1-e^{-\varepsilon}) - \mathbb{P}^*(k(\mathcal{M}(S')) = e^{-\varepsilon}-1)\cdot\varepsilon(1-e^{-\varepsilon})\\ &+ \mathbb{P}_Q(k(\mathcal{M}(S')) \ge 0)\cdot\varepsilon(e^{\varepsilon}-1) - \mathbb{P}^*(k(\mathcal{M}(S')) = e^{-\varepsilon}-1)\cdot\varepsilon(e^{-\varepsilon}-1).\end{aligned}$$

Note that

$$\begin{aligned}\mathbb{P}_Q(k(\mathcal{M}(S')) < 0) =& \mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon} - 1) - \mathbb{P}_Q(k(\mathcal{M}(S')) \geq 0) \\ &+ \mathbb{P}^*(k(\mathcal{M}(S')) = e^{-\varepsilon} - 1).\end{aligned}$$

Therefore, together with the condition eq. (15),

$$\begin{aligned}&\mathbb{E}_{\mathcal{M}(S')\sim Q}(k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)I_{k(\mathcal{M}(S')\leq 0)}) \\ \leq&(\mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon} - 1) - \mathbb{P}_Q(k(\mathcal{M}(S')) \geq 0)) \cdot \varepsilon(1 - e^{-\varepsilon}) \\ &+ \mathbb{P}^*(k(\mathcal{M}(S')) = e^{-\varepsilon} - 1) \cdot \varepsilon(1 - e^{-\varepsilon}).\end{aligned} \tag{16}$$

Also,

$$\mathbb{E}_{\mathcal{M}(S')\sim Q}(k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)I_{k(\mathcal{M}(S'))>0}) \leq \mathbb{P}_Q(k(\mathcal{M}(S')) \geq 0) \cdot \varepsilon(e^{\varepsilon} - 1). \tag{17}$$

Therefore, combined inequalities eqs. (16) and (17), we have

$$\mathbb{E}_{\mathcal{M}(S')\sim Q}(k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)) \leq \mathbb{E}_{\mathcal{M}(S')\sim Q^*}(k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)).$$

Since the distribution Q is arbitrary, the distribution Q^* maximizes the $k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)$.

Case 2: $\mathbb{P}_Q(k(\mathcal{M}(S')) \geq 0) > \mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon} - 1)$

We first prove that if $\mathbb{P}_Q(1 - e^{-\varepsilon} < k(\mathcal{M}(S')) < 0) \neq 0$, there exists a distribution Q' such that

$$\begin{aligned}\mathbb{P}_{Q'}(k(\mathcal{M}(S')) \geq 0) &= \mathbb{P}_Q(k(\mathcal{M}(S')) \geq 0), \\ \mathbb{P}_{Q'}(k(\mathcal{M}(S')) < 0) &= \mathbb{P}_Q(k(\mathcal{M}(S')) < 0), \\ \mathbb{P}_{Q'}(k(\mathcal{M}(S')) < 0) &= \mathbb{P}_{Q'}(k(\mathcal{M}(S') = e^{-\varepsilon} - 1), \\ \mathbb{E}_{Q'}(k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)) &> \mathbb{E}_{Q'}(k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)),\end{aligned}$$

while the two conditions (eqs. 13, 15) still hold.

Additionally, we have assumed that

$$\mathbb{P}_Q(k(\mathcal{M}(S')) \geq 0) > \mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon} - 1).$$

Therefore,

$$\mathbb{P}_Q(k(\mathcal{M}(S')) \leq 0) < \mathbb{P}^*(k(\mathcal{M}(S')) = e^{-\varepsilon} - 1).$$

Also, since the distribution Q' is arbitrary, let it satisfy

$$\mathbb{P}_{Q'}(k(\mathcal{M}(S')) < 0) = \mathbb{P}_Q(k(\mathcal{M}(S')) < 0) = \mathbb{P}_{Q'}(k(\mathcal{M}(S') = e^{-\varepsilon} - 1).$$

Then, in order to meet the condition eq. (13), let

$$\mathbb{P}_{Q'}(k(\mathcal{M}(S') = e^{\varepsilon} - 1) > \mathbb{P}_Q(k(\mathcal{M}(S') = e^{\varepsilon} - 1),$$

and

$$\mathbb{P}_{Q'}(0 < k(\mathcal{M}(S')) < e^{\varepsilon} - 1) \leq \mathbb{P}_Q(0 < k(\mathcal{M}(S')) < e^{\varepsilon} - 1),$$

Since $x\log(x + 1)$ increases when $x > 0$ and decreases when $x < 0$, we have

$$\mathbb{E}_{Q'}(k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)) > \mathbb{E}_Q(k(\mathcal{M}(S')\log(k(\mathcal{M}(S')) + 1)).$$

Therefore, we have proved that the argument when $\mathbb{P}_Q(k(\mathcal{M}(S')) < 0) \neq \mathbb{P}_Q(k(\mathcal{M}(S')) = e^{-\varepsilon} - 1)$. We now prove the case that

$$\mathbb{P}_Q(k(\mathcal{M}(S')) < 0) = \mathbb{P}_Q(k(\mathcal{M}(S')) = e^{-\varepsilon} - 1),$$

where

$$\mathbb{E}_Q(k(\mathcal{M}(S')\log(k(\mathcal{M}(S'))+1)I_{k(\mathcal{M}(S'))<0}) = \varepsilon(1-e^{-\varepsilon})\mathbb{P}_Q(k(\mathcal{M}(S'))<0).$$

Applying Jensen's inequality to bound the $\mathbb{E}_Q(k(\mathcal{M}(S')\log(k(\mathcal{M}(S'))+1)I_{k(\mathcal{M}(S'))\geq 0})$, we have

$$\begin{aligned}
&\mathbb{E}_Q(k(\mathcal{M}(S'))\log(k(\mathcal{M}(S'))+1)I_{k(\mathcal{M}(S'))\geq 0}) \\
=&\mathbb{P}_Q(\mathcal{M}(S')\geq 0)\mathbb{E}_{Q'}(k(\mathcal{M}(S'))\log(k(\mathcal{M}(S'))+1)|k(\mathcal{M}(S'))\geq 0) \\
\overset{(*)}{\leq}&\mathbb{P}_Q(\mathcal{M}(S')\geq 0)\mathbb{E}_Q\left(k(\mathcal{M}(S'))|k(\mathcal{M}(S'))\geq 0\right)\cdot\log(\mathbb{E}_Q\left(k(\mathcal{M}(S'))|k(\mathcal{M}(S'))\geq 0\right)+1),
\end{aligned} \tag{18}$$

where the inequality $(*)$ uses Jensen's inequality ($x\log(1+x)$ is convex with respect to x when $x>0$). The upper bound in eq. (18) is achieved as long as

$$\mathbb{P}_Q(k(\mathcal{M}(S'))\geq 0) = \mathbb{P}_Q(k(\mathcal{M}(S')) = \mathbb{E}_Q(k(\mathcal{M}(S'))|k(\mathcal{M}(S'))\geq 0)).$$

Furthermore,

$$\mathbb{P}_Q(k(\mathcal{M}(S'))<0) = \mathbb{P}_Q(k(\mathcal{M}(S')) = e^{-\varepsilon}-1).$$

Therefore, the distribution Q is determined by the cumulative density functions $\mathbb{P}_Q(k(\mathcal{M}(S'))<0)$ and $\mathbb{P}_Q(k(\mathcal{M}(S'))\geq 0)$.

Hence, maximizing $\mathbb{E}_Q(k(\mathcal{M}(S')\log(k(\mathcal{M}(S'))+1))$ is equivalent to maximizing the following object function,

$$g(q) = q(1-e^{-\varepsilon})\log e^{\varepsilon} + (1-q)\frac{q}{1-q}(1-e^{-\varepsilon})\log\left(\frac{q}{1-q}(1-e^{-\varepsilon})+1\right),$$

subject to

$$\frac{q}{1-q}\leq e^{\varepsilon}, \tag{19}$$

where $g(q)$ is the maximum of eq. (14) subject to $\mathbb{P}_Q(k(\mathcal{M}(S'))<0) = q$, and the condition eq. (19) comes from the $\mathbb{P}_Q(k(\mathcal{M}(S'))\geq 0) > \mathbb{P}^*(k(\mathcal{M}(S')) = e^{\varepsilon}-1)$ (the assumption of Case 2).

Additionally, $g(q)$ can be represented as follows,

$$q(1-e^{-\varepsilon})\log\left(\frac{q}{1-q}(e^{\varepsilon}-1)+\varepsilon\right).$$

Since both q and $\frac{q}{1-q}$ monotonously increase, $g(q)$ monotonously increases. Therefore, Q^* maximize eq. (14), which finishes the proof. □

B.3 PROOF OF THEOREM 5

Based on Lemma 2, we can prove the following composition theorem for ε-differential privacy as a preparation theorem of the general case.

Proof of Theorem 5. We begin by calculating $\log \frac{\mathbb{P}^S\left(\{W_i\}_{i=0}^T\right)}{\mathbb{P}^{S'}\left(\{W_i\}_{i=0}^T\right)}$ as follows,

$$\begin{aligned}
\log \frac{\mathbb{P}^S\left(\{W_i\}_{i=0}^T\right)}{\mathbb{P}^{S'}\left(\{W_i\}_{i=0}^T\right)} &= \log\left(\prod_{i=0}^T \frac{\mathbb{P}^S\left(W_i|W_{i-1},...,W_0\right)}{\mathbb{P}^{S'}\left(W_i|W_{i-1},...,W_0\right)}\right) \\
&= \sum_{i=0}^T \log\left(\frac{\mathbb{P}^S\left(W_i|W_{i-1},...,W_0\right)}{\mathbb{P}^{S'}\left(W_i|W_{i-1},...,W_0\right)}\right) \\
&\overset{(*)}{=} \sum_{i=1}^T \log\left(\frac{\mathbb{P}^S\left(W_i|W_{i-1},...,W_0\right)}{\mathbb{P}^{S'}\left(W_i|W_{i-1},...,W_0\right)}\right) \\
&= \sum_{i=1}^T \log\left(\frac{\mathbb{P}^S\left(\mathcal{M}_i(W_{i-1},S)=W_i|W_{i-1},...,W_0\right)}{\mathbb{P}^{S'}\left(\mathcal{M}_i(W_{i-1},S')=W_i|W_{i-1},...,W_0\right)}\right) \\
&\overset{(**)}{=} \sum_{i=1}^T \log\left(\frac{\mathbb{P}^{S,W_{i-1}}\left(\mathcal{M}_i(W_{i-1},S)=W_i\right)}{\mathbb{P}^{S',W_{i-1}}\left(\mathcal{M}_i(W_{i-1},S')=W_i\right)}\right),
\end{aligned}$$

where eq. $(*)$ comes from the independence of W_0 with respect to S and eq. $(**)$ is because the independence of $\mathcal{M}_i$ to W_k $(k<i)$ when the W_{i-1} is fixed.

By the definition of ε-differential privacy, one has for arbitrary W_{i-1},

$$\begin{aligned}
D_\infty\left(\mathcal{M}_i(W_{i-1},S)\|\mathcal{M}_i(W_{i-1},S')\right) &< \varepsilon_i, \\
D_\infty\left(\mathcal{M}_i(W_{i-1},S')\|\mathcal{M}_i(W_{i-1},S)\right) &< \varepsilon_i.
\end{aligned}$$

Thus, by Lemma 2, we have that

$$\begin{aligned}
&\mathbb{E}^S\left(\log\left(\frac{\mathbb{P}\left(\mathcal{M}_i(W_{i-1},S)=W_i\right)}{\mathbb{P}\left(\mathcal{M}_i(W_{i-1},S')=W_i\right)}\right)|W_{i-1},\cdots,W_0\right) \\
=&D_{KL}(\mathcal{M}_i(W_{i-1},S)\|\mathcal{M}_i(W_{i-1},S')) \\
\leq&\varepsilon_i\frac{e^{\varepsilon_i}-1}{e^{\varepsilon_i}+1}.
\end{aligned} \tag{20}$$

Combining Azuma Lemma (Lemma 3) and eq. (20), we can finally derive the following equation

$$\mathbb{P}^S\left(\{W_i'\}_{i=0}^T : \frac{\mathbb{P}^S\left(W_i=W_i', i\in\{0,\cdots,T\}\right)}{\mathbb{P}\left(W_i=W_i', i\in\{0,\cdots,T\}\right)} > e^{\varepsilon'}\right) < \delta',$$

where S and S' are adjacent sample sets.

Therefore, the algorithm $\mathcal{A}$ is ε'-differentially private.

The proof is completed.

□

B.4 PROOF OF THEOREM 6

Now, we can prove our composition theorems for (ε,δ)-differential privacy. We first prove a composition algorithm of (ε,δ)-differential privacy whose estimate of ε' is somewhat looser than the existing results. Then, we tighten the results and obtain a composition theorem that strictly tighter than the current estimate.

Proof of Theorem 6. It has been proved that the optimal privacy preservation can be achieved by a sequence of independent iterations (see [Kairouz et al., 2017], Theorem 3.5). Therefore, without loss of generality, we assume that the iterations in our theorem are independent with each other.

Fixed any two adjacent sample sets S and S', and rewrite $W_i(S)$ as W_i^0, and $W_i(S')$ as W_i^1. Then, by Lemma 3, for $i \geq 1$ there exist random variables $\tilde{W}_i^0$ and $\tilde{W}_i^1$, such that

$$\Delta\left(W_i^0 \| \tilde{W}_i^0\right) \leq \frac{\delta_i}{1+e^{\varepsilon_i}}, \tag{21}$$

$$\Delta\left(W_i^1 \| \tilde{W}_i^1\right) \leq \frac{\delta_i}{1+e^{\varepsilon_i}}, \tag{22}$$

$$D_\infty\left(\tilde{W}_i^0 \| \tilde{W}_i^1\right) \leq \varepsilon_i, \tag{23}$$

$$D_\infty\left(\tilde{W}_i^1 \| \tilde{W}_i^0\right) \leq \varepsilon_i. \tag{24}$$

Applying Theorem 6 (here, $\delta = \tilde{\delta}$), we have that

$$D_\infty^{\tilde{\delta}}\left(\{\tilde{W}_i^0\}_{i=0}^T \| \{\tilde{W}_i^1\}_{i=0}^T\right) \leq \varepsilon',$$
$$D_\infty^{\tilde{\delta}}\left(\{\tilde{W}_i^1\}_{i=0}^T \| \{\tilde{W}_i^0\}_{i=0}^T\right) \leq \varepsilon'.$$

Apparently, for any sequence of hypothesis sets $\mathcal{H}_0, \cdots, \mathcal{H}_T$,

$$\mathbb{P}(W_i^0 \in \mathcal{H}_i) - \min\left\{\frac{\delta_i}{1+e^{\varepsilon_i}}, \mathbb{P}(W_i^0 \in \mathcal{H}_i)\right\} \geq 0.$$

Therefore,

$$\begin{aligned}
&\mathbb{P}(W_0^0 \in \mathcal{H}_0)\left(\mathbb{P}(W_1^0 \in \mathcal{H}_1) - \min\left\{\frac{\delta_1}{1+e^{\varepsilon_1}}, \mathbb{P}(W_1^0 \in \mathcal{H}_1)\right\}\right) \\
&\cdots\left(\mathbb{P}(W_T^0 \in \mathcal{H}_T) - \min\left\{\frac{\delta_T}{1+e^{\varepsilon_T}}, \mathbb{P}(W_T^0 \in \mathcal{H}_1)\right\}\right) \\
\leq &\mathbb{P}(\tilde{W}_0^0 \in \mathcal{H}_0)\cdots\mathbb{P}(\tilde{W}_T^0 \in \mathcal{H}_T) \\
\leq &e^{\varepsilon'}\mathbb{P}(\tilde{W}_0^1 \in \mathcal{H}_0)\cdots\mathbb{P}(\tilde{W}_T^1 \in \mathcal{H}_T) + \tilde{\delta}.
\end{aligned} \tag{25}$$

Furthermore, by eq. (24), we also have that

$$\mathbb{P}(\tilde{W}_i^0 \in \mathcal{H}_i) \leq \min\left\{e^{\varepsilon_i}, \frac{1}{\mathbb{P}(\tilde{W}_i^1 \in \mathcal{H}_i)}\right\}\mathbb{P}(\tilde{W}_i^1 \in \mathcal{H}_i).$$

Therefore,

$$\mathbb{P}(\tilde{W}_0^0 \in \mathcal{H}_0)\cdots\mathbb{P}(\tilde{W}_n^0 \in \mathcal{H}_T) \leq \prod_{i=1}^T \min\left\{e^{\varepsilon_i}, \frac{1}{\mathbb{P}(\tilde{W}_i^1 \in \mathcal{H}_i)}\right\}\mathbb{P}(\tilde{W}_0^1 \in \mathcal{H}_0)\cdots\mathbb{P}(\tilde{W}_T^1 \in \mathcal{H}_T) + \tilde{\delta}.$$

Then, we prove this theorem in two cases: (1) $\prod_{\mathbf{i=1}}^{\mathbf{T}} \min\left\{\mathbf{e}^{\varepsilon_\mathbf{i}}, \frac{\mathbf{1}}{\mathbb{P}(\tilde{\mathbf{W}}_\mathbf{i}^\mathbf{1} \in \mathcal{H}_\mathbf{i})}\right\} \leq \mathbf{e}^{\varepsilon'}$; and (2) $\prod_{\mathbf{i=1}}^{\mathbf{T}} \min\left\{\mathbf{e}^{\varepsilon_\mathbf{i}}, \frac{\mathbf{1}}{\mathbb{P}(\tilde{\mathbf{W}}_\mathbf{i}^\mathbf{1} \in \mathcal{H}_\mathbf{i})}\right\} > \mathbf{e}^{\varepsilon'}$.

Case 1- $\prod_{\mathbf{i=1}}^{\mathbf{T}} \min\left\{\mathbf{e}^{\varepsilon_\mathbf{i}}, \frac{\mathbf{1}}{\mathbb{P}(\tilde{\mathbf{W}}_\mathbf{i}^\mathbf{1} \in \mathcal{H}_\mathbf{i})}\right\} \leq \mathbf{e}^{\varepsilon'}$.

We have that

$$\begin{aligned}
&\mathbb{P}(\tilde{W}_0^1 \in \mathcal{H}_0)\left(\mathbb{P}(\tilde{W}_1^1 \in \mathcal{H}_1) - \frac{\delta_1}{1+e^{\varepsilon_1}}\right)\cdots\left(\mathbb{P}(\tilde{W}_T^1 \in \mathcal{H}_T) - \frac{\delta_T}{1+e^{\varepsilon_T}}\right) \\
\leq &\mathbb{P}(W_0^1 \in \mathcal{H}_0)\cdots\mathbb{P}(W_T^1 \in \mathcal{H}_T).
\end{aligned}$$

By simple calculation, we have that

$$\begin{aligned}
&\prod_{i=1}^{T}\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}\mathbb{P}(\tilde{W}_0^1\in\mathcal{H}_0)\cdots\mathbb{P}(\tilde{W}_T^1\in\mathcal{H}_T)\\
\leq&\prod_{i=1}^{T}\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}\mathbb{P}(W_0^1\in\mathcal{H}_0)\cdots\mathbb{P}(W_T^1\in\mathcal{H}_T)\\
&+\prod_{i=1}^{T}\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}\mathbb{P}(\tilde{W}_0^1\in\mathcal{H}_0)\cdots\mathbb{P}(\tilde{W}_T^1\in\mathcal{H}_T)\\
&-\prod_{i=1}^{n}\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}\mathbb{P}(\tilde{W}_0^1\in\mathcal{H}_0)\\
&\qquad\left(\mathbb{P}(\tilde{W}_1^1\in\mathcal{H}_1)-\frac{\delta_1}{1+e^{\varepsilon_1}}\right)\cdots\left(\mathbb{P}(\tilde{W}_T^1\in\mathcal{H}_T)-\frac{\delta_T}{1+e^{\varepsilon_T}}\right).
\end{aligned}$$

Apparently,

$$\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}\mathbb{P}(\tilde{W}_0^i\in\mathcal{H}_i)\leq 1,$$

and when $A>B$, $f(x)=Ax-(x-a)B$ increases when x increases.

Therefore, we have that

$$\begin{aligned}
&\prod_{i=1}^{T}\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}\mathbb{P}(\tilde{W}_0^1\in\mathcal{H}_0)\cdots\mathbb{P}(\tilde{W}_T^1\in\mathcal{H}_T)\\
&-\prod_{i=1}^{T}\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}P(\tilde{W}_0^1\in\mathcal{H}_0)\\
&\qquad\left(\mathbb{P}(\tilde{W}_1^1\in\mathcal{H}_1)-\frac{\delta_1}{1+e^{\varepsilon_1}}\right)\cdots\left(\mathbb{P}(\tilde{W}_T^1\in\mathcal{H}_T)-\frac{\delta_T}{1+e^{\varepsilon_T}}\right)\\
\leq&1-\prod_{i=1}^{T}\left(1-\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}\frac{\delta_i}{1+e^{\varepsilon_i}}\right).
\end{aligned}$$

Combining with eq. (25), we have that

$$\delta'\leq 1-\prod_{i=1}^{T}\left(1-\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\}\frac{\delta_i}{1+e^{\varepsilon_i}}\right)+1-\prod_{i=1}^{T}\left(1-\frac{\delta_i}{1+e^{\varepsilon_i}}\right)+\tilde{\delta}.$$

Case 2- $\prod_{\mathbf{i=1}}^{\mathbf{T}}\min\left\{\mathbf{e}^{\varepsilon_\mathbf{i}},\frac{\mathbf{1}}{\mathbb{P}(\tilde{\mathbf{W}}_\mathbf{i}^\mathbf{1}\in\mathcal{H}_\mathbf{i})}\right\}>\mathbf{e}^{\varepsilon'}$:

There exists a sequence of reals $\{\alpha_i\}_{i=1}^{T}$ such that

$$\begin{aligned}
e^{\alpha_i}&\leq\min\left\{e^{\varepsilon_i},\frac{1}{\mathbb{P}(\tilde{W}_i^1\in\mathcal{H}_i)}\right\},\\
\sum_{i=1}^{T}\alpha_i&=\varepsilon'.
\end{aligned}$$

Therefore, similar to Case 1, we have that

$$\delta'\leq 1-\prod_{i=1}^{T}\left(1-e^{\alpha_i}\frac{\delta_i}{1+e^{\varepsilon_i}}\right)+1-\prod_{i=1}^{T}\left(1-\frac{\delta_i}{1+e^{\varepsilon_i}}\right).$$

Overall, we have proven that

$$\delta' \leq 1 - \prod_{i=1}^{T} \left(1 - e^{\alpha_i} \frac{\delta_i}{1 + e^{\varepsilon_i}}\right) + 1 - \prod_{i=1}^{T} \left(1 - \frac{\delta_i}{1 + e^{\varepsilon_i}}\right),$$

where $\sum_{i=1}^{T} \alpha_i \leq \varepsilon'$ and $\alpha_i \leq \varepsilon_i$.

From Lemma 4, the minimum is realised on the boundary, which is exactly this theorem claims.

The proof is completed. □

Then, we can prove prove Theorem 4.

Proof of Theorem 4. Applying Theorem 3.5 in [Kairouz et al., 2017] and replacing ε' in the proof of Theorem 6 as

$$\varepsilon' = \min\left\{\varepsilon'_1, \varepsilon'_2, \varepsilon'_3\right\}.$$

The proof is completed. □

B.5 PROOF OF COROLLARY 2

Proof of Corollary 2. Let $\mathcal{P}_0$ and $\mathcal{P}_1$ be two distributions whose cumulative distribution functions P_0 and P_1 are respectively defined as following:

$$P_0(x) = \begin{cases} \delta\,, & x = 0 \\ \dfrac{(1-\delta)e^{\varepsilon}}{1+e^{\varepsilon}}\,, & x = 1 \\ \dfrac{1-\delta}{1+e^{\varepsilon}}\,, & x = 2 \\ 0\,, & x = 3 \end{cases},$$

and

$$P_1(x) = \begin{cases} 0\,, & x = 0 \\ \dfrac{(1-\delta)e^{\varepsilon}}{1+e^{\varepsilon}}\,, & x = 1 \\ \dfrac{1-\delta}{1+e^{\varepsilon}}\,, & x = 2 \\ \delta\,, & x = 3 \end{cases}.$$

By Theorem 3.4 of [Kairouz et al., 2017], the largest magnitude of the (ε', δ')-differential privacy can be calculated from the $\mathcal{P}_0^{\bigotimes T}$ and $\mathcal{P}_1^{\bigotimes T}$.

Construct $\tilde{\mathcal{P}}_0$ and $\tilde{\mathcal{P}}_1$, whose cumulative distribution functions are as follows,

$$\tilde{P}_0(x) = \begin{cases} \dfrac{e^{\varepsilon}\delta}{1+e^{\varepsilon}}\,, & x = 0 \\ \dfrac{(1-\delta)e^{\varepsilon}}{1+e^{\varepsilon}}\,, & x = 1 \\ \dfrac{1-\delta}{1+e^{\varepsilon}}\,, & x = 2 \\ \dfrac{\delta}{1+e^{\varepsilon}}\,, & x = 3 \end{cases},$$

and

$$\tilde{P}_1(x)\begin{cases}\dfrac{\delta}{1+e^{\varepsilon}}, & x=0\\ \dfrac{(1-\delta)e^{\varepsilon}}{1+e^{\varepsilon}}, & x=1\\ \dfrac{1-\delta}{1+e^{\varepsilon}}, & x=2\\ \dfrac{e^{\varepsilon}\delta}{1+e^{\varepsilon}}, & x=3\end{cases}.$$

One can easily verify that

$$\begin{aligned}\Delta(\mathcal{P}_0\|\tilde{\mathcal{P}}_0)&\le\frac{\delta}{1+e^{\varepsilon}},\\ \Delta(\mathcal{P}_1\|\tilde{\mathcal{P}}_1)&\le\frac{\delta}{1+e^{\varepsilon}},\\ D_{\infty}(\tilde{\mathcal{P}}_0\|\tilde{\mathcal{P}}_1)&\le\varepsilon,\\ D_{\infty}(\tilde{\mathcal{P}}_1\|\tilde{\mathcal{P}}_0)&\le\varepsilon_i.\end{aligned}$$

Let $V_i(x_i)=\log\left(\frac{\tilde{\mathcal{P}}_0(x_i)}{\tilde{\mathcal{P}}_1(x_i)}\right)$ and $S(x_1,\cdots,x_T)=\sum_{i=1}^{T}V_i(x_i)$.

We have that for any $t>0$,

$$\mathbb{P}_{\tilde{\mathcal{P}}_0^T}(\{x_i\}:S(\{x_i\})>\varepsilon')\le e^{-\varepsilon' t}\mathbb{E}_{\tilde{\mathcal{P}}_0^{\otimes T}}(e^{tS})=e^{-\varepsilon' t}\left(\frac{e^{t\varepsilon+\varepsilon}}{1+e^{\varepsilon}}+\frac{e^{-t\varepsilon}}{1+e^{\varepsilon}}\right)^{T}=e^{-\varepsilon' t-Tt\varepsilon}\left(\frac{e^{2t\varepsilon+\varepsilon}}{1+e^{\varepsilon}}+\frac{1}{1+e^{\varepsilon}}\right)^{T}. \tag{26}$$

By calculating the derivative,, we have that the minimum of the RHS of eq. (26) is achieved at

$$e^{2\varepsilon t}=e^{-\varepsilon}\frac{T\varepsilon+\varepsilon'}{T\varepsilon-\varepsilon'}. \tag{27}$$

Since $\varepsilon'\ge T\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1}$,

$$e^{-\varepsilon}\frac{T\varepsilon+\varepsilon'}{T\varepsilon-\varepsilon'}>1.$$

Therefore, by applying eq.(27) into the RHS of eq. (26), we have that

$$\mathbb{P}_{\tilde{\mathcal{P}}_0^T}(\{x_i\}:S(\{x_i\})>\varepsilon')\le e^{-\frac{\varepsilon'+T\varepsilon}{2}}\left(\frac{1}{1+e^{\varepsilon}}\left(\frac{2T\varepsilon}{T\varepsilon-\varepsilon'}\right)\right)^{T}\left(\frac{T\varepsilon+\varepsilon'}{T\varepsilon-\varepsilon'}\right)^{-\frac{\varepsilon'+T\varepsilon}{2\varepsilon}}. \tag{28}$$

Define RHS of eq. (28) as δ'. We have $(\tilde{\mathcal{P}}^b)^{\otimes T}$ $(b=0,1)$ have ε' δ'-approximate max divergence. Then, using similar analysis of the Proof of Theorem 6, we prove this theorem. □

B.6 TIGHTNESS OF THEOREM 2

This section analyses the tightness of Theorem 2. Specifically, we compare it with our Theorem 4.

In the proof of Theorem 4 (see Section B.3), ε_3' is derived through Azuma Lemma (Lemma 3). Specifically, the δ' is derived by

$$\begin{aligned}\mathbb{P}\left[S_T\ge\varepsilon'-T\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1}\right]&\le e^{-t(\varepsilon'-T\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1})}\mathbb{E}\left[e^{tS_T}\right]\\ &=e^{-t(\varepsilon'-T\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1})}\mathbb{E}_{\tilde{\mathcal{P}}_0^{\otimes(T-1)}}\left[e^{tS_{T-1}}\mathbb{E}\left[e^{t\mathbf{V}_T}|x_1,\ldots,x_{T-1}\right]\right]\\ &\le e^{-t(\varepsilon'-T\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1})}\mathbb{E}_{\tilde{\mathcal{P}}_0^{\otimes(T-1)}}\left[e^{tS_{T-1}}\right]e^{4t^2\varepsilon^2/8}\\ &\le e^{-t(\varepsilon'-T\frac{e^{\varepsilon}-1}{e^{\varepsilon}+1})}e^{Tt^2\varepsilon^2/2},\end{aligned}$$

where V_i is defined as $\log \frac{\mathbb{P}_{\tilde{\mathcal{P}}_0}(x_i)}{\mathbb{P}_{\tilde{\mathcal{P}}_1}(x_i)} - \mathbb{E}_{\tilde{\mathcal{P}}_0}\left[\log \frac{\mathbb{P}_{\tilde{\mathcal{P}}_0}(x_i)}{\mathbb{P}_{\tilde{\mathcal{P}}_1}(x_i)} | x_1, \cdots, x_{i-1}\right]$ and S_j is defined as $\sum_{i=1}^{j} V_i$.

Since $\mathbb{P}\left[S_T \geq \varepsilon' - T\frac{e^\varepsilon - 1}{e^\varepsilon + 1}\right]$ does not depend on t,

$$\mathbb{P}\left[S_T \geq \varepsilon' - T\frac{e^\varepsilon - 1}{e^\varepsilon + 1}\right] \leq \min_{t>0} e^{-\frac{(\varepsilon' - T\frac{e^\varepsilon - 1}{e^\varepsilon + 1})^2}{2T\varepsilon^2}} = \delta',$$

By contrary, the approach here directly calculates $\mathbb{E}[e^{tS_T}]$, without the shrinkage in the proof of Theorem 4 (see Section B.3). Specifically,

$$e^{-\varepsilon' t - Tt\varepsilon}\left(\frac{e^{2t\varepsilon+\varepsilon}}{1+e^\varepsilon} + \frac{1}{1+e^\varepsilon}\right)^T = e^{-t\varepsilon}\mathbb{E}\left[e^{tS_T}\right] \leq e^{-t(\varepsilon' - T\frac{e^\varepsilon - 1}{e^\varepsilon + 1})} e^{Tt^2\varepsilon^2/2}.$$

Therefore,

$$\min_{t>0} e^{-\varepsilon' t - Tt\varepsilon}\left(\frac{e^{2t\varepsilon+\varepsilon}}{1+e^\varepsilon} + \frac{1}{1+e^\varepsilon}\right)^T \leq \min_{t>0} e^{-t(\varepsilon' - T\frac{e^\varepsilon - 1}{e^\varepsilon + 1})} e^{Tt^2\varepsilon^2/2},$$

which leads to

$$e^{-\frac{\varepsilon' + T\varepsilon}{2}}\left(\frac{1}{1+e^\varepsilon}\left(\frac{2T\varepsilon}{T\varepsilon - \varepsilon'}\right)\right)^T \left(\frac{T\varepsilon + \varepsilon'}{T\varepsilon - \varepsilon'}\right)^{-\frac{\varepsilon' + T\varepsilon}{2\varepsilon}} \leq \delta'.$$

It ensures that this estimate further tightens δ' than Section B.3 (which is also the $\tilde{\delta}$ in Theorem 4) if the ε' is the same.

C SUPPLEMENTARY MATERIALS OF THE APPLICATIONS

This appendix collects the formal description of SGLD and agnostic federated learning, together with the proof for the application in IGMM and the asymptotic generalization bound for agnostic federated learning.

C.1 DETAILED DESCRIPTION OF SGLD AND AGNOSTIC FEDERATED LEARNING

SGLD and agnostic federated learning are described respectively as the following two charts.

Algorithm 1 Stochastic Gradient Langevin Dynamics (SGLD)

Require: Sample $S = \{z_1, ...z_N\}$, Gaussian noise variance σ, size of mini-batch τ, iteration steps T, learning rate $\{\eta_1, ...\eta_T\}$, loss function $\ell(z, W)$, and the diameter of the gradient space $D \triangleq \max_{W,z,z'} \|\nabla\ell(z, W) - \nabla\ell(z', W)\|$.

1: Initialize W_0 randomly.
2: For $t = 1$ to T do:
3: Uniformly sample a mini-batch $\mathcal{B}_t$ of size τ from S without replacement;
4: Sample g_t from $\sigma\mathcal{N}(0, \mathbb{I})$;
5: Update $W_t \leftarrow W_{t-1} - \eta_t\left[\frac{1}{\tau}\sum_{z\in\mathcal{B}_t}\nabla\ell(z, W_{t-1}) + g_t\right]$.

Algorithm 2 Differentially Private Federated Learning

Require: Clients $\{c_1, ...c_N\}$, Gaussian noise variance σ, size of mini-batch τ, learning rate $\{\eta_1, \cdots, \eta_T\}$, iteration steps T, positive constant L.

1: Initialize W_0 randomly.
2: For $t = 1$ to T do:
3: Uniformly sample a mini-batch of clients $\mathcal{B}_t$ of size τ without replacement;
4: Randomly sample g_t from $\mathcal{N}(0, L^2\sigma^2 I)$;
5: Central curator distributes W_{t-1} to the clients in the mini-batch $\mathcal{B}_t$;
6: Update $W_t \leftarrow W_{t-1} + \eta_t\left(\frac{1}{\tau}\sum_{c\in\mathcal{B}_t}\frac{\text{ClientUpdate}(c, W_{t-1})}{\max\left(1, \frac{\|\text{ClientUpdate}(c, W_{t-1})\|_2}{L}\right)} + g_t\right)$.

By the above two charts, one can easily observe that SGLD is a special case of IGMM with $g = \nabla\ell$, and agnostic federated learning is also a special case of IGMM with $g = \frac{\text{ClientUpdate}}{\max\left(1, \frac{\|\text{ClientUpdate}\|_2}{L}\right)}$ and $D = 2L$.

C.2 PROOF OF THEOREM 7

IGMM applies the sub-sampling technique (i.e., mini-batch) to amplify differential privacy. Therefore, before the proof of Theorem 7, we first present a lemma from [Balle et al., 2018] which provide bound of differential privacy parameters after sub-sampling uniformly without replacement.

Lemma 5 (c.f. Theorem 9, [Balle et al., 2018])**.** *Let $\mathcal{M}^o : \mathcal{Z}^m \mapsto \Delta\mathcal{H}$ be any mechanism preserving (ε, δ) differential privacy. Let $\mathcal{M}^{wo} : \mathcal{Z}^N \mapsto \Delta\mathcal{Z}^m$ be the uniform sub-sampling without replacement mechanism. Then mechanism $\mathcal{M}^o \circ \mathcal{M}^{wo}$ satisfy $(\log(1 + (m/N)(e^\varepsilon - 1)), m\delta/N)$ differential privacy.*

Proof of Theorem 7. Before the start of the proof, we define several notations. We denote $G_{\mathcal{B}}(W) \triangleq \frac{1}{\|\mathcal{B}\|}\sum_{z\in\mathcal{B}} g(z, W)$ as the mean of g over $\mathcal{B}$ for brevity. We also use $\boldsymbol{p}$ as the probability density, with $\boldsymbol{p}^V$ the probability density conditional on any random variable V.

We first calculate the differential privacy of each step. To begin with, step 3 in Algorithm 1 is equivalent to uniformly sampling a mini-batch $\mathcal{I}_t$ from index set $[N]$ with size τ without replacement and letting $\mathcal{B}_t = S_{\mathcal{I}_t}$. Furthermore, for fixed W_{t-1}, $\mathcal{I}$, and any two adjacent sample sets S and S', we have

$$\begin{aligned}\frac{\boldsymbol{p}^{S,\mathcal{I}_t}(W_t = W|W_{t-1})}{\boldsymbol{p}^{S',\mathcal{I}_t}(W_t = W|W_{t-1})} &= \frac{\boldsymbol{p}^{S,\mathcal{I}_t}(\eta_t(G_{S_{\mathcal{I}}}(W_{t-1}) + \mathcal{N}(0,\sigma^2\mathbb{I})) = W - W_{t-1})}{\boldsymbol{p}^{S',\mathcal{I}_t}(\eta_t(G_{S'_{\mathcal{I}}}(W_{t-1}) + \mathcal{N}(0,\sigma^2\mathbb{I})) = W - W_{t-1})}\\ &= \frac{\boldsymbol{p}^{\mathcal{I}_t, W_{t-1}}(\mathcal{N}(0,\sigma^2\mathbb{I}) = W')}{\boldsymbol{p}^{S,S',\mathcal{I}_t,W_{t-1}}(G_{S'_{\mathcal{I}}}(W_{t-1}) - G_{S_{\mathcal{I}}}(W_{t-1}) + \mathcal{N}(0,\sigma^2\mathbb{I}) = W')},\end{aligned}$$

where $\eta_t W' = W - W_{t-1} - \eta_t G_{S_{\mathcal{I}}}(W_{t-1})$. Therefore, if $W \sim W_{t-1} + \eta_t(G_{S_{\mathcal{I}}}(W_{t-1}) + \mathcal{N}(0,\sigma\mathbb{I}))$, then $W' \sim G_{S_{\mathcal{I}}}(W_{t-1}) + \mathcal{N}(0,\sigma\mathbb{I})$. Define

$$D^{S,S',\mathcal{I}_t,W_{t-1}}(W') = \log\frac{\boldsymbol{p}^{\mathcal{I}_t, W_{t-1}}(\mathcal{N}(0,\sigma^2\mathbb{I}) = W')}{\boldsymbol{p}^{S,S',\mathcal{I}_t,W_{t-1}}(G_{S'_{\mathcal{I}}}(W_{t-1}) - G_{S_{\mathcal{I}}}(W_{t-1}) + \mathcal{N}(0,\sigma^2\mathbb{I}) = W')},$$

which by the definition of Gaussian distribution further leads to

$$\begin{aligned}D(W') =& -\frac{\|W'\|^2}{2\sigma^2} + \frac{\|W' - G_{S'_{\mathcal{I}}}(W_{t-1}) + G_{S_{\mathcal{I}}}(W_{t-1})\|^2}{2\sigma^2}\\ =& \frac{2\langle W', -G_{S'_{\mathcal{I}}}(W_{t-1}) + G_{S_{\mathcal{I}}}(W_{t-1})\rangle + \|G_{S'_{\mathcal{I}}}(W_{t-1}) - G_{S_{\mathcal{I}}}(W_{t-1})\|^2}{2\sigma^2}.\end{aligned}$$

Denote $-G_{S'_{\mathcal{I}}}(W_{t-1}) + G_{S_{\mathcal{I}}}(W_{t-1})$ as $\mathbf{v}$. By the definition of D, we have that

$$\|\mathbf{v}\| < \frac{1}{\tau}D.$$

On the other hand, since $\langle \mathbf{v}, W'\rangle \sim \mathcal{N}(0, \|\mathbf{v}\|^2\sigma^2)$, by Chernoff Bound technique,

$$\begin{aligned}\mathbb{P}\left(\langle \mathbf{v}, W'\rangle \geq \frac{\sqrt{2}D\sigma}{\tau}\sqrt{\log\frac{1}{\delta}}\right) &\leq \mathbb{P}\left(\langle \mathbf{v}, W'\rangle \geq \sqrt{2}\|\mathbf{v}\|\sigma\sqrt{\log\frac{1}{\delta}}\right)\\ &\leq \min_t e^{-\sqrt{2}t\|\mathbf{v}\|\sigma\sqrt{\log\frac{1}{\delta}}}\mathbb{E}(e^{t\langle \mathbf{v}, W'\rangle})\\ &= \delta.\end{aligned}$$

Therefore, with probability at least $1 - \delta$ with respect to W', we have that

$$D(W') \leq \frac{\sqrt{2}D\sigma\frac{1}{\tau}\sqrt{\log\frac{1}{\delta}} + \frac{1}{\tau^2}D^2}{2\sigma^2}.$$

By Lemma 5, we have that step t is $(\tilde{\varepsilon}, \frac{\tau}{N}\delta)$-differentially private. Applying Theorem 2 with $\varepsilon = \varepsilon_3'$ we can prove the differential privacy.

By applying Theorem 1, we can prove the generalization bound.

The proof is completed. □

C.3 PROOF OF COROLLARY 3

Proof of Corollary 3. We first calculate the asymptotic bounds for privacy parameters. Let $\delta = \Theta(1/N^2)$, we have

$$\begin{aligned}
\tilde{\varepsilon} =& \log\left(\frac{N-\tau}{N} + \frac{\tau}{N}\exp\left(\frac{\sqrt{2}D\sigma\frac{1}{\tau}\sqrt{\log\frac{1}{\delta}} + \frac{1}{\tau^2}D^2}{2\sigma^2}\right)\right) \\
=& \log\left(\frac{N-\tau}{N} + \frac{\tau}{N}\exp\left(\Theta(\sqrt{\log N})\right)\right) \\
=& \log\left(1 + \frac{\tau}{N}\left(\left(\exp\left(\Theta(\sqrt{\log N})\right)\right) - 1\right)\right) \\
=& \tilde{\Theta}\left(\frac{1}{N}\right). \qquad (29)
\end{aligned}$$

Furthermore, let $\tilde{\delta} = \Theta(\frac{1}{N^2})$. By eq.(29), we have

$$\begin{aligned}
\varepsilon' =& \sqrt{2T\log\left(\frac{1}{\tilde{\delta}}\right)\tilde{\varepsilon}^2} + T\tilde{\varepsilon}\frac{e^{\tilde{\varepsilon}}-1}{e^{\tilde{\varepsilon}}+1} \\
=& \Theta\left(\sqrt{\log N}\tilde{\Theta}\left(\frac{1}{N}\right)\right) + \Theta\left(\tilde{\Theta}\left(\frac{1}{N}\right)^2\right) = \tilde{\Theta}\left(\frac{1}{N}\right).
\end{aligned}$$

On the other hand, by $\tilde{\delta}' \leq \tilde{\delta} = \mathcal{O}(1/N^2)$, and

$$\begin{aligned}
& 2 - \left(1 - \frac{\delta}{1+e^{\tilde{\varepsilon}}}\right)^T - \left(1 - \frac{\delta e^{\tilde{\varepsilon}}}{1+e^{\tilde{\varepsilon}}}\right)^{\left\lceil\frac{N\varepsilon'}{\tau\tilde{\varepsilon}}\right\rceil}\left(1 - \frac{\delta}{1+e^{\tilde{\varepsilon}}}\right)^{T-\left\lceil\frac{N\varepsilon'}{\tau\tilde{\varepsilon}}\right\rceil} \\
=& \mathcal{O}(\delta) = \mathcal{O}\left(\frac{1}{N^2}\right),
\end{aligned}$$

we have

$$\delta' = 2 - \left(1 - \frac{\delta}{1+e^{\tilde{\varepsilon}}}\right)^T - \left(1 - \frac{\delta e^{\tilde{\varepsilon}}}{1+e^{\tilde{\varepsilon}}}\right)^{\left\lceil\frac{N\varepsilon'}{\tau\tilde{\varepsilon}}\right\rceil}\left(1 - \frac{\delta}{1+e^{\tilde{\varepsilon}}}\right)^{T-\left\lceil\frac{N\varepsilon'}{\tau\tilde{\varepsilon}}\right\rceil} + \tilde{\delta}' = \mathcal{O}\left(\frac{1}{N^2}\right).$$

When N is large enough, we have $\mathcal{A}^{fed}$ is $\left(\tilde{\Theta}(1/N), \tilde{\mathcal{O}}(1/N)\right)$-differentially private, and thus $\left(\tilde{\Theta}(1/\sqrt{N}), \tilde{\mathcal{O}}(1/N)\right)$-differentially private. By

$$\frac{2e^{-2\varepsilon'}\delta'}{\varepsilon'}\ln\left(\frac{2}{\varepsilon'}\right) = \frac{2e^{-2\tilde{\Theta}(\frac{1}{\sqrt{N}})}\mathcal{O}(\frac{1}{N^2})}{\tilde{\Theta}(\frac{1}{\sqrt{N}})}\ln\left(\frac{2}{\tilde{\Theta}(\frac{1}{\sqrt{N}})}\right) = \tilde{\mathcal{O}}\left(\frac{1}{N^{\frac{3}{2}}}\right).$$

Since when N is large enough,

$$N \geq \frac{2}{0.077\tilde{\Theta}(\frac{1}{\sqrt{N}})^2}\ln\left(\frac{43}{254e^{-1.7\tilde{\Theta}(\frac{1}{\sqrt{N}})}\tilde{\mathcal{O}}\left(\frac{1}{N^{\frac{3}{2}}}\right)}\right),$$

by Theorem 1, the proof is completed. □