A Details of Section 3

A.1 Randomized Response with Constant Probability p_{const}

Algorithm 2 Randomized Response Majority (RR)

Input: K (ϵ, Δ)-DP mechanisms $\{M_i\}_{i=1}^K$, noise function $\gamma : \{0, \dots, K\} \to [0, 1]$, dataset $\mathcal{D}$, privacy allowance $1 \le m \le K$, failure probability $\delta \ge \Delta \ge 0$
Output: $(m\epsilon, \delta)$-DP majority vote of $\{M_i\}_{i=1}^K$
Compute a *constant* probability $p_{const} \in [0, 1]$
Flip the p_{const}- biased coin
if Head (with probability p_{const}) **then**
  $\mathcal{S} = \{S_1, .., S_k\}$, where $S_i \sim M_i(\mathcal{D})$
  $\mathcal{L} = \sum_{i=1}^K S_i$
  Output $\mathbb{I}\{\frac{1}{K}\mathcal{L} \ge \frac{1}{2}\}$
else
  Output 0/1 with equal probability
end if

We show the magnitude of p_{const} in RR (Algorithm 2) to solve Problem 1.1, such that the output is $(m\epsilon, \delta)$-DP, in Lemma A.1.

Lemma A.1. *Consider using RR (Algorithm 2) to solve Problem 1.1. Let the majority of K (ϵ, Δ)-differentially private mechanisms be $(\tau\epsilon, \lambda)$-differentially private, where $\tau \in [1, K]$ and $\lambda \in [0, 1)$ are computed by simple composition (Theorem 2.2) or general composition (Theorem 2.3). If*

$$p_{const} \le \frac{e^{m\epsilon} - 1 + 2\delta}{\frac{2(e^{\tau\epsilon} - e^{m\epsilon} + (1+e^{m\epsilon})\lambda)}{e^{\tau\epsilon}+1}} + e^{m\epsilon} - 1 \tag{4}$$

then RR is $(m\epsilon, \delta)$-differentially private.

Proof of Lemma A.1. Let $x \in \{0, 1\}$ denote the output of RR. Let $q_x = \Pr[\mathcal{L}(\mathcal{D}) = x]$ and $q'_x = \Pr[\mathcal{L}(\mathcal{D}') = x]$, where $\mathcal{L}(\mathcal{D}) = \sum_{i=1}^K M_i(\mathcal{D})$, $\mathcal{L}(\mathcal{D}') = \sum_{i=1}^K M_i(\mathcal{D}')$ and $\mathcal{D}$, $\mathcal{D}'$ are adjacent datasets. Recall each mechanism M_i is (ϵ, Δ)-differentially private, and the majority of the outputs of $\{M_i\}_{i=1}^K$ is $(\tau\epsilon, \lambda)$-differentially private. When $\Delta = 0$, using simple composition, $\tau = K$ and $\lambda = 0$. When $\Delta > 0$, using general composition $\tau \approx \sqrt{K}$ and $\lambda \approx K\Delta$. By definition of differential privacy (Definition 2.1), all of the following four constraints on q_x, q'_x apply:

$$\begin{aligned} q_x &\le e^{\tau\epsilon} q'_x + \lambda, \quad \text{and} \quad 1 - q'_x \le e^{\tau\epsilon}(1 - q_x) + \lambda \\ q'_x &\le e^{\tau\epsilon} q_x + \lambda, \quad \text{and} \quad 1 - q_x \le e^{\tau\epsilon}(1 - q'_x) + \lambda \end{aligned}$$

To ensure RR is $(m\epsilon, \delta)$-differentially private, p_{const} needs to be such that for all possible $q_x, q'_x \in [0, 1]$,

$$\Pr[\mathsf{RR}(\mathcal{D}) = x] \le e^{m\epsilon} \Pr[\mathsf{RR}(\mathcal{D}') = x] + \delta \tag{5}$$

$$p_{const} \cdot q_x + \frac{1}{2}(1 - p_{const}) \le e^{m\epsilon}(p_{const} \cdot q'_x + \frac{1}{2}(1 - p_{const})) + \delta \tag{6}$$

$$(q_x - e^{m\epsilon} q'_x + \frac{1}{2}e^{m\epsilon} - \frac{1}{2}) \cdot p_{const} \le \frac{1}{2}e^{m\epsilon} - \frac{1}{2} + \delta \tag{7}$$

Let $h(q_x, q'_x) := q_x - e^{m\epsilon} q'_x + \frac{1}{2}e^{m\epsilon} - \frac{1}{2}$. The above inequality of p_{const} (Eq. 7) needs to hold for worst case output probabilities q^*_x, q'^*_x that cause the maximum privacy loss. That is, p_{const} needs to satisfy

$$p_{const} \cdot max_{q_x, q'_x} h(q_x, q'_x) \le \frac{1}{2}e^{m\epsilon} - \frac{1}{2} + \delta \tag{8}$$

To find the worst case output probabilities, we solve the following Linear Programming (LP) problem:

$$\text{Objective:} \quad \max_{q_x, q'_x} \quad h(q_x, q'_x) := q_x - e^{m\epsilon} q'_x + \frac{1}{2} e^{m\epsilon} - \frac{1}{2} \tag{9}$$

$$\text{Subject to:} \quad 0 \leq q_x \leq 1, 0 \leq q'_x \leq 1 \tag{10}$$

$$q_x \leq e^{\tau\epsilon} q'_x + \lambda, 1 - q'_x \leq e^{\tau\epsilon}(1 - q_x) + \lambda \tag{11}$$

$$q'_x \leq e^{\tau\epsilon} q_x + \lambda, 1 - q_x \leq e^{\tau\epsilon}(1 - q'_x) + \lambda \tag{12}$$

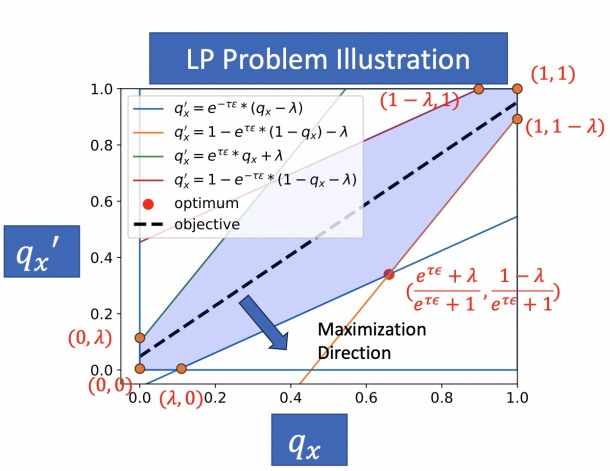

Figure 3: A visualization of the above LP problem.

The optimum of any LP problem is at the corners of the feasible region, which is bounded by the optimization constraints. We plot the feasible region $\mathcal{F}$ and the objective of the above LP problem in Figure 3. Here, $(q_x^*, q_x'^*) = \arg\max_{q_x, q'_x} h(q_x, q'_x) \in \{(0,0), (1,1), (0,\lambda), (\lambda, 0), (1-\lambda, 1), (1, 1-\lambda), (\frac{1-\lambda}{e^{\tau\epsilon}+1}, \frac{e^{\tau\epsilon}+\lambda}{e^{\tau\epsilon}+1}), (\frac{e^{\tau\epsilon}+\lambda}{e^{\tau\epsilon}+1}, \frac{1-\lambda}{e^{\tau\epsilon}+1})\}$. The optimum of the LP problem – that is, the worse case probabilities $q_x^*, q_x'^*$ – is,

$$q_x^* = \frac{e^{\tau\epsilon} + \lambda}{e^{\tau\epsilon} + 1}, \quad q_x'^* = \frac{1 - \lambda}{e^{\tau\epsilon} + 1} \tag{13}$$

By Eq. 8,

$$p_{const} \cdot \left(\frac{e^{\tau\epsilon} + \lambda}{e^{\tau\epsilon} + 1} - e^{m\epsilon} \frac{1 - \lambda}{e^{\tau\epsilon} + 1} + \frac{1}{2} e^{m\epsilon} - \frac{1}{2} \right) \leq \frac{1}{2}(e^{m\epsilon} - 1) + \delta \tag{14}$$

$$p_{const} \cdot \left(\frac{e^{\tau\epsilon} - e^{m\epsilon} + (1 + e^{m\epsilon})\lambda}{e^{\tau\epsilon} + 1} + \frac{1}{2}(e^{m\epsilon} - 1) \right) \leq \frac{1}{2}(e^{m\epsilon} - 1) + \delta \tag{15}$$

$$p_{const} \leq \frac{e^{m\epsilon} - 1 + 2\delta}{\frac{2(e^{\tau\epsilon} - e^{m\epsilon} + (1 + e^{m\epsilon})\lambda)}{e^{\tau\epsilon}+1} + e^{m\epsilon} - 1} \tag{16}$$

For small m, ϵ, K, using the approximation $e^y \approx 1 + y$ and that $\tau\epsilon < 2$,

$$p_{const} \approx \frac{m\epsilon + 2\delta}{\frac{2(\tau\epsilon - m\epsilon + (2 + m\epsilon)\lambda)}{\tau\epsilon + 2} + m\epsilon} \approx \frac{m\epsilon + 2\delta}{\tau\epsilon + (2 + m\epsilon)\lambda} \tag{17}$$

In the pure differential privacy setting, $\delta = 0, \lambda = 0, \tau = K$, and so $p_{const} \approx \frac{m}{K}$; and in the approximate differential privacy setting, $\lambda \approx 0, \delta \approx 0, \tau \approx \sqrt{K}$, and so $p_{const} \approx \frac{m}{\sqrt{K}}$. □

Algorithm 3 Subsampling Majority (SubMaj)

1: Input: K (ϵ, Δ)-DP mechanisms $\{M_i\}_{i=1}^K$, noise function $\gamma : \{0, \dots, K\} \to [0, 1]$, dataset $\mathcal{D}$, privacy allowance $1 \le m \le K$, failure probability $\delta \ge \Delta \ge 0$
2: Output: $(m\epsilon, \delta)$-DP majority vote of $\{M_i\}_{i=1}^K$
3: $\mathcal{S} = \{S_1, .., S_k\}$, where $S_i \sim M_i(\mathcal{D})$
4: $\mathcal{J}_m \leftarrow m$ indices chosen uniformly at random from $[K]$ without replacement
5: $\widehat{\mathcal{L}} = \sum_{j \in \mathcal{J}} S_j$
6: Output $\mathbb{I}\{\frac{1}{m}\widehat{\mathcal{L}} \ge \frac{1}{2}\}$

A.2 Proof of Lemma 3.1

Lemma A.2 (Restatement of Lemma 3.1)**.** *Consider Problem 1.1, with the privacy allowance $m \in [K]$. Consider the data-dependent algorithm that computes $\mathcal{L}(\mathcal{D})$ and then applies RR with probability p_γ. If $p_\gamma = \gamma_{Sub}(l)$, where $l \in \{0, 1, \dots, K\}$ is the value of $\mathcal{L}(\mathcal{D})$, i.e., the (random) sum of observed outcomes on dataset $\mathcal{D}$, and $\gamma_{Sub} : \{0, 1, \dots, K\} \to [0, 1]$ is*

$$\begin{aligned}\gamma_{Sub}(l) &= \gamma_{Sub}(K - l) \\ &= \begin{cases} 1 - 2\sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} & \text{if } m \text{ is odd} \\ 1 - 2\sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} - \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}} & \text{if } m \text{ is even} \end{cases}\end{aligned}$$

then the majority of m out of K subsampled mechanisms without replacement and the output of our data-dependent RR algorithm have the same distribution.

Proof of Lemma 3.1. Let $\mathcal{L} = \sum_{i=1}^K S_i$ be the sum of observed outcomes from K mechanisms. Following Algorithm 3, $\mathcal{J}_m$ denotes the m indices chosen uniformly at random from $[K]$ without replacement. Conditioned on $\mathcal{L}$, notice the output of SubMaj follows a hypergeometric distribution. The output probability of SubMaj is

$$\Pr[\mathsf{SubMaj}(\mathcal{D}) = 1] = \sum_{l=0}^{K} \Pr[\mathsf{SubMaj}(\mathcal{D}) = 1 \mid \mathcal{L} = l] \cdot \Pr[\mathcal{L} = l] \tag{18}$$

$$= \sum_{l=0}^{K} \Pr[\sum_{j \in \mathcal{J}_m} S_j \ge \frac{m}{2} \mid \mathcal{L} = l] \cdot \Pr[\mathcal{L} = l] \tag{19}$$

$$= \begin{cases} \sum_{l=0}^{K} (\sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}}) \cdot \Pr[\mathcal{L} = l] & \text{if } m \text{ is odd} \\ \sum_{l=0}^{K} (\sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} + \frac{1}{2}\frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}}) \cdot \Pr[\mathcal{L} = l] & \text{if } m \text{ is even} \end{cases} \tag{20}$$

Consider an arbitrary noise function $\gamma_{Sub} : \{0, 1, \dots, K\} \to [0, 1]$. Let $\mathsf{RR\text{-}d}(\mathcal{D})$ denote the output of the data-dependent RR-d on dataset $\mathcal{D}$, where RR-d has the *non-constant* probability set by γ_{Sub}. The output probability of RR is,

$$\Pr[\mathsf{RR\text{-}d}(\mathcal{D}) = 1] = \sum_{l=0}^{K} \Pr[\mathsf{RR\text{-}d}(\mathcal{D}) = 1 \mid \mathcal{L} = l] \cdot \Pr[\mathcal{L} = l] \tag{21}$$

$$= \sum_{l=0}^{K} (\gamma_{Sub}(l) \cdot \mathbb{I}\{l \ge \frac{K+1}{2}\} + \frac{1}{2}(1 - \gamma_{Sub}(l))) \cdot \Pr[\mathcal{L} = l] \tag{22}$$

We want $\Pr[\mathsf{RR\text{-}d}(\mathcal{D}) = 1] = \Pr[\mathsf{Submaj}(\mathcal{D}) = 1]$.

If m is odd, for any $l \leq \frac{K-1}{2}$, this is

$$\frac{1}{2}(1-\gamma_{Sub}(l)) = \sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}}$$
$$\Rightarrow \gamma_{Sub}(l) = 1 - 2\sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} \tag{23}$$

and for any $l \geq \frac{K+1}{2}$, this is

$$\frac{1}{2} + \frac{1}{2}\gamma_{Sub}(l) = \sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}}$$
$$\Rightarrow \gamma_{Sub}(l) = 2\sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} - 1 \tag{24}$$

Similarly, if m is even, for any $l \leq \frac{K-1}{2}$, this is

$$\frac{1}{2}(1-\gamma_{Sub}(l)) = \sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} + \frac{1}{2}\frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}}$$
$$\Rightarrow \gamma_{Sub}(l) = 1 - 2\sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} - \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}} \tag{25}$$

and for any $l \geq \frac{K+1}{2}$, this is

$$\frac{1}{2} + \frac{1}{2}\gamma_{Sub}(l) = \sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} + \frac{1}{2}\frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}}$$
$$\Rightarrow \gamma_{Sub}(l) = 2\sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} + \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}} - 1 \tag{26}$$

Next, we show the above γ_{Sub} is indeed symmetric around $\frac{K}{2}$. For any $l \leq \frac{K-1}{2}$, there is $K - l \geq \frac{K+1}{2}$. If m is odd,

$$\begin{aligned}\gamma_{Sub}(K-l) &= 2\sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{K-l}{j}\binom{l}{m-j}}{\binom{K}{m}} - 1 = 2\Big(1 - \sum_{j=1}^{\frac{m-1}{2}} \frac{\binom{K-l}{j}\binom{l}{m-j}}{\binom{K}{m}}\Big) - 1 \\ &= 1 - 2\sum_{j=1}^{\frac{m-1}{2}} \frac{\binom{K-l}{j}\binom{l}{m-j}}{\binom{K}{m}} = 1 - 2\sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} \\ &= \gamma_{Sub}(l)\end{aligned} \tag{27}$$

Similarly, if m is even,

$$\begin{aligned}\gamma_{Sub}(K-l) &= 2\sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{K-l}{j}\binom{l}{m-j}}{\binom{K}{m}} + \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}} - 1 = 2\Big(1 - \sum_{j=1}^{\frac{m}{2}-1} \frac{\binom{K-l}{j}\binom{l}{m-j}}{\binom{K}{m}} - \frac{1}{2}\frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}}\Big) - 1 \\ &= 1 - 2\sum_{j=1}^{\frac{m}{2}-1} \frac{\binom{K-l}{j}\binom{l}{m-j}}{\binom{K}{m}} - \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}} = 1 - 2\sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} - \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}}\end{aligned}$$

$$= \gamma_{Sub}(l) \tag{28}$$

Now, combining Eq. 23, Eq. 24 and Eq. 27, if m is odd, setting γ_{Sub} as

$$\gamma_{Sub}(l) = \gamma_{Sub}(K-l) = 1 - 2\sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} \tag{29}$$

makes RR-d have the same output distribution as SubMaj.

Similarly, combining Eq. 25, Eq. 26 and Eq. 28, if m is even, setting γ_{Sub} as

$$\gamma_{Sub}(l) = \gamma_{Sub}(K-l) = 1 - 2\sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} - \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}} \tag{30}$$

makes RR-d have the same output distribution as SubMaj.

□

A.3 Proof of Lemma 3.2

Lemma A.3 (Restatement of Lemma 3.2)**.** *Let $\mathcal{A}$ be an (ϵ, δ)-differentially private algorithm, where $\epsilon \in (0, \frac{1}{2})$ and $\delta \in [0, \frac{1}{2})$, that computes the majority of K (ϵ, δ)-differentially private mechanisms $M_1, \ldots, M_K$, where $M_i : \mathcal{D} \to \{0, 1\}$ on dataset $\mathcal{D}$ and $\Pr[M_i(\mathcal{D}) = 1] = p_i, \forall i \in [K]$. Then, the error $\mathcal{E}(\mathcal{A}) \geq |\Pr[g(\mathcal{S}) = 1] - \frac{1}{K}\sum_{i=1}^{K} p_i|$, where $g(\mathcal{S})$ is the probability of the true majority output being 1 as defined in Definition 1.1.*

Proof. Consider the setting where M_i's are i.i.d., i.e., $\Pr[M_i(\mathcal{D}) = 1] = p, \forall i \in [K]$ for some $p \in [0, 1]$ on any dataset $\mathcal{D}$. Then, it suffices to show $\mathcal{E}(\mathcal{A}) \geq |\Pr[g(\mathcal{S})] = 1 - p|$, because a lower bound in this special case would indicate a lower bound for the more general case, where p_i's can be different.

Construct a dataset $\mathcal{D}_0$ and K mechanisms $\{M_i\}_{i=1}^{K}$ such that $\Pr[\mathcal{M}_i(\mathcal{D}_0) = 1] = \Pr[\mathcal{M}_i(\mathcal{D}_0) = 0] = \frac{1}{2}$ and without loss of generality, we may assume $\Pr[\mathcal{A}(\mathcal{D}_0) = 1] \leq \frac{1}{2}$.

Next, we construct a sequence of datasets $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_L$, such that $\mathcal{D}_j$ and $\mathcal{D}_{j+1}$ are neighboring datasets tha t differ in one entry, for all $j \in [L-1]$, and $\Pr[M_i(\mathcal{D}_j) = 1] = \frac{1}{2}e^{j\epsilon} + \sum_{l=0}^{j-1} e^{l\epsilon}\delta$, $\forall i \in [K]$, $\forall j \in [L]$. Choose $L \in \mathbb{N}$ such that $\frac{1}{2}e^{L\epsilon} + \sum_{l=0}^{L-1} e^{\epsilon l}\delta = p$, for some $1 \geq p > \frac{1}{2}$.

Now, by definition of differential privacy,

$$\begin{aligned}
&\Pr[\mathcal{A}(\mathcal{D}_1) = 1] \leq e^{\epsilon}\Pr[\mathcal{A}(\mathcal{D}_0) = 1] + \delta \\
&\Pr[\mathcal{A}(\mathcal{D}_2) = 1] \leq e^{\epsilon}\Pr[\mathcal{A}(\mathcal{D}_1) = 1] + \delta \leq e^{2\epsilon}\Pr[\mathcal{A}(\mathcal{D}_0) = 1] + e^{\epsilon}\delta + \delta \\
&\ldots \\
&\Pr[\mathcal{A}(\mathcal{D}_L) = 1] \leq e^{L\epsilon}\Pr[\mathcal{A}(\mathcal{D}_0) = 1] + \sum_{l=0}^{L-1} e^{\epsilon l}\delta \leq e^{L\epsilon}\frac{1}{2} + \sum_{l=0}^{L-1} e^{\epsilon l}\delta = p
\end{aligned}$$

Since the probability of true majority being 1 on dataset $\mathcal{D}_L$ is $\Pr[g(\mathcal{S}) = 1] \geq p > \frac{1}{2}$, there is

$$\mathcal{E}(\mathcal{A}) = |\Pr[g(\mathcal{S}) = 1] - \Pr[\mathcal{A}(\mathcal{D}_L) = 1]| \geq \Pr[g(\mathcal{S}) = 1] - p$$

□

A.4 Proof of Lemma 3.3

Lemma A.4 (Restatement of Lemma 3.3)**.** *Let $\mathcal{A}$ be any randomized algorithm to compute the majority function g on $\mathcal{S}$ such that for all $\mathcal{S}$, $\Pr[\mathcal{A}(\mathcal{S}) = g(\mathcal{S})] \geq 1/2$ (i.e. $\mathcal{A}$ is at least as good as a random guess). Then, there exists a a general function $\gamma : \{0,1\}^{K+1} \to [0,1]$ such that if one sets p_γ by $\gamma(\mathcal{S})$ in DaRRM, the output distribution of DaRRM_γ is the same as the output distribution of $\mathcal{A}$.*

Proof of Lemma 3.3. For some $\mathcal{D}$ and conditioned on $\mathcal{S}$, we see that by definition $\Pr[\mathsf{DaRRM}_\gamma(\mathcal{S}) = g(\mathcal{S})] = \gamma(\mathcal{S}) + (1/2)(1 - \gamma(\mathcal{S}))$. We want to set γ such that $\Pr[\mathsf{DaRRM}_\gamma(\mathcal{S}) = g(\mathcal{S})] = \Pr[\mathcal{A}(\mathcal{S}) = g(\mathcal{S})]$. Therefore, we set $\gamma(\mathcal{S}) = 2\Pr[\mathcal{A}(\mathcal{S}) = g(\mathcal{S})] - 1$.

Lastly, we need to justify that $\gamma \in [0,1]$. Clearly, $\gamma(\mathcal{S}) \leq 2 - 1 \leq 1$ since $\Pr[\mathcal{A}(\mathcal{S}) = g(\mathcal{S})] \leq 1$. Note that the non-negativity follows from assumption. □

A.5 Proof of Lemma 3.4

Lemma A.5 (Restatement of Lemma 3.4)**.** *Consider using DaRRM (Algorithm 1) to solve Problem 1.1, let $\alpha_l = \Pr[\mathcal{L}(\mathcal{D}) = l]$ and $\alpha'_l = \Pr[\mathcal{L}(\mathcal{D}') = l]$, where $\mathcal{D}$ and $\mathcal{D}'$ are adjacent datasets and $l \in \{0, \ldots, K\}$. For a noise function $\gamma : \{0, 1, \ldots, K\} \to [0,1]$ such that $\gamma(l) = \gamma(K - l), \forall l$, DaRRM_γ is $(m\epsilon, \delta)$-differentially private if and only if for all α_l, α'_l, the following holds,*

$$f(p_1, \ldots, p_K, p'_1, \ldots, p'_K; \gamma) \leq e^{m\epsilon} - 1 + 2\delta \tag{31}$$

where f is called the ***privacy cost objective*** *and*

$$f(p_1, \ldots, p_K, p'_1, \ldots, p'_K; \gamma) := \sum_{l=0}^{\frac{K-1}{2}} (e^{m\epsilon}\alpha'_l - \alpha_l) \cdot \gamma(l) + \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l) \cdot \gamma(l)$$

Proof of Lemma 3.4. By the definition of differential privacy (Definition 2.1),

$$\begin{aligned}
&\mathsf{DaRRM}_\gamma \text{ is } (m\epsilon, \delta)\text{-differentially private} \\
&\iff \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1] \leq e^{m\epsilon} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1] + \delta, \\
&\text{and } \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 0] \leq e^{m\epsilon} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 0] + \delta, \quad \forall \text{ adjacent datasets } \mathcal{D}, \mathcal{D}'
\end{aligned} \tag{32}$$

Let random variables $\mathcal{L}(\mathcal{D}) = \sum_{i=1}^{K} S(\mathcal{D})$ and $\mathcal{L}(\mathcal{D}') = \sum_{i=1}^{K} S(\mathcal{D}')$ be the sum of observed outcomes on adjacent datasets $\mathcal{D}$ and $\mathcal{D}'$, based on which one sets p_γ in DaRRM. Let $\alpha_l = \Pr[\mathcal{L}(\mathcal{D}) = l]$ and $\alpha'_l = \Pr[\mathcal{L}(\mathcal{D}') = l]$, $\forall l \in \{0, 1, \ldots, K\}$.

Consider the output being 1.

$$\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1] \leq e^{m\epsilon} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1] + \delta \tag{33}$$

$$\iff \sum_{l=0}^{K} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1 \mid \mathcal{L}(\mathcal{D}) = l] \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] \tag{34}$$

$$\leq e^{m\epsilon} \Big(\sum_{l=0}^{K} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1 \mid \mathcal{L}(\mathcal{D}') = l] \cdot \Pr[\mathcal{L}(\mathcal{D}') = l] \Big) + \delta$$

$$\iff \sum_{l=0}^{K} \Big(\gamma(l) \cdot \mathbb{I}\{l \geq \frac{K}{2}\} + \frac{1}{2}(1 - \gamma(l)) \Big) \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] \tag{35}$$

$$\leq e^{m\epsilon} \Big(\sum_{l=0}^{K} \Big(\gamma(l) \cdot \mathbb{I}\{l \geq \frac{K}{2}\} + \frac{1}{2}(1 - \gamma(l)) \Big) \Big) \cdot \Pr[\mathcal{L}(\mathcal{D}') = l] \Big) + \delta$$

$$\iff \sum_{l=\frac{K+1}{2}}^{K} \Big(\gamma(l) + \frac{1}{2}(1-\gamma(l))\Big) \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] + \sum_{l=0}^{\frac{K-1}{2}} \frac{1}{2}(1-\gamma(l)) \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] \tag{36}$$

$$\leq e^{m\epsilon}\Big(\sum_{l=\frac{K+1}{2}}^{K} \Big(\gamma(l) + \frac{1}{2}(1-\gamma(l))\Big) \cdot \Pr[\mathcal{L}(\mathcal{D}) = l]\Big) + e^{m\epsilon}\Big(\sum_{l=0}^{\frac{K-1}{2}} \frac{1}{2}(1-\gamma(l)) \cdot \Pr[\mathcal{L}(\mathcal{D}') = l]\Big) + \delta$$

$$\iff \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}\gamma(l)\alpha_l - \sum_{l=0}^{\frac{K-1}{2}} \frac{1}{2}\gamma(l)\alpha_l + \frac{1}{2} \tag{37}$$

$$\leq e^{m\epsilon} \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}\gamma(l)\alpha'_l - e^{m\epsilon} \sum_{l=0}^{\frac{K-1}{2}} \frac{1}{2}\gamma(l)\alpha'_l + \frac{1}{2}e^{m\epsilon} + \delta$$

$$\iff \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) - \sum_{l=0}^{\frac{K-1}{2}} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) \leq e^{m\epsilon} - 1 + 2\delta \tag{38}$$

Similarly, consider the output being 0.

$$\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 0] \leq e^{m\epsilon} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 0] + \delta \tag{39}$$

$$\iff \sum_{l=0}^{K} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 0 \mid \mathcal{L}(\mathcal{D}) = l] \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] \tag{40}$$

$$\leq e^{m\epsilon}\Big(\sum_{l=0}^{K} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 0 \mid \mathcal{L}(\mathcal{D}') = l] \cdot \Pr[\mathcal{L}(\mathcal{D}') = l]\Big) + \delta$$

$$\iff \sum_{l=0}^{K} \Big(\gamma(l) \cdot \mathbb{I}\{l < \frac{K}{2}\} + \frac{1}{2}(1-\gamma(l))\Big) \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] \tag{41}$$

$$\leq e^{m\epsilon}\Big(\sum_{l=0}^{K} \gamma(l) \cdot \mathbb{I}\{l < \frac{K}{2}\} + \frac{1}{2}(1-\gamma(l))\Big) \cdot \Pr[\mathcal{L}(\mathcal{D}') = l] + \delta$$

$$\iff \sum_{l=0}^{\frac{K-1}{2}} \Big(\gamma(l) + \frac{1}{2}(1-\gamma(l))\Big) \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] + \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}(1-\gamma(l)) \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] \tag{42}$$

$$\leq e^{m\epsilon}\Big(\sum_{l=0}^{\frac{K-1}{2}} \Big(\gamma(l) + \frac{1}{2}(1-\gamma(l))\Big) \cdot \Pr[\mathcal{L}(\mathcal{D}') = l] + \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}(1-\gamma(l)) \cdot \Pr[\mathcal{L}(\mathcal{D}') = l]\Big) + \delta$$

$$\iff \sum_{l=0}^{\frac{K-1}{2}} \frac{1}{2}\gamma(l)\alpha_l - \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}\gamma(l)\alpha_l + \frac{1}{2} \tag{43}$$

$$\leq e^{m\epsilon} \sum_{l=0}^{\frac{K-1}{2}} \frac{1}{2}\gamma(l)\alpha'_l - e^{m\epsilon} \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}\gamma(l)\alpha'_l + \frac{1}{2}e^{m\epsilon} + \delta$$

$$\iff \sum_{l=0}^{\frac{K-1}{2}} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) - \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) \leq e^{m\epsilon} - 1 + 2\delta \tag{44}$$

Therefore, plugging Eq. 38 and Eq. 44 into Eq. 32,

$$\mathsf{DaRRM}_\gamma \text{ is } (m\epsilon, \delta)\text{-differentially private}$$

$$\iff \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) - \sum_{l=0}^{\frac{K-1}{2}} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) \leq e^{m\epsilon} - 1 + 2\delta \tag{45}$$

$$\text{and } \sum_{l=0}^{\frac{K-1}{2}} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) - \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) \leq e^{m\epsilon} - 1 + 2\delta \tag{46}$$

where $\alpha_l = \Pr[\mathcal{L}(\mathcal{D}) = l]$ and $\alpha'_l = \Pr[\mathcal{L}(\mathcal{D}') = l]$, $\forall l \in \{0, 1, \ldots, K\}$ and $\mathcal{D}, \mathcal{D}'$ are any adjacent datasets.

Next, we show if γ is symmetric around $\frac{K}{2}$, i.e., $\gamma(l) = \gamma(K - l)$, satisfying either one of Eq. 45 or Eq. 46 implies satisfying the other one. Following Eq. 45,

$$\sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) - \sum_{l=0}^{\frac{K-1}{2}} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) \leq e^{m\epsilon} - 1 + 2\delta \tag{47}$$

$$\iff \sum_{l=0}^{\frac{K-1}{2}} (\alpha_{K-l} - e^{m\epsilon}\alpha'_{K-l}) \cdot \gamma(K-l) - \sum_{l=\frac{K-1}{2}}^{K} (\alpha_{K-l} - e^{m\epsilon}\alpha'_{K-l}) \cdot \gamma(K-l) \leq e^{m\epsilon} - 1 + 2\delta \tag{48}$$

$$\iff \sum_{l=0}^{\frac{K-1}{2}} (\alpha_{K-l} - e^{m\epsilon}\alpha'_{K-l}) \cdot \gamma(l) - \sum_{l=\frac{K-1}{2}}^{K} (\alpha_{K-l} - e^{m\epsilon}\alpha'_{K-l}) \cdot \gamma(l) \leq e^{m\epsilon} - 1 + 2\delta \tag{49}$$

Since $\gamma(l) = \gamma(K - l)$

For analysis purpose, we rewrite Eq. 46 as

$$\sum_{l=0}^{\frac{K-1}{2}} (\widetilde{\alpha}_l - e^{m\epsilon}\widetilde{\alpha}'_l) \cdot \gamma(l) - \sum_{l=\frac{K-1}{2}}^{K} (\widetilde{\alpha}_l - e^{m\epsilon}\widetilde{\alpha}'_l) \cdot \gamma(l) \leq e^{m\epsilon} - 1 + 2\delta \tag{50}$$

and proceed by showing Eq. 49 $\iff$ Eq. 50.

Recall $p_i = \Pr[M_i(\mathcal{D}) = 1]$ and $p'_i = \Pr[M_i(\mathcal{D}') = 1]$. Observe $\mathcal{L}(\mathcal{D}) \sim \text{PoissonBinomial}(\{p_i\}_{i=1}^K)$ and $\mathcal{L}(\mathcal{D}') \sim \text{PoissonBinomial}(\{p'_i\}_{i=1}^K)$. Let $F_l = \{\mathcal{A} : |\mathcal{A}| = l, \mathcal{A} \subseteq [K]\}$, for any $l \in \{0, \ldots, K\}$, denote the set of all subsets of l integers that can be selected from $[K]$. Let $\mathcal{A}^c = [K] \setminus \mathcal{A}$ be $\mathcal{A}$'s complement set. Notice $F_{K-l} = \{\mathcal{A}^c : \mathcal{A} \in F_l\}$.

Since α denotes the pmf of the Poisson Binomial distribution at l, it follows that

$$\alpha_l = \Pr[\mathcal{L}(\mathcal{D}) = l] = \sum_{\mathcal{A} \in F_l} \Pi_{i \in \mathcal{A}} p_i \Pi_{j \in \mathcal{A}^c} (1 - p_j) \tag{51}$$

Consider $\beta_i = 1 - p_i, \forall i \in [K]$ and a new random variable $\mathcal{L}^\beta \sim \text{PoissonBinomial}(\{\beta_i\}_{i=1}^K)$, and let $\widetilde{\alpha}_l = \Pr[\mathcal{L}^\beta = 1]$. Observe that

$$\begin{aligned}
\widetilde{\alpha}'_l = \Pr[\mathcal{L}^\beta = l] &= \sum_{\mathcal{A} \in F_l} \Pi_{j \in \mathcal{A}} \beta_i \Pi_{i \in \mathcal{A}^c} (1 - \beta_i) = \sum_{\mathcal{A} \in F_l} \Pi_{j \in \mathcal{A}} (1 - p_j) \Pi_{i \in \mathcal{A}^c} p_i \\
&= \sum_{\mathcal{A}^c \in F_{K-l}} \Pi_{j \in \mathcal{A}} (1 - p_i) \Pi_{i \in \mathcal{A}^c} p_i = \sum_{\mathcal{A} \in F_{K-l}} \Pi_{i \in \mathcal{A}} p_i \Pi_{j \in \mathcal{A}^c} (1 - p_i)
\end{aligned}$$

$$= \alpha_{K-l} \tag{52}$$

Similarly, consider $\beta'_i = 1 - p'_i, \forall i \in [K]$ and a new random variable $\mathcal{L}'^{\beta} \sim \text{PoissonBinomial}(\beta'_i\}_{i=1}^{L})$, and let $\widetilde{\alpha}'_l = \Pr[\mathcal{L}'^{\beta} = 1]$. Then, $\widetilde{\alpha}'_l = \alpha'_{K-l}$.

Since Eq. 49 holds for all possible α_{K-l}, α'_{K-l}, Eq. 50 then holds for all $\widetilde{\alpha}_l, \widetilde{\alpha}'_l$ in the K-simplex, and so Eq. 50 follows by relabeling α_{K-l} as $\widetilde{\alpha}_l$ and α'_{K-l} as $\widetilde{\alpha}'_l$.

The above implies Eq. 45 $\iff$ Eq. 46. Therefore,

$$\begin{aligned} &\mathsf{DaRRM}_\gamma \text{ is } (m\epsilon, \delta)\text{-differentially private} \\ \iff &\underbrace{\sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) - \sum_{l=0}^{\frac{K-1}{2}} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l)}_{:= f(p_1,\ldots,p_K,p'_1,\ldots,p'_K;\gamma)} \leq e^{m\epsilon} - 1 + 2\delta \end{aligned} \tag{53}$$

□

B Details of Section 4: Provable Privacy Amplification

In this section, we consider Problem 1.1 in the pure differential privacy and i.i.d. mechanisms setting. That is, $\delta = \Delta = 0$ and $p = p_i = \Pr[M_i(\mathcal{D}) = 1], p' = p'_i = \Pr[M_i(\mathcal{D}') = 1], \forall i \in [K]$. Our goal is to search for a good noise function γ such that: 1) DaRRM_γ is $m\epsilon$-DP, and 2) DaRRM_γ achieves higher utility than that of the baselines (see Section 3) under a fixed privacy loss. Our main finding of such a γ function is presented in Theorem 4.1, which states given a privacy allowance $m \in [K]$, one can indeed output the majority of $2m-1$ subsampled mechanisms, instead of just m as indicated by simple composition. Later, we formally verify in Lemma B.11, Section B.3 that taking the majority of more mechanisms strictly increases the utility.

To start, by Lemma 3.4, for any noise function γ, γ satisfying goal 1) is equivalent to satisfying

$$f(p, p'; \gamma) \leq e^\epsilon - 1 \tag{54}$$

where $f(p, p'; \gamma) = \sum_{l=0}^{\frac{K-1}{2}} (e^{m\epsilon}\alpha'_l - \alpha_l) \cdot \gamma(l) + \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l) \cdot \gamma(l)$ refers to the privacy cost objective (see Lemma 3.4) in the i.i.d. mechanisms setting, and recall $\alpha_l = \Pr[\mathcal{L}(\mathcal{D}) = l]$ and $\alpha'_l = \Pr[\mathcal{L}(\mathcal{D}') = l]$, $\forall l \in \{0, 1, \ldots, K\}$. Notice in this setting, $\mathcal{L}(\mathcal{D}) \sim \text{Binomial}(p)$, and $\mathcal{L}(\mathcal{D}') \sim \text{Binomial}(p')$.

Monotonicity Assumption. For analysis, we restrict our search for a γ function with good utility to the class with a mild monotonicity assumption: $\gamma(l) \geq \gamma(l+1), \forall l \leq \frac{K-1}{2}$ and $\gamma(l) \leq \gamma(l+1), \forall l \geq \frac{K+1}{2}$. This matches our intuition that as $\mathcal{L}(\mathcal{D}) = \sum_{i=1}^{K} S_i$, i.e., the number of mechanisms outputting 1, approaches 0 or K, there is a clearer majority and so not much noise is needed to ensure privacy, which implies a larger value of γ.

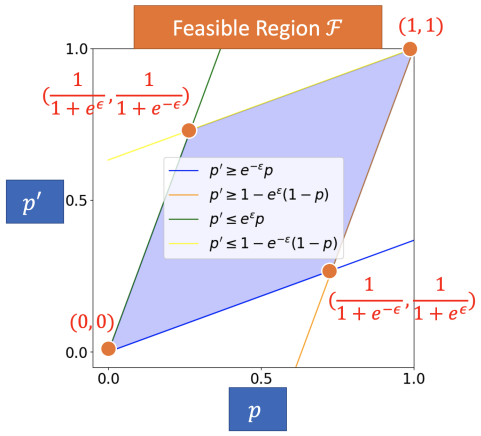

Figure 4: The feasible region $\mathcal{F}$ is plotted as the blue area. The four boundaries are implied by p, p' satisfying ϵ-differential privacy.

Roadmap of Proof of Theorem 4.1. Since γ needs to enable Eq. 54 to be satisfied for all $p, p' \in [0, 1]$, we begin by showing characteristics of **the worst case probabilities**, i.e., $(p^*, p'^*) = \arg\max_{(p,p')} f(p, p'; \gamma)$, given any $\gamma : \{0, 1, \ldots, K\} \to [0, 1]$ that is symmetric around $\frac{K}{2}$ and that satisfies the above monotonicity assumption, in Lemma B.1, Section B.1. We call (p^*, p'^*) the worst case probabilities, since they incur the largest privacy loss. Later in Section B.2, we present the main proof of Theorem 4.1, where we focus on searching for a good γ that enables $f(p^*, p'^*; \gamma) \leq e^\epsilon - 1$, based on the characteristics of (p^*, p'^*) in Lemma B.1, to ensure DaRRM_γ is $m\epsilon$-differentially private.

B.1 Characterizing the Worst Case Probabilities

First, note (p, p') are close to each other and lie in a feasible region $\mathcal{F}$, due to each mechanism M_i being ϵ-differentially private; and so does (p^*, p'^*). The feasible region, as illustrated in Figure 4, is bounded by (a) $p' \leq e^\epsilon p$ (b) $p \leq e^\epsilon p'$ (c) $1 - p' \leq e^\epsilon(1-p)$, and (d) $1 - p \leq e^\epsilon(1-p')$, where the four boundaries are derived from the definition of differential privacy. Therefore, we only need to search for $(p^*, p'^*) = \arg\max_{(p,p') \in \mathcal{F}} f(p, p'; \gamma)$.

Next, we show that given γ satisfying certain conditions, (p^*, p'^*) can only be on two of the four boundaries of $\mathcal{F}$ in Lemma B.1 — that is, either $p^* = e^\epsilon p'$, i.e., on the blue line in Figure 4, or $1 - p'^* = e^\epsilon(1 - p^*)$, i.e., on the orange line in Figure 4.

Lemma B.1 (Characteristics of worst case probabilities)**.** *For any noise function $\gamma : \{0, 1, \ldots, K\} \to [0, 1]$ that is 1) symmetric around $\frac{K}{2}$, 2) satisfies the monotonicity assumption, and 3) $\gamma(\frac{K-1}{2}) > 0$ and $\gamma(\frac{K+1}{2}) > 0$, the worst case probabilities given γ, $(p^*, p'^*) = \arg\max_{(p,p') \in \mathcal{F}} f(p, p'; \gamma)$, must satisfy one of the following two equalities:*

$$p^* = e^\epsilon p'^*, \qquad \forall p^* \in [0, \frac{1}{e^{-\epsilon}+1}], p'^* \in [0, \frac{1}{1+e^\epsilon}]$$

$$or \quad 1-p'^* = e^\epsilon(1-p^*), \qquad \forall p^* \in [\frac{1}{1+e^{-\epsilon}}, 1], p'^* \in [\frac{1}{1+e^\epsilon}, 1]$$

To show Lemma B.1, we first show in Lemma B.2 that the search of (p^*, p'^*) can be refined to one of the four boundaries of $\mathcal{F}$, via a careful gradient analysis of $f(p, p'; \gamma)$ in $\mathcal{F}$, and then show in Lemma B.3 that the search of (p^*, p'^*) can be further refined to two of the four boundaries, due to symmetry of p, p'. Lemma B.1 directly follows from the two.

Lemma B.2. *For any noise function $\gamma : \{0, 1, \ldots, K\} \to [0, 1]$ that is 1) symmetric around $\frac{K}{2}$, 2) satisfies the monotonicity assumption, and 3) $\gamma(\frac{K-1}{2}) > 0$ and $\gamma(\frac{K+1}{2}) > 0$, the worst case probabilities given γ, $(p^*, p'^*) = \arg\max_{(p,p')\in\mathcal{F}} f(p, p'; \gamma)$, must satisfy one of the following four equalities:*

$$\begin{aligned}
&p'^* = e^\epsilon p^*, && \forall p^* \in [0, \frac{1}{1+e^\epsilon}], p'^* \in [0, \frac{1}{1+e^{-\epsilon}}] \\
&p^* = e^\epsilon p'^*, && \forall p^* \in [0, \frac{1}{e^{-\epsilon}+1}], p'^* \in [0, \frac{1}{1+e^\epsilon}] \\
&1-p^* = e^\epsilon(1-p'^*), && \forall p^* \in [\frac{1}{1+e^\epsilon}, 1], p'^* \in [\frac{1}{1+e^{-\epsilon}}, 1] \\
&1-p'^* = e^\epsilon(1-p^*), && \forall p^* \in [\frac{1}{1+e^{-\epsilon}}, 1], p'^* \in [\frac{1}{1+e^\epsilon}, 1]
\end{aligned}$$

Proof of Lemma B.2. Recall the privacy cost objective (as defined in Lemma 3.4) is now

$$f(p, p'; \gamma) = \sum_{l=0}^{\frac{K-1}{2}} (e^{m\epsilon}\alpha'_l - \alpha_l) \cdot \gamma(l) + \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l) \cdot \gamma(l)$$

where $\alpha_l = \Pr[\mathcal{L}(\mathcal{D}) = l]$ and $\alpha'_l = \Pr[\mathcal{L}(\mathcal{D}') = l]$, $\forall l \in \{0, 1, \ldots, K\}$. Since $\mathcal{L}(\mathcal{D}) \sim \text{Binomial}(p)$ and $\mathcal{L}(\mathcal{D}') \sim \text{Binomial}(p')$ in the i.i.d. mechanisms setting, and using the pmf of the Binomial distribution, f can be written as

$$f(p, p'; \gamma) = \sum_{l=0}^{\frac{K-1}{2}} (e^{m\epsilon}\binom{K}{l}p'^l(1-p')^{K-l} - \binom{K}{l}p^l(1-p)^{K-l}) \cdot \gamma(l) + \sum_{l=\frac{K+1}{2}}^{K} (\binom{K}{l}p^l(1-p)^{K-l} - e^{m\epsilon}\binom{K}{l}p'^l(1-p')^{K-l})$$

The gradients w.r.t. p and p' are

$$\begin{aligned}
\nabla_p f(p, p'; \gamma) = &\underbrace{\sum_{l=0}^{\frac{K-1}{2}} -\binom{K}{l}\gamma(l) \cdot (lp^{l-1}(1-p)^{K-l} - p^l(K-l)(1-p)^{K-l-1})}_{:=A} \\
&\underbrace{+ \sum_{l=\frac{K+1}{2}}^{K} \binom{K}{l}\gamma(l) \cdot (lp^{l-1}(1-p)^{K-l} - p^l(K-l)(1-p)^{K-l-1})}_{:=B}
\end{aligned} \tag{55}$$

and

$$\begin{aligned}
\nabla_{p'} f(p, p'; \gamma) = &\sum_{l=0}^{\frac{K-1}{2}} e^{m\epsilon}\binom{K}{l}\gamma(l) \cdot (lp'^{l-1}(1-p')^{K-l} - p'^l(K-l)(1-p')^{K-l-1}) \\
&+ \sum_{l=\frac{K+1}{2}}^{K} -e^{m\epsilon}\binom{K}{l}\gamma(l) \cdot (lp'^{l-1}(1-p')^{K-l} - p'^l(K-l)(1-p')^{K-l-1})
\end{aligned} \tag{56}$$

We show in the following $\forall p \in (0,1)$, $\nabla_p f(p,p';\gamma) > 0$ and $\nabla_{p'} f(p,p';\gamma) < 0$. This implies there is no local maximum inside $\mathcal{F}$, and so $(p^*, p'^*) = \arg\max_{p,p'} f(p,p';\gamma)$ must be on one of the four boundaries of $\mathcal{F}$. Also, if $p = 0$, then $p' = 0$, and $(0,0)$ is a corner point at the intersection of two boundaries. Similarly, if $p = 1$, then $p' = 1$, and $(1,1)$ is also a corner point. This concludes $\forall p \in [0,1]$, $(p^*, p'^*) = \arg\max_{p,p'} f(p,p';\gamma)$ must be on one of the four boundaries of $\mathcal{F}$.

To show $\nabla_p f(p,p';\gamma) > 0$ for $p \in (0,1)$, we write $\nabla_p f(p,p';\gamma) = A + B$ as in Eq. 55, and show that $A > 0$ and $B > 0$.

To show $A > 0$, first note

$$A := \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K}{l} \cdot (p^l (K-l)(1-p)^{K-l-1} - l p^{l-1}(1-p)^{K-l}) > 0 \tag{57}$$

$$\iff \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K}{l} \cdot p^l (K-l)(1-p)^{K-l-1} > \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K}{l} \cdot l p^{l-1}(1-p)^{K-l}) \tag{58}$$

$$\iff \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K-1}{l} \frac{K}{K-l} \cdot p^l (K-l)(1-p)^{K-l-1} > \sum_{l=1}^{\frac{K-1}{2}} \gamma(l) \binom{K-1}{l-1} \frac{K}{l} \cdot l p^{l-1}(1-p)^{K-l} \tag{59}$$

$$\iff K \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K-1}{l} p^l (1-p)^{K-l-1} > K \sum_{l=1}^{\frac{K-1}{2}} \gamma(l) \binom{K-1}{l-1} p^{l-1}(1-p)^{K-l} \tag{60}$$

$$\iff \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K-1}{l} p^l (1-p)^{K-l-1} > \sum_{l=0}^{\frac{K-1}{2}-1} \gamma(l+1) \binom{K-1}{l} p^l (1-p)^{K-l-1} \tag{61}$$

Since $\forall l \leq \frac{K-1}{2}$, $\gamma(l) \geq \gamma(l+1)$ and $p \in (0,1)$, there is for $l \in \{0, \ldots, \frac{K-1}{2} - 1\}$,

$$\gamma(l) \binom{K-1}{l} p^l (1-p)^{K-l-1} \geq \gamma(l+1) \binom{K-1}{l} p^l (1-p)^{K-l-1} \tag{62}$$

Furthermore, since $\gamma(\frac{K-1}{2}) > 0$ and $p \in (0,1)$,

$$\gamma(\frac{K-1}{2}) \binom{K-1}{\frac{K-1}{2}} p^{\frac{K-1}{2}} (1-p)^{\frac{K-1}{2}} > 0 \tag{63}$$

Eq. 62 and Eq. 63 combined implies

$$\gamma(\frac{K-1}{2}) \binom{K-1}{\frac{K-1}{2}} p^{\frac{K-1}{2}} (1-p)^{\frac{K-1}{2}} + \sum_{l=0}^{\frac{K-1}{2}-1} \gamma(l) \binom{K-1}{l} p^l (1-p)^{K-l-1} > \sum_{l=0}^{\frac{K-1}{2}-1} \gamma(l+1) \binom{K-1}{l} p^l (1-p)^{K-l-1} \tag{64}$$

and hence, Eq. 61 holds. This further implies $A > 0$.

Next, to show $B > 0$, note that

$$B := \sum_{l=\frac{K+1}{2}}^{K} \binom{K}{l} \gamma(l) \cdot (l p^{l-1}(1-p)^{K-l} - p^l (K-l)(1-p)^{K-l-1}) > 0 \tag{65}$$

$$\iff \sum_{l=\frac{K+1}{2}}^{K} \binom{K}{l} \gamma(l) \cdot l p^{l-1}(1-p)^{K-l} > \sum_{l=\frac{K+1}{2}}^{K} \binom{K}{l} p^l (K-l)(1-p)^{K-l-1} \tag{66}$$

$$\iff \sum_{l=\frac{K+1}{2}}^{K} \gamma(l)\binom{K-1}{l-1}\frac{K}{l}\cdot lp^{l-1}(1-p)^{K-l} \tag{67}$$

$$> \sum_{l=\frac{K+1}{2}}^{K-1} \gamma(l)\binom{K-1}{l}\frac{K}{K-l}\cdot p^{l}(K-l)(1-p)^{K-l-1}$$

$$\iff K\sum_{l=\frac{K+1}{2}}^{K} \gamma(l)\binom{K-1}{l-1}\cdot p^{l-1}(1-p)^{K-l} \tag{68}$$

$$> K\sum_{l=\frac{K+1}{2}}^{K-1} \gamma(l)\binom{K-1}{l}\cdot p^{l}(1-p)^{K-l-1}$$

$$\iff \sum_{l=\frac{K+1}{2}}^{K} \gamma(l)\binom{K-1}{l-1}\cdot p^{l-1}(1-p)^{K-l} > \sum_{l=\frac{K+1}{2}+1}^{K} \gamma(l-1)\binom{K-1}{l-1}\cdot p^{l-1}(1-p)^{K-l} \tag{69}$$

Since $\forall l \geq \frac{K+1}{2}$, $\gamma(l) \geq \gamma(l-1)$ and $p \in (0,1)$, there is for $l \in \{\frac{K+1}{2}+1, \ldots, K\}$,

$$\gamma(l)\binom{K-1}{l-1}p^{l-1}(1-p)^{K-l} \geq \gamma(l-1)\binom{K-1}{l-1}p^{l-1}(1-p)^{K-l} \tag{70}$$

Furthermore, since $\gamma(\frac{K+1}{2}) > 0$ and $p \in (0,1)$,

$$\gamma(\frac{K+1}{2})\binom{K-1}{\frac{K-1}{2}}p^{\frac{K-1}{2}}(1-p)^{\frac{K-1}{2}} > 0 \tag{71}$$

Eq. 70 and Eq. 71 combined implies

$$\gamma(\frac{K+1}{2})\binom{K-1}{\frac{K-1}{2}}p^{\frac{K-1}{2}}(1-p)^{\frac{K-1}{2}} + \sum_{l=\frac{K+1}{2}+1}^{K} \gamma(l)\binom{K-1}{l-1}\cdot p^{l-1}(1-p)^{K-l} > \sum_{l=\frac{K+1}{2}+1}^{K} \gamma(l-1)\binom{K-1}{l-1}\cdot p^{l-1}(1-p)^{K-l} \tag{72}$$

and hence Eq. 69 holds. This further implies $B > 0$.

Following Eq.55, for $p \in (0,1)$ and γ satisfying the three assumptions,

$$\nabla_p f(p, p'; \gamma) = A + B > 0 \tag{73}$$

Following similar techniques, one can show for $p \in (0,1)$ and γ satisfying the three conditions,

$$\nabla_{p'} f(p, p'; \gamma) < 0 \tag{74}$$

This implies there is no local minima or local maxima inside the feasible region $\mathcal{F}$. Also recall $(p, p') \in \{(0,0),(1,1)\}$ are two special cases where (p, p') is at the intersection of two boundaries. Hence, we conclude the worst case probability $(p^*, p'^*) = \arg\max_{p,p' \in \mathcal{F}} f(p, p'; \gamma)$ is on one of the four boundaries of $\mathcal{F}$ — that is, (p^*, p'^*) satisfy one of the following:

$$p'^* = e^{\epsilon}p^*, \qquad \forall p \in [0, \frac{1}{1+e^{\epsilon}}], p' \in [0, \frac{1}{1+e^{-\epsilon}}]$$

$$p^* = e^{\epsilon}p'^*, \qquad \forall p \in [0, \frac{1}{e^{-\epsilon}+1}], p' \in [0, \frac{1}{1+e^{\epsilon}}]$$

$$1-p^* = e^{\epsilon}(1-p'^*), \qquad \forall p \in [\frac{1}{1+e^{\epsilon}}, 1], p' \in [\frac{1}{1+e^{-\epsilon}}, 1]$$

$$1-p'^* = e^{\epsilon}(1-p^*), \qquad \forall p \in [\frac{1}{1+e^{-\epsilon}}, 1], p' \in [\frac{1}{1+e^{\epsilon}}, 1]$$

□

Lemma B.3. *For any noise function $\gamma : \{0, 1, \ldots, K\} \to [0, 1]$ function that is 1) symmetric around $\frac{K}{2}$ and 2) satisfies the monotonicity assumption, the privacy cost objective $f(p, p'; \gamma)$ is maximized when $p \geq p'$.*

Proof of Lemma B.3. Following Eq. 33 and Eq. 38 in the proof of Lemma 3.4, and that $\delta = 0$,

$$\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1] \leq e^{m\epsilon} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1] \tag{75}$$

$$\iff \underbrace{\sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l) - \sum_{l=0}^{\frac{K-1}{2}} (\alpha_l - e^{m\epsilon}\alpha'_l)\gamma(l)}_{=f(p,p';\gamma)} \leq e^{m\epsilon} - 1 \tag{76}$$

where $\alpha_l = \Pr[\mathcal{L}(\mathcal{D}) = l]$ and $\alpha'_l = \Pr[\mathcal{L}(\mathcal{D}') = l]$, $\forall l \in \{0, 1, \ldots, K\}$. This implies

$$f(p, p'; \gamma) = \frac{\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1]}{\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1]} - 1 \tag{77}$$

Hence, $f(p, p'; \gamma)$ is maximized when $\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1] \geq \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1]$.

$$\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1] = \sum_{l=0}^{K} \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1 \mid \mathcal{L}(\mathcal{D}) = 1] \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] \tag{78}$$

$$= \sum_{l=0}^{K} \left(\gamma(l) \cdot \mathbb{I}\{l \geq \frac{K}{2}\} + \frac{1}{2}(1 - \gamma(l))\right) \cdot \Pr[\mathcal{L}(\mathcal{D}) = l] \tag{79}$$

$$= \sum_{l=0}^{\frac{K-1}{2}} \frac{1}{2}(1 - \gamma(l)) \cdot \alpha_l + \sum_{l=\frac{K+1}{2}}^{K} \left(\gamma(l) + \frac{1}{2}(1 - \gamma(l))\right) \cdot \alpha_l \tag{80}$$

$$= \frac{1}{2} \sum_{l=\frac{K+1}{2}}^{K} \gamma(l) \binom{K}{l} p^l (1-p)^{K-l} - \frac{1}{2} \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K}{l} p^l (1-p)^{K-l-1} + \frac{1}{2} \tag{81}$$

where the last line follows from the observation that in the i.i.d. mechanisms setting, $\mathcal{L}(\mathcal{D}) \sim \text{Binomial}(p)$ and α_l is hence the pmf of the Binomial distribution at l.

Similarly,

$$\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1] = \frac{1}{2} \sum_{l=\frac{K+1}{2}}^{K} \gamma(l) \binom{K}{l} p'^l (1-p')^{K-l} - \frac{1}{2} \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K}{l} p'^l (1-p')^{K-l-1} + \frac{1}{2} \tag{82}$$

Now define the objective

$$h(\beta) = \frac{1}{2} \sum_{l=\frac{K+1}{2}}^{K} \gamma(l) \binom{K}{l} \beta^l (1-\beta)^{K-l} - \frac{1}{2} \sum_{l=0}^{\frac{K-1}{2}} \gamma(l) \binom{K}{l} \beta^l (1-\beta)^{K-l-1} + \frac{1}{2} \tag{83}$$

for $\beta \in [0, 1]$ and it follows that $\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1] = h(p)$ and $\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1] = h(p')$. We now analyze the monotonicity of $h(\beta)$ in β.

For ease of presentation, define $g(l) := \begin{cases} -\frac{1}{2}\gamma(l) & \forall l \leq \frac{K}{2} \\ \frac{1}{2}\gamma(l) & \forall l \geq \frac{K}{2} \end{cases}$. Since $\gamma(l) \geq \gamma(l+1), \forall l \leq \frac{K}{2}$ and $\gamma(l+1) \geq \gamma(l), \forall l \geq \frac{K}{2}$, there is $g(l+1) \geq g(l), \forall l \in \{0, \ldots, K\}$. And replacing $\gamma(l)$ with $g(l)$ in Eq. 83,

$$h(\beta) = \sum_{l=0}^{K} g(l) \binom{K}{l} \beta^l (1-\beta)^{K-l} \tag{84}$$

$$\nabla_\beta h(\beta) = \sum_{l=0}^{K} g(l)\binom{K}{l}\Big(l\beta^{l-1}(1-\beta)^{K-l} - (K-l)\beta^l(1-\beta)^{K-l-1}\Big) \tag{85}$$

$$= \sum_{l=1}^{K} g(l)\binom{K-1}{l-1}\frac{K}{l}l\beta^{l-1}(1-\beta)^{K-l} - \sum_{l=0}^{K-1}\binom{K-1}{l}\frac{K}{K-l}(K-l)\beta^l(1-\beta)^{K-l-1} \tag{86}$$

$$= K\sum_{l=1}^{K}\binom{K-1}{l-1}\beta^{l-1}(1-\beta)^{K-l} - K\sum_{l=0}^{K-1}\binom{K-1}{l}\beta^l(1-\beta)^{K-l-1} \tag{87}$$

$$= K\sum_{l=0}^{K-1} g(l+1)\binom{K-1}{l}\beta^l(1-\beta)^{K-l-1} - K\sum_{l=0}^{K-1} g(l)\binom{K-1}{l}\beta^l(1-\beta)^{K-l-1} \tag{88}$$

$$= K\sum_{l=0}^{K-1}\Big(g(l+1)-g(l)\Big)\binom{K-1}{l}\beta^l(1-\beta)^{K-l-1} \tag{89}$$

Since $g(l+1) \geq g(l)$ and $\binom{K-1}{l}\beta^l(1-\beta)^{K-l-1} \geq 0$, $\nabla_\beta h(\beta) \geq 0$. This implies $h(\beta)$ is monotonically non-decreasing in β and hence,

$$\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}) = 1] \geq \Pr[\mathsf{DaRRM}_\gamma(\mathcal{D}') = 1] \iff p \geq p' \tag{90}$$

Therefore, $f(p, p'; \gamma)$ is maximzied when $p \geq p'$. □

B.2 Proof of Privacy Amplification (Theorem 4.1)

Theorem B.4 (Restatement of Theorem 4.1)**.** *Consider using DaRRM (Algorithm 1) to solve Problem 1.1, with i.i.d. mechanisms $\{M_i\}_{i=1}^K$, i.e., $p_i = p$, $p_i' = p'$, $\forall i \in [K]$, the privacy allowance $m \in [K]$ and $\delta = \Delta = 0$. Let the noise function $\gamma : \{0, 1, \ldots, K\} \to [0, 1]$ be that:*
if $m \geq \frac{K+1}{2}$,

$$\gamma(l) = 1$$

and if $m \leq \frac{K-1}{2}$,

$$\gamma(l) = \begin{cases} 1 - 2h(l) & \forall l \leq \frac{K-1}{2} \\ 2h(l) - 1 & \forall l \geq \frac{K+1}{2} \end{cases}$$

where $h(l) = \sum_{i=m}^{2m-1} \frac{\binom{l}{i}\binom{K-l}{2m-1-i}}{\binom{K}{2m-1}}$, then DaRRM_γ is $m\epsilon$-differentially private.

Roadmap. Theorem 4.1 consists of two parts: γ under a large privacy allowance $m \geq \frac{K+1}{2}$ and γ under a small privacy allowance $m \leq \frac{K-1}{2}$. We first show in Lemma B.5, Section B.2.1 that if $m \geq \frac{K+1}{2}$, setting $\gamma = 1$ suffices to ensure DaRRM_γ to be $m\epsilon$-differentially private, and hence one can always output the true majority of K mechanisms. In contrast, simple composition indicates only when $m = K$ can one output the true majority of K mechanisms. Next, we show in Lemma B.10, Section B.2.2 that if $m \leq \frac{K-1}{2}$, one can set γ to be γ_{DSub}, which corresponds to outputting the majority of $2m-1$ subsampled mechanisms (and hence the name "Double Subsampling", or DSub). In contrast, simple compositon indicates one can only output the majority of m subsampled mechanisms to make sure the output is $m\epsilon$-differentially private. **Theorem 4.1 follows directly from combining Lemma B.5 and Lemma B.10.**

B.2.1 Privacy Amplification Under A Large Privacy Allowance $m \geq \frac{K+1}{2}$

The proof of Lemma B.5 is straightforward. We show that given the constant $\gamma_{max}(l) = 1$, if $m \geq \frac{K+1}{2}$, the worst case probabilities are $(p^*, p'^*) = \arg\max_{(p,p')\in\mathcal{F}} f(p, p'; \gamma_{max}) = (0, 0)$ and notice that $f(0, 0; \gamma_{max}) = e^{m\epsilon} - 1$, which satisfies the condition in Lemma 3.4. Hence, $\mathsf{DaRRM}_{\gamma_{max}}$ is $m\epsilon$-differentially private.

Lemma B.5 (Privacy amplification, $m \geq \frac{K+1}{2}$)**.** *Consider using DaRRM (Algorithm 1) to solve Problem 1.1, with i.i.d. mechanisms $\{M_i\}_{i=1}^K$, i.e., $p_i = p$, $p_i' = p'$, $\forall i \in [K]$, the privacy allowance $m \geq \frac{K+1}{2}, m \in \mathbb{Z}$ and $\delta = \Delta = 0$. Let the noise function be the constant $\gamma_{max}(l) = 1, \forall l \in \{0, 1, \ldots, K\}$. Then, $\mathsf{DaRRM}_{\gamma_{max}}$ is $m\epsilon$-differentially private.*

Proof of Lemma B.5. First, notice $\gamma_{max}(l) = 1, \forall l \in \{0, 1, \ldots, K\}$ is: 1) symmetric around $\frac{K}{2}$, 2) satisfies the monotonicity assumption, and 3) $\gamma_{max}(\frac{K-1}{2}) > 0$ and $\gamma_{max}(\frac{K+1}{2}) > 0$. Therefore, by Lemma B.1, the worst case probabilities given γ_{max}, i.e., $(p^*, p'^*) = \arg\max_{(p,p')\in\mathcal{F}} f(p, p'; \gamma_{max})$, are on one of the two boundaries of $\mathcal{F}$, satisfying

$$\begin{aligned} &p^* = e^\epsilon p'^*, && \forall p^* \in [0, \frac{1}{e^{-\epsilon}+1}], p'^* \in [0, \frac{1}{1+e^\epsilon}] \\ \text{or} \quad &1 - p'^* = e^\epsilon (1 - p^*), && \forall p^* \in [\frac{1}{1+e^{-\epsilon}}, 1], p'^* \in [\frac{1}{1+e^\epsilon}, 1] \end{aligned}$$

We now find the local maximums on the two possible boundaries, i.e.,

$$(p^*_{local}, p'^*_{local}) = \underset{(p,p'): p = e^\epsilon p', p \in [0, \frac{1}{e^{-\epsilon}+1}]}{\arg\max} f(p, p'; \gamma_{max})$$

and

$$(p^*_{local}, p'^*_{local}) = \underset{(p,p'): 1-p' = e^\epsilon (1-p), p \in [\frac{1}{1+e^{-\epsilon}}, 1]}{\arg\max} f(p, p'; \gamma_{max})$$

separately.

Part I: Local worst case probabilities on the boundary $p = e^\epsilon p'$.

Plugging $p = e^\epsilon p'$ into the privacy cost objective $f(p, p'; \gamma_{max})$, one gets

$$\begin{aligned} f(p'; \gamma_{max}) = &\sum_{l=0}^{\frac{K-1}{2}} (e^{m\epsilon} \binom{K}{l} p'^l (1-p')^{K-l} - \binom{K}{l} (e^\epsilon p')^l (1 - e^\epsilon p')^{K-l}) \\ &+ \sum_{l=\frac{K+1}{2}}^{K} (\binom{K}{l} (e^\epsilon p')^l (1 - e^\epsilon p')^{K-l} - e^{m\epsilon} \binom{K}{l} p'^l (1-p')^{K-l}) \end{aligned} \tag{91}$$

The gradient w.r.t. p' is

$$\begin{aligned} \nabla_{p'} f(p'; \gamma_{max}) = &\sum_{l=0}^{\frac{K-1}{2}} \Big(e^{m\epsilon} \binom{K}{l} (l p'^{l-1} (1-p')^{K-l} - p'^l (K-l)(1-p')^{K-l-1}) \\ &- e^\epsilon \binom{K}{l} (l (e^\epsilon p')^{l-1} (1 - e^\epsilon p')^{K-l} - e^{\epsilon l} p'^l (K-l)(1-e^\epsilon p')^{K-l-1}) \Big) \\ &+ \sum_{l=\frac{K+1}{2}}^{K} \Big(e^\epsilon \binom{K}{l} (l (e^\epsilon p')^{l-1} (1 - e^\epsilon p')^{K-l} - e^{\epsilon l} p'^l (K-l)(1-e^\epsilon p')^{K-l-1}) \\ &- e^{m\epsilon} \binom{K}{l} (l p'^{l-1} (1-p')^{K-l} - p'^l (K-l)(1-p')^{K-l-1}) \Big) \end{aligned} \tag{92}$$

$$\begin{aligned} = &-K \sum_{l=0}^{\frac{K-1}{2}} e^{m\epsilon} \binom{K-1}{l} p'^l (1-p')^{K-l-1} + K \sum_{l=\frac{K+1}{2}}^{K-1} e^{m\epsilon} \binom{K-1}{l} p'^l (1-p')^{K-l-1} \\ &+ K \sum_{l=0}^{\frac{K-1}{2}} e^\epsilon \binom{K-1}{l} (\epsilon p')^\epsilon (1 - e^\epsilon p')^{K-l-1} - K \sum_{l=\frac{K+1}{2}}^{K-1} e^\epsilon \binom{K-1}{l} (e^\epsilon p')^l (1 - e^\epsilon p')^{K-l-1} \end{aligned} \tag{93}$$

$$
\begin{aligned}
&+ K \sum_{l=0}^{\frac{K-1}{2}-1} e^{m\epsilon}\binom{K-1}{l} p'^{l}(1-p')^{K-l-1} - K \sum_{l=\frac{K-1}{2}}^{K-1} e^{m\epsilon}\binom{K-1}{l} p'^{l}(1-p')^{K-l-1} \\
&- K \sum_{l=0}^{\frac{K-1}{2}-1} e^{\epsilon}\binom{K-1}{l} (e^{\epsilon}p')^{l}(1-e^{\epsilon}p')^{K-l-1} + K \sum_{l=\frac{K-1}{2}}^{K-1} e^{\epsilon}\binom{K-1}{l} (e^{\epsilon}p')^{l}(1-e^{\epsilon}p')^{K-l-1} \\
&= \underbrace{-2K\, e^{m\epsilon}\binom{K-1}{\frac{K-1}{2}} p'^{\frac{K-1}{2}}(1-p')^{\frac{K-1}{2}}}_{:=A} + \underbrace{2K\, e^{\epsilon}\binom{K-1}{\frac{K-1}{2}} (e^{\epsilon}p')^{\frac{K-1}{2}}(1-e^{\epsilon}p')^{\frac{K-1}{2}}}_{:=B}
\end{aligned} \tag{94}
$$

Notice that

$$
\frac{A}{B} = \frac{e^{m\epsilon}\binom{K-1}{\frac{K-1}{2}} p'^{\frac{K-1}{2}}(1-p')^{\frac{K-1}{2}}}{e^{\epsilon}\binom{K-1}{\frac{K-1}{2}} (e^{\epsilon}p')^{\frac{K-1}{2}}(1-e^{\epsilon}p')^{\frac{K-1}{2}}} = \frac{e^{m\epsilon}}{e^{\frac{K+1}{2}\epsilon}} \cdot \left(\frac{1-p'}{1-e^{\epsilon}p'}\right)^{\frac{K-1}{2}} \tag{95}
$$

Since $\frac{1-p'}{1-e^{\epsilon}p'} \geq 1$ and $m \geq \frac{K+1}{2}$, $\frac{A}{B} \geq 1$. This implies $\nabla_{p'} f(p'; \gamma_{max}) \leq 0$. Hence, $f(p'; \gamma_{max})$ is monotonically non-increasing on the boundary, for $p' \in [0, \frac{1}{1+e^{\epsilon}}]$.

Therefore, $\arg\max_{p':p' \in [0, \frac{1}{1+e^{\epsilon}}]} f(p'; \gamma_{max}) = 0$. Since $p = e^{\epsilon}p'$, $p' = 0$ implies $p = 0$.

Hence,

$$
(p^*_{local}, p'^*_{local}) = \underset{(p,p'):p=e^{\epsilon}p', p \in [0, \frac{1}{e^{-\epsilon}+1}]}{\arg\max} f(p, p'; \gamma_{max}) = (0, 0)
$$

and

$$
\max_{(p,p'):p=e^{\epsilon}p', p \in [0, \frac{1}{e^{-\epsilon}+1}]} f(p, p'; \gamma_{max}) = f(0, 0; \gamma_{max}) = e^{m\epsilon} - 1
$$

Part II: Local worst case probabilities on the boundary $1 - p' = e^{\epsilon}(1-p)$.

For simplicity, let $q = 1 - p$ and $q' = 1 - p'$. Note on this boundary $p \in [\frac{1}{1+e^{-\epsilon}}, 1]$ and $p' \in [\frac{1}{1+e^{\epsilon}}, 1]$, and hence, $q \in [0, \frac{1}{1+e^{\epsilon}}]$ and $q' \in [0, \frac{1}{1+e^{-\epsilon}}]$.

Plugging q and q' into the privacy cost objective $f(p, p'; \gamma_{max})$, one gets a new objective in q, q' as

$$
\begin{aligned}
f(q, q'; \gamma_{max}) &= \sum_{l=0}^{\frac{K-1}{2}} \left(e^{m\epsilon}\binom{K}{l}(1-q')^{l}q'^{K-l} - \binom{K}{l}(1-q)^{l}q^{K-l}\right) \cdot \gamma_{max}(l) \qquad (96) \\
&\quad + \sum_{l=\frac{K+1}{2}}^{K} \left(\binom{K}{l}(1-q)^{l}q^{K-l} - e^{m\epsilon}\binom{K}{l}(1-q')^{l}q'^{K-l}\right) \cdot \gamma_{max}(l) \\
&= \sum_{l=0}^{\frac{K-1}{2}} \left(e^{m\epsilon}\binom{K}{l}(1-q')^{l}q'^{K-l} - \binom{K}{l}(1-q)^{l}q^{K-l}\right) \qquad (97) \\
&\quad + \sum_{l=\frac{K+1}{2}}^{K} \left(\binom{K}{l}(1-q)^{l}q^{K-l} - e^{m\epsilon}\binom{K}{l}(1-q')^{l}q'^{K-l}\right)
\end{aligned}
$$

Since on this boundary, $1 - p' = e^{\epsilon}(1-p)$, writing this in q, q', this becomes $q' = e^{\epsilon}q$. Plugging $q' = e^{\epsilon}q$ into $f(q, q'; \gamma_{max})$, one gets

$$
f(q; \gamma_{max}) = \sum_{l=0}^{\frac{K-1}{2}} \left(e^{m\epsilon}\binom{K}{l}(1-e^{\epsilon}q)^{l}(e^{\epsilon}q)^{K-l} - \binom{K}{l}(1-q)^{l}q^{K-l}\right) \tag{98}
$$

$$+\sum_{l=\frac{K+1}{2}}^{K}\Big(\binom{K}{l}(1-q)^l q^{K-l}-e^{m\epsilon}\binom{K}{l}(1-e^\epsilon q)^l(e^\epsilon q)^{K-l}\Big)$$

The gradient w.r.t. q is

$$\begin{aligned}
\nabla_q f(q) &= \sum_{l=0}^{\frac{K-1}{2}}\Big(e^{m\epsilon}\binom{K}{l}\Big((-e^\epsilon)l(1-e^\epsilon q)^{l-1}(e^\epsilon q)^{K-l}+e^\epsilon(K-l)(1-e^\epsilon q)^l(e^\epsilon q)^{K-l-1}\Big) \qquad (99)\\
&\quad-\binom{K}{l}\Big(-l(1-q)^{l-1}q^{K-l}+(K-l)(1-q)^l q^{K-l-1}\Big)\Big)\\
&\quad+\sum_{l=\frac{K+1}{2}}^{K}\Big(\binom{K}{l}\Big(-l(1-q)^{l-1}q^{K-l}+(K-l)(1-q)^l q^{K-l-1}\Big)\\
&\quad-e^{m\epsilon}\binom{K}{l}\Big((-e^\epsilon)l(1-e^\epsilon q)^{l-1}(e^\epsilon q)^{K-l}+e^\epsilon(K-l)(1-e^\epsilon q)^l(e^\epsilon q)^{K-l-1}\Big)\Big)\\
&= -\sum_{l=1}^{\frac{K-1}{2}}e^{(m+1)\epsilon}\binom{K-1}{l-1}\frac{K}{l}l(1-e^\epsilon q)^{l-1}(e^\epsilon q)^{K-l}+\sum_{l=0}^{\frac{K-1}{2}}e^{(m+1)\epsilon}\binom{K-1}{l}\frac{K}{K-l}(K-l)(1-e^\epsilon q)^l(e^\epsilon q)^{K-l-1} \qquad (100)\\
&\quad+\sum_{l=1}^{\frac{K-1}{2}}\binom{K-1}{l-1}\frac{K}{l}l(1-q)^{l-1}q^{K-l}-\sum_{l=0}^{\frac{K-1}{2}}\binom{K-1}{l}\frac{K}{K-l}(K-l)(1-q)^l q^{K-l-1}\\
&\quad-\sum_{l=\frac{K+1}{2}}^{K}\binom{K-1}{l-1}\frac{K}{l}l(1-q)^{l-1}q^{K-l}+\sum_{l=\frac{K+1}{2}}^{K-1}\binom{K-1}{l}\frac{K}{K-l}(K-l)(1-q)^l q^{K-l-1}\\
&\quad+\sum_{l=\frac{K+1}{2}}^{K}e^{(m+1)\epsilon}\binom{K-1}{l-1}\frac{K}{l}l(1-e^\epsilon q)^{l-1}(e^\epsilon q)^{K-l}-\sum_{l=\frac{K+1}{2}}^{K-1}e^{(m+1)\epsilon}\binom{K-1}{l}\frac{K}{K-l}(K-l)(1-e^\epsilon q)^l(e^\epsilon q)^{K-l-1}\\
&= -K\sum_{l=1}^{\frac{K-1}{2}}e^{(m+1)\epsilon}\binom{K-1}{l-1}(1-e^\epsilon q)^{l-1}(e^\epsilon q)^{K-l}+K\sum_{l=0}^{\frac{K-1}{2}}e^{(m+1)\epsilon}\binom{K-1}{l}(1-e^\epsilon q)^l(e^\epsilon q)^{K-l-1} \qquad (101)\\
&\quad+K\sum_{l=1}^{\frac{K-1}{2}}\binom{K-1}{l-1}(1-q)^{l-1}q^{K-l}-K\sum_{l=0}^{\frac{K-1}{2}}\binom{K-1}{l}(1-q)^l q^{K-l-1}\\
&\quad-K\sum_{l=\frac{K+1}{2}}^{K}\binom{K-1}{l-1}(1-q)^{l-1}q^{K-l}+K\sum_{l=\frac{K+1}{2}}^{K-1}\binom{K-1}{l}(1-q)^l q^{K-l-1}\\
&\quad+K\sum_{l=\frac{K+1}{2}}^{K}e^{(m+1)\epsilon}\binom{K-1}{l-1}(1-e^\epsilon q)^{l-1}(e^\epsilon q)^{K-l}-K\sum_{l=\frac{K+1}{2}}^{K-1}e^{(m+1)\epsilon}\binom{K-1}{l}(1-e^\epsilon q)^l(e^\epsilon q)^{K-l-1}\\
&= 2Ke^{(m+1)\epsilon}\binom{K-1}{\frac{K-1}{2}}(1-e^\epsilon q)^{\frac{K-1}{2}}(e^\epsilon q)^{\frac{K-1}{2}}-2K\binom{K-1}{\frac{K-1}{2}}(1-q)^{\frac{K-1}{2}}q^{\frac{K-1}{2}} \qquad (102)
\end{aligned}$$

Recall $q\in[0,\frac{1}{1+e^\epsilon}]$ and so $(1-e^\epsilon q)(e^\epsilon q)\geq(1-q)q$. Furthermore, since $e^{(m+1)\epsilon}\geq 1$, there is $\nabla_q f(q)\geq 0$. This implies $f(q)$ is monotonically non-decreasing in q, and so the local maximum on this boundary is

$$(q^*_{local},q'^*_{local})=\underset{(q,q'):q'=e^\epsilon q,q\in[0,\frac{1}{1+e^\epsilon}]}{\arg\max}f(q,q';\gamma_{max})=(\frac{1}{1+e^\epsilon},\frac{1}{1+e^{-\epsilon}}) \qquad (103)$$

That is,

$$(p^*_{local}, p'^*_{local}) = \underset{(p,p'):1-p'=e^\epsilon(1-p), p\in[\frac{1}{1+e^{-\epsilon}},1]}{\arg\max} f(p,p';\gamma_{max}) = (1-q^*_{local}, 1-q'^*_{local}) = (\frac{1}{1+e^{-\epsilon}}, \frac{1}{1+e^{\epsilon}}) \tag{104}$$

Part III: The global worst case probabilities.

Notice that $(\frac{1}{1+e^{-\epsilon}}, \frac{1}{1+e^{\epsilon}})$, the maximum on the second boundary $1-p' = e^\epsilon(1-p), \forall p \in [\frac{1}{1+e^{-\epsilon}}, 1]$, is indeed the minimum on the first boundary $p = e^\epsilon p', \forall p \in [0, \frac{1}{1+e^{-\epsilon}+1}]$.

Therefore, the global maximum given γ_{max} is

$$(p^*, p'^*) = \underset{(p,p')\in\mathcal{F}}{\arg\max} f(p,p';\gamma_{max}) = \underset{(p,p'):p=e^\epsilon p', p\in[0,\frac{1}{1+e^{-\epsilon}}]}{\arg\max} f(p,p';\gamma_{max}) = (0,0) \tag{105}$$

and recall that $f(0,0;\gamma_{max}) = e^{m\epsilon} - 1$.

Hence, if $m \geq \frac{K+1}{2}$, by Lemma 3.4 $\mathsf{DaRRM}_{\gamma_{max}}$ is $m\epsilon$-differentially private.

□

B.2.2 Privacy Amplification Under A Small Privacy Allowance $m \leq \frac{K-1}{2}$

The proof of Lemma B.10 is slightly more involved. First, recall by Lemma 3.1, γ_{Sub}, the noise function that makes the output of $\mathsf{DaRRM}_{\gamma_{Sub}}$ and the subsampling baseline the same, is

$$\begin{aligned}\gamma_{Sub}(l) &= \gamma_{Sub}(K=l) \\ &= \begin{cases} 1 - 2\sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} & \text{if } m \text{ is odd} \\ 1 - 2\sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} - \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}} & \text{if } m \text{ is even} \end{cases}\end{aligned}$$

for $l \in \{0,1,\ldots,K\}$, suppose the privacy allowance $m \in \mathbb{Z}$.

If we define $h(l) := \begin{cases} \sum_{j=\frac{m+1}{2}}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} & \text{if } m \text{ is odd} \\ \sum_{j=\frac{m}{2}+1}^{m} \frac{\binom{l}{j}\binom{K-l}{m-j}}{\binom{K}{m}} - \frac{\binom{l}{\frac{m}{2}}\binom{K-l}{\frac{m}{2}}}{\binom{K}{m}} & \text{if } m \text{ is even} \end{cases}$, then $\gamma_{Sub}(l)$ can be written as $\gamma_{Sub}(l) = \begin{cases} 1-2h(l) & \text{if } l \leq \frac{K-1}{2} \\ 2h(l)-1 & \text{if } l \geq \frac{K+1}{2} \end{cases}$.

This can be generalized to a broader class of γ functions — which we call the "symmetric form family" — as follows

Definition B.6. *$\gamma : \{0,1,\ldots,K\} \to [0,1]$ is a member of the "symmetric form family" if γ follows*

$$\gamma(l) = \begin{cases} 1-2h(l) & \text{if } l \leq \frac{K-1}{2} \\ 2h(l)-1 & \text{if } l \geq \frac{K+1}{2} \end{cases} \tag{106}$$

where $h : \{0,1,\ldots,K\} \to [0,1]$ and

$$h(l) + h(K-l) = 1, \quad h(l+1) \geq h(l), \quad \forall l \in \{0,1,\ldots,K\}, \quad \text{and} \quad \gamma(\frac{K-1}{2}) > 0, \gamma(\frac{K+1}{2}) > 0$$

It is easy to verify any γ function that belongs to the "symmetric form family" satisfies: 1) symmetric around $\frac{K}{2}$ and 2) the monotonicity assumption. Hence, Lemma B.1 can be invoked to find the worst case probabilities

given such γ, i.e., $(p^*, p'^*) = \arg\max_{(p,p')\in\mathcal{F}} f(p, p'; \gamma)$, which in turn gives us the guarantee of DaRRM_γ being $m\epsilon$-differentially private.

Roadmap. In this section, we restrict our search of a good γ that maximizes the utility of DaRRM_γ to in the "symmetric form family". To show the main privacy amplification result under a small m in Lemma B.10, Section B.2.4, we need a few building blocks, shown in Section B.2.3. We first show in Lemma B.7, Section B.2.3 two clean sufficient conditions that if a "symmetric form family" γ satisfies, then DaRRM_γ is $m\epsilon$-differentially private, in terms of the expectation of the γ function applied to Binomial random variables. The Binomial random variables appear in the lemma, because recall the sum of the observed outcomes on a dataset $\mathcal{D}$, $\mathcal{L}(\mathcal{D})$, follows a Binomial distribution in the i.i.d. mechanisms setting. Next, we show a recurrence relationship that connects the expectation of Binomial random variables to Hypergeometric random variables in Lemma B.9. This is needed because observe that for γ functions that makes DaRRM_γ have the same output as the majority of subsampled mechanisms, the h function is now a sum of pmfs of the Hypergeometric random variable.

Finally, the proof of the main result under a small m (Lemma B.10) is presented in Section B.2.4, based on Lemma B.7 and Lemma B.9. We show in Lemma B.10 that γ_{DSub}, i.e., the γ function that enables the output of $\mathsf{DaRRM}_{\gamma_{DSub}}$ and outputting the majority of $2m-1$ subsampled mechanisms to be the same, belongs to the "symmetric form family" and satisfies the sufficient conditions as stated in Lemma B.7, implying $\mathsf{DaRRM}_{\gamma_{DSub}}$ being $m\epsilon$-differentially private.

B.2.3 Building Blocks

Lemma B.7 (Privacy conditions of the "symmetric form family" functions)**.** *Let random variables $X \sim$ Binomial$(K-1, p')$, $Y \sim$ Binomial$(K-1, e^\epsilon p')$, $\hat{X} \sim$ Binomial$(K-1, 1-e^\epsilon(1-p))$ and $\hat{Y} \sim$ Binomial$(K-1, p)$. For a function $\gamma : \{0, 1, \ldots, K\} \to [0, 1]$ that belongs to the "symmetric form family" (Definition B.6), if γ also satisfies both conditions as follows:*

$$e^{m\epsilon}\mathbb{E}_X[h(X+1) - h(X)] \geq e^\epsilon \mathbb{E}_Y[h(Y+1) - h(Y)], \quad \forall p' \in [0, \frac{1}{1+e^\epsilon}] \tag{107}$$

$$e^{(m+1)\epsilon}\mathbb{E}_{\hat{X}}[h(\hat{X}+1) - h(\hat{X})] \geq \mathbb{E}_{\hat{Y}}[h(\hat{Y}+1) - h(\hat{Y})], \quad \forall p \in [\frac{1}{1+e^{-\epsilon}}, 1] \tag{108}$$

then Algorithm DaRRM_γ is $m\epsilon$-differentially private.

Proof of Lemma B.7. Since $h(l+1) \geq h(l)$ on $l \in \{0, \ldots, K\}$, $\gamma(l) \geq \gamma(l+1), \forall l \leq \frac{K}{2}$ and $\gamma(l+1) \geq \gamma(l), \forall l \geq \frac{K}{2}$. Furthermore, since $h(l)+h(K-l) = 1$, $\gamma(\frac{K-1}{2}) = 1-2h(\frac{K-1}{2}) = 1-2(1-h(\frac{K+1}{2})) = 2h(\frac{K+1}{2})-1$. Hence, any γ that belongs to the "symmetric form family" satisfies: 1) symmetric around $\frac{K}{2}$, 2) the monotonicity assumption, and 3) $\gamma(\frac{K-1}{2}) = \gamma(\frac{K+1}{2}) > 0$.

Therefore, by Lemma B.1, the worst case probabilities $(p^*, p'^*) = \arg\max_{(p,p')\in\mathcal{F}} f(p, p'; \gamma)$ are on one of the two boundaries of $\mathcal{F}$, satisfying

$$p^* = e^\epsilon p'^*, \qquad \forall p^* \in [0, \frac{1}{e^{-\epsilon}+1}], p'^* \in [0, \frac{1}{1+e^\epsilon}] \tag{109}$$

$$\text{or} \quad 1 - p'^* = e^\epsilon(1-p^*), \qquad \forall p^* \in [\frac{1}{1+e^{-\epsilon}}, 1], p'^* \in [\frac{1}{1+e^\epsilon}, 1] \tag{110}$$

We now derive the sufficient conditions that if any γ from the "symmetric form family" satisfy, then DaRRM_γ is $m\epsilon$-differentially private, from the two boundaries as in Eq. 109 and Eq. 110 separately.

Part I: Deriving a sufficient condition from Eq. 109 for "symmetric form family" γ.

Consider the boundary of $\mathcal{F}$, $p = e^\epsilon p'$, $\forall p \in [0, \frac{1}{1+e^{-\epsilon}}], p' \in [0, \frac{1}{1+e^\epsilon}]$.

Given any γ, plugging $p = e^\epsilon p'$ into the privacy cost objective $f(p, p'; \gamma)$, one gets

$$f(p';\gamma) = \sum_{l=0}^{\frac{K-1}{2}} (e^{m\epsilon}\binom{K}{l}p'^l(1-p')^{K-l} - \binom{K}{l}(e^\epsilon p')^l(1-e^\epsilon p')^{K-l}) \cdot \gamma(l) \tag{111}$$
$$+ \sum_{l=\frac{K+1}{2}}^{K} (\binom{K}{l}(e^\epsilon p')^l(1-e^\epsilon p')^{K-l} - e^{m\epsilon}\binom{K}{l}p'^l(1-p')^{K-l}) \cdot \gamma(l)$$

The gradient w.r.t. p' is

$$\frac{\nabla_{p'} f(p');\gamma}{K} = e^{m\epsilon} \sum_{l=0}^{\frac{K-1}{2}-1} \binom{K-1}{l} p'^l (1-p')^{K-l-1} \Big(\gamma(l+1) - \gamma(l)\Big) - 2e^{m\epsilon}\binom{K-1}{\frac{K-1}{2}} p'^{\frac{K-1}{2}}(1-p')^{\frac{K-1}{2}}\gamma(\frac{K-1}{2}) \tag{112}$$
$$+ e^{m\epsilon} \sum_{l=\frac{K+1}{2}}^{K-1} \binom{K-1}{l} p'^l (1-p')^{K-l-1} \Big(\gamma(l) - \gamma(l+1)\Big)$$
$$+ e^{\epsilon} \sum_{l=0}^{\frac{K-1}{2}-1} \binom{K-1}{l} (e^\epsilon p')^l (1-e^\epsilon p')^{K-l-1} \Big(\gamma(l) - \gamma(l+1)\Big) + 2e^{\epsilon}\binom{K-1}{\frac{K-1}{2}} (e^\epsilon p')^{\frac{K-1}{2}}(1-e^\epsilon p')^{\frac{K-1}{2}}\gamma(\frac{K-1}{2})$$
$$+ e^{\epsilon} \sum_{l=\frac{K+1}{2}}^{K-1} \binom{K-1}{l} (e^\epsilon p')^l (1-e^\epsilon p')^{K-l-1} \Big(\gamma(l+1) - \gamma(l)\Big)$$

Consider $l \in \{0, 1, \ldots, K\}$ in the above Eq. 112. For any function γ that belongs to the "symmetric form family",

1. If $l \le \frac{K}{2}$, $\gamma(l) - \gamma(l+1) = (1 - 2h(l)) - (1 - 2h(l+1)) = 2h(l+1) - 2h(l)$
2. If $l \ge \frac{K}{2}$, $\gamma(l+1) - \gamma(l) = (2h(l+1) - 1) - (2h(l) - 1) = 2h(l+1) - 2h(l)$
3. Since $\gamma(\frac{K-1}{2}) = \gamma(\frac{K+1}{2})$,

$$2\gamma(\frac{K-1}{2}) = \Big(\gamma(\frac{K-1}{2}) + \gamma(\frac{K+1}{2})\Big) \tag{113}$$
$$= \Big(1 - 2h(\frac{K-1}{2}) + 2h(\frac{K+1}{2}) - 1\Big) \tag{114}$$
$$= 2h(\frac{K+1}{2}) - 2h(\frac{K-1}{2}) \tag{115}$$

Hence, following Eq. 112, the gradient, $\nabla_{p'} f(p';\gamma)$, given a "symmetric form family" γ can be written as

$$\frac{\nabla_{p'} f(p';\gamma)}{K} = -e^{m\epsilon} \sum_{l=0}^{K-1} \binom{K-1}{l} p'^l (1-p')^{K-l} \Big(2h(l+1) - 2h(l)\Big) \tag{116}$$
$$+ e^{\epsilon} \sum_{l=0}^{K-1} \binom{K-1}{l} (e^\epsilon p')^l (1-e^\epsilon p')^{K-l-1} \Big(2h(l+1) - 2h(l)\Big)$$
$$= -2e^{m\epsilon}\mathbb{E}_X[h(X+1) - h(X)] + 2e^{\epsilon}\mathbb{E}_Y[h(Y+1) - h(Y)] \tag{117}$$

where $X \sim \text{Binomial}(K-1, p')$ and $Y \sim \text{Binomial}(K-1, e^\epsilon p')$. The above implies

$$\nabla_{p'} f(p';\gamma) \le 0 \iff e^{\epsilon}\mathbb{E}_Y[h(Y+1) - h(Y)] \le e^{m\epsilon}\mathbb{E}_X[h(X+1) - h(X)] \tag{118}$$

If $\nabla_{p'} f(p';\gamma) \leq 0$, then we know the local worst case probabilities on the boundary $p = e^{\epsilon}p', \forall p \in [0, \frac{1}{1+e^{-\epsilon}}]$ given any γ is $(p^*_{local}, p'^*_{local}) = \arg\max_{(p,p'):p=e^{\epsilon}p', p\in[0,\frac{1}{1+e^{-\epsilon}}]} f(p,p';\gamma) = (0,0)$. Furthermore, recall the privacy cost objective given any γ is

$$
\begin{aligned}
&f(p,p';\gamma) \\
&= \sum_{l=0}^{\frac{K-1}{2}} (e^{m\epsilon}\alpha'_l - \alpha_l)\cdot\gamma(l) + \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha'_l)\cdot\gamma(l) \\
&= \sum_{l=0}^{\frac{K-1}{2}} \left(e^{m\epsilon}\binom{K}{l}p'^l(1-p')^{K-l} - \binom{K}{l}p^l(1-p)^{K-l}\right)\cdot\gamma(l) + \sum_{l=\frac{K+1}{2}}^{K} \left(\binom{K}{l}p^l(1-p)^{K-l} - e^{m\epsilon}\binom{K}{l}p'^l(1-p')^{K-l}\right)\cdot\gamma(l)
\end{aligned}
$$

and so for any γ,

$$
f(0,0;\gamma) = (e^{m\epsilon}-1)\cdot\gamma(0) \leq e^{m\epsilon}-1 \tag{119}
$$

Also, notice the local minimum on this boundary is

$$
(p_{min}, p'_{min}) = \underset{(p,p'):p=e^{\epsilon}p', p\in[0,\frac{1}{1+e^{-\epsilon}}]}{\arg\min} f(p,p'l;\gamma) = (\frac{1}{1+e^{-\epsilon}}, \frac{1}{1+e^{\epsilon}}) \tag{120}
$$

Part II: Deriving a sufficient condition from Eq. 110 for "symmetric form family" γ.

Consider the boundary of $\mathcal{F}$, $1-p' = e^{\epsilon}(1-p)$, $\forall p \in [\frac{1}{1+e^{-\epsilon}}, 1], p' \in [\frac{1}{1+e^{\epsilon}}, 1]$. For simplicity, let $q = 1-p \in [0, \frac{1}{1+e^{\epsilon}}]$ and $q' = 1-p' \in [0, \frac{1}{1+e^{-\epsilon}}]$. Plugging $q' = e^{\epsilon}q$ into the privacy cost objective, one gets, given any γ,

$$
\begin{aligned}
f(q;\gamma) = &\sum_{l=0}^{\frac{K-1}{2}} \left(e^{m\epsilon}\binom{K}{l}(1-e^{\epsilon}q)^l(e^{\epsilon}q)^{K-l} - \binom{K}{l}(1-q)^l q^{K-l}\right)\cdot\gamma(l) \\
&+ \sum_{l=\frac{K+1}{2}}^{K} \left(\binom{K}{l}(1-q)^l q^{K-l} - e^{m\epsilon}\binom{K}{l}(1-e^{\epsilon}q)^l(e^{\epsilon}q)^{K-l}\right)\cdot\gamma(l)
\end{aligned} \tag{121}
$$

The gradient w.r.t. q is

$$
\begin{aligned}
\frac{\nabla_q f(q;\gamma)}{K} = &\sum_{l=0}^{\frac{K-1}{2}-1} e^{(m+1)\epsilon}\binom{K-1}{l}(1-e^{\epsilon}q)^l(e^{\epsilon}q)^{K-l-1}\cdot\Big(\gamma(l)-\gamma(l+1)\Big) \\
&+ \sum_{l=\frac{K+1}{2}}^{K-1}\binom{K-1}{l}(1-e^{\epsilon}q)^l(e^{\epsilon}q)^{K-l-1}\cdot\Big(\gamma(l+1)-\gamma(l)\Big) + 2e^{(m+1)\epsilon}\binom{K-1}{\frac{K-1}{2}}(1-e^{\epsilon}q)^{\frac{K-1}{2}}(e^{\epsilon}q)^{\frac{K-1}{2}}\cdot\gamma(\frac{K-1}{2}) \\
&+ \sum_{l=0}^{\frac{K-1}{2}-1}\binom{K-1}{l}(1-q)^l q^{K-l-1}\cdot\Big(\gamma(l+1)-\gamma(l)\Big) \\
&+ \sum_{l=\frac{K+1}{2}}^{K-1}(1-q)^l q^{K-l-1}\cdot\Big(\gamma(l)-\gamma(l+1)\Big) - 2\binom{K-1}{\frac{K-1}{2}}(1-q)^{\frac{K-1}{2}}q^{\frac{K-1}{2}}\cdot\gamma(\frac{K-1}{2})
\end{aligned} \tag{122}
$$

For any function γ that belongs to the "symmetric form family", the gradient $\nabla_q f(q;\gamma)$ can be written as

$$
\frac{\nabla_q f(q;\gamma)}{K} = e^{(m+1)\epsilon}\sum_{l=0}^{K-1}\binom{K-1}{l}(1-e^{\epsilon}q)^l(e^{\epsilon}q)^{K-l-1}\cdot\Big(2h(l+1)-2h(l)\Big) \tag{123}
$$

$$
\begin{aligned}
&- \sum_{l=0}^{K} \binom{K-1}{l} (1-q)^l q^{K-l-1} \cdot \Big(2h(l+1) - 2h(l)\Big) \\
&= 2e^{(m+1)\epsilon} \mathbb{E}_{\hat{X}}[h(\hat{X}+1) - h(\hat{X})] - 2\mathbb{E}_{\hat{Y}}[h(\hat{Y}+1) - h(\hat{Y})] \qquad (124)
\end{aligned}
$$

where $\hat{X} \sim \text{Binomial}(K-1, 1-e^{\epsilon}(1-p))$ and $\hat{Y} \sim \text{Binomial}(K-1, p)$. The above implies

$$
\nabla_q f(q;\gamma) \geq 0 \iff e^{(m+1)\epsilon} \mathbb{E}_{\hat{X}}[h(\hat{X}+1) - h(\hat{X})] \geq \mathbb{E}_{\hat{Y}}[h(\hat{Y}+1) - h(\hat{Y})] \qquad (125)
$$

If $\nabla_q f(q;\gamma) \geq 0$, then since $q \in [0, \frac{1}{1+e^{\epsilon}}]$, we know that the local maximum given any γ is $(q^*_{local}, q'^*_{local}) = \arg\max_{(q,q'):q'=e^{\epsilon}q, q\in[0,\frac{1}{1+e^{\epsilon}}]} f(q,q';\gamma) = (\frac{1}{1+e^{\epsilon}}, \frac{1}{1+e^{-\epsilon}})$. That is,

$$
(p^*_{local}, p'^*_{local}) = \underset{(p,p'):1-p'=e^{\epsilon}(1-p), p\in[\frac{1}{1+e^{-\epsilon}},1]}{\arg\max} f(p,p';\gamma) = (1-q^*_{local}, 1-q'^*_{local}) = (\frac{1}{1+e^{-\epsilon}}, \frac{1}{1+e^{\epsilon}})
$$

Notice by Eq. 120, the above $(\frac{1}{1+e^{-\epsilon}}, \frac{1}{1+e^{\epsilon}})$ is the local minimum on the first boundary $p = e^{\epsilon}p', \forall p \in [0, \frac{1}{1+e^{-\epsilon}}]$. Therefore, given an arbitrary γ function, if it satisfies both of the following:

1. On the boundary $p = e^{\epsilon}p', \forall p \in [0, \frac{1}{1+e^{-\epsilon}}]$, $\nabla_{p'} f(p';\gamma) \leq 0$
2. On the boundary $1-p' = e^{\epsilon}(1-p), \forall p \in [\frac{1}{1+e^{-\epsilon}}, 1]$, $\nabla_{q'} f(q';\gamma) \geq 0$ where $q' = 1-p'$

then the global worst case probabilities given this γ is $(p^*, p'^*) = \arg\max_{(p,p')\in\mathcal{F}} f(p,p';\gamma) = (0,0)$. Furthermore, since by Eq. 119, $f(0,0;\gamma) \leq e^{m\epsilon} - 1$ for any γ, this implies DaRRM_{γ} is $m\epsilon$-differentially private by Lemma 3.4.

Now, if γ belongs to the "symmetric form family", by Eq. 118 and Eq. 125, the sufficient conditions for γ that enables DaRRM_{γ} to be $m\epsilon$-differentially private are hence

$$
\begin{aligned}
&e^{\epsilon}\mathbb{E}_Y[h(Y+1) - h(Y)] \leq e^{m\epsilon}\mathbb{E}_X[h(X+1) - h(X)], \quad \forall p' \in [0, \frac{1}{1+e^{\epsilon}}] \\
&\text{and} \quad e^{(m+1)\epsilon}\mathbb{E}_{\hat{X}}[h(\hat{X}+1) - h(\hat{X})] \geq \mathbb{E}_{\hat{Y}}[h(\hat{Y}+1) - h(\hat{Y})], \quad \forall p \in [\frac{1}{1+e^{-\epsilon}}, 1]
\end{aligned}
$$

where $X \sim \text{Binomial}(K-1, p')$, $Y \sim \text{Binomial}(K-1, e^{\epsilon}p')$, $\hat{X} \sim \text{Binomial}(K-1, 1-e^{\epsilon}(1-p))$ and $\hat{Y} \sim \text{Binomial}(K-1, p)$.

□

Lemma B.8 (Binomial Expectation Recurrence Relationship (Theorem 2.1 of Zhang et al. (2019))). *Let $X_{(K-1)} \sim$ Binomial$(K-1,p)$ and $X_{(K)} \sim$ Binomial(K,p). Let $g(x)$ be a function with $-\infty < \mathbb{E}[g(X_{(K-1)})] < \infty$ and $-\infty < g(-1) < \infty$, then*

$$
Kp\mathbb{E}_{X_{(K-1)}}[g(X_{(K-1)})] = \mathbb{E}_{X_{(K)}}[X_{(K)}g(X_{(K)} - 1)] \qquad (126)
$$

Lemma B.9. *Given $i, m, K \in \mathbb{Z}$, $K \geq 1$, $0 \leq i \leq m \leq K$, let $X_{(K)} \sim$ Binomial(K,p) for some $p \in [0,1]$, there is*

$$
\frac{1}{\binom{K}{m}} \mathbb{E}_{X_{(K)}} \left[\binom{X}{i} \binom{K-X}{m-i} \right] = \binom{m}{i} p^i (1-p)^{m-i} \qquad (127)
$$

Proof of Lemma B.9. We show the above statement in Eq. 127 by induction on K and m.

Base Case: $K = 1$.

1. If $m = 0$, then $i = 0$. $\frac{1}{\binom{1}{0}} \mathbb{E}_{X_{(1)}}[\binom{X}{0}\binom{1-X}{0}] = \mathbb{E}_{X_{(1)}}[1] = 1$, and $\binom{0}{0}p^0(1-p)^0 = 1$.

2. If $m = 1$,
   (a) $i = 0$, $\frac{1}{\binom{1}{1}}\mathbb{E}_{X_{(1)}}[\binom{X}{0}\binom{1-X}{1}] = \mathbb{E}_{X_{(1)}}[1 - X] = 1 - p$, and $\binom{1}{0}p^0(1-p)^1 = 1 - p$
   (b) $i = 1$, $\frac{1}{\binom{1}{1}}\mathbb{E}_{X_{(1)}}[\binom{X}{1}\binom{1-X}{0}] = \mathbb{E}_{X_{(1)}}[X] = p$, and $\binom{1}{1}p^1(1-p)^0 = p$.

Hence, Eq. 127 holds for the base case.

Induction Hypothesis: Suppose the statement holds for some $K \geq 1$ and $0 \leq i \leq m \leq K$. Consider $1 \leq i \leq m \leq K + 1$,

$$\frac{1}{\binom{K+1}{m}}\mathbb{E}_{X_{(K+1)}}\left[\binom{X}{i}\binom{K+1-X}{m-i}\right] \tag{128}$$

$$= \frac{1}{\binom{K+1}{m}}\mathbb{E}_{X_{(K+1)}}[\frac{X!}{i!(X-i)!}\frac{(K+1-X)!}{(m-i)!(K+1-X-(m-i))!}] \tag{129}$$

$$= \frac{1}{\binom{K+1}{m}i!(m-i)!}\mathbb{E}_{X_{(K+1)}}[X\frac{(X-1)!}{((X-1)-(i-1))!}\frac{(K-(X-1))!}{(K-(X-1)-((m-1)-(i-1)))!}] \tag{130}$$

$$= \frac{1}{\binom{K+1}{m}i!(m-i)!}\mathbb{E}_{X_{(K)}}[\frac{X!}{(X-(i-1))!}\frac{(K-X)!}{(K-X-((m-1)-(i-1)))!}] \tag{131}$$

(By Lemma B.8)

$$= \frac{(i-1)!(m-i)!}{\binom{K+1}{m}i!(m-i)!}\mathbb{E}_{X_{(K)}}[\binom{X}{i-1}\binom{K-X}{(m-1)-(i-1)}] \tag{132}$$

$$= \frac{(i-1)!}{\binom{K+1}{m}i!}(K+1)p\binom{K}{m-1}\binom{m-1}{i-1}p^{i-1}(1-p)^{m-i} \tag{133}$$

(By Induction Hypothesis)

$$= \frac{m!(K+1-m)!}{(K+1)!i}\frac{K!}{(m-1)!(K-m+1)!}\frac{(m-1)!}{(i-1)!(m-i)!}(K+1)p^i(1-p)^{m-i} \tag{134}$$

$$= \frac{m!}{i!(m-i)!}p^i(1-p)^{m-i} = \binom{m}{i}p^i(1-p)^{m-i} \tag{135}$$

Now we consider the edge cases when $0 = i \leq m$.

If $i = 0$ and $m = 0$,

$$\frac{1}{\binom{K+1}{0}}\mathbb{E}_{X_{(K+1)}}[\binom{X}{0}\binom{K+1-X}{0}] = 1 \cdot \mathbb{E}_{X_{(K+1)}}[1] = 1 = \binom{0}{0}p^0(1-p)^0 \tag{136}$$

If $i = 0$ and $m > 0$,

$$\frac{1}{\binom{K+1}{m}}\mathbb{E}_{X_{(K+1)}}[\binom{K+1-X}{m}] \tag{137}$$

$$= \frac{1}{\binom{K+1}{m}}\sum_{x=0}^{K+1}\binom{K+1-x}{m}\binom{K+1}{x}p^x(1-p)^{K+1-x} \tag{138}$$

$$= \frac{1}{\binom{K+1}{m}}\sum_{x=0}^{K+1}\binom{K+1-x}{m}\left(\binom{K}{x} + \binom{K}{x-1}\mathbb{I}\{x \geq 1\}\right)p^x(1-p)^{K+1-x} \tag{139}$$

$$= \frac{1}{\binom{K+1}{m}}\sum_{x=0}^{K}\binom{K+1-x}{m}\binom{K}{x}p^x(1-p)^{K+1-x} + \frac{1}{\binom{K+1}{m}}\sum_{x=1}^{K+1}\binom{K+1-x}{m}\binom{K}{x-1}p^x(1-p)^{K+1-x} \tag{140}$$

(Since when $x = K + 1$ and $m > 0$, $\binom{K+1-x}{m} = 0$)

$$= \frac{1}{\binom{K+1}{m}}\Big(\sum_{x=0}^{K}\binom{K-x}{m}\binom{K}{x}p^x(1-p)^{K+1-x} + \sum_{x=0}^{K}\binom{K-x}{m-1}\binom{K}{x}p^x(1-p)^{K+1-x}\Big) \tag{141}$$

$$+ \frac{1}{\binom{K+1}{m}}\sum_{x=0}^{K}\binom{K-x}{m}\binom{K}{x}p^{x+1}(1-p)^{K-x}$$

$$\text{(Since } \binom{K+1-x}{m} = \binom{K-x}{m} + \binom{K-x}{m-1}\text{)}$$

$$= \frac{1}{\binom{K+1}{m}}\Big((1-p)\mathbb{E}_{X_{(K)}}[\binom{K-X}{m}] + (1-p)\mathbb{E}_{X_{(k)}}[\binom{K-X}{m-1}]\Big) + \frac{1}{\binom{K+1}{m}}p\mathbb{E}_{X_{(K)}}[\binom{K-X}{m}] \tag{142}$$

$$= \frac{1}{\binom{K+1}{m}}\Big(\mathbb{E}_{X_{(K)}}[\binom{K-X}{m}] + (1-p)\mathbb{E}_{X_{(K)}}[\binom{K-X}{m-1}]\Big) \tag{143}$$

$$= \frac{1}{\binom{K+1}{m}}\Big(\binom{K}{m}(1-p)^m + (1-p)\binom{K}{m-1}(1-p)^{m-1}\Big) \tag{144}$$

$$\text{(By Induction Hypothesis)} \tag{145}$$

$$= \frac{1}{\binom{K+1}{m}}\binom{K+1}{m}(1-p)^m \tag{146}$$

$$= (1-p)^m \tag{147}$$

Hence, Eq. 127 holds for all $K \geq 1$ and $0 \leq i \leq m \leq K$.

□

B.2.4 Main Result: Privacy Amplification Under a Small m

Lemma B.10 (Privacy amplification, $m \leq \frac{K-1}{2}$)**.** *Consider using DaRRM (Algorithm 1) to solve Problem 1.1, with i.i.d. mechanisms $\{M_i\}_{i=1}^K$, $p_i = p$, $p_i' = p'$, $\forall i \in [K]$, the privacy allowance $1 \leq m \leq \frac{K-1}{2}, m \in \mathbb{Z}$ and $\delta = \Delta = 0$. Let the noise function be that*

$$\gamma_{DSub}(l) = \begin{cases} 1 - 2h(l) & \forall l \in \{0, 1, \ldots, \frac{K-1}{2}\} \\ 2h(l) - 1 & \forall l \in \{\frac{K+1}{2}, \ldots, K\} \end{cases} \tag{148}$$

where $h : \{0, 1, \ldots, K\} \to [0, 1]$ and $h(l) = \sum_{i=m}^{2m-1} \frac{\binom{l}{i}\binom{K-l}{2m-1-i}}{\binom{K}{2m-1}}$, $\forall l \in \{0, 1, \ldots, K\}$, then Algorithm DaRRM$_{\gamma_{DSub}}$ is $m\epsilon$-differentially private.

Proof of Lemma B.10. First, note γ_{DSub} belongs to the "symmetric form family". We show γ_{DSub} satisfies the two sufficient conditions in Lemma B.7 and hence by Lemma B.7, DaRRM$_{\gamma_{DSub}}$ is $m\epsilon$-differentially private. Specifically, we consider $h(l) = \sum_{i=m}^{2m-1} \frac{\binom{l}{i}\binom{K-l}{2m-1-i}}{\binom{K}{2m-1}}$, $\forall l \in \{0, 1, \ldots, K\}$ and $1 \leq m \leq K$.

Two show the first condition is satisfied, let $X_{(K-1)} \sim \text{Binomial}(K-1, p)$ and $Y_{(K-1)} \sim \text{Binomial}(K-1, e^\epsilon p)$, and consider $p \in [0, \frac{1}{1+e^\epsilon}]$.

$$\mathbb{E}_{X_{(K-1)}}[h(X+1)] = \frac{1}{\binom{K}{2m-1}}\sum_{i=m}^{2m-1}\mathbb{E}_{X_{(K-1)}}[\binom{X+1}{i}\binom{K-X-1}{2m-1-i}] \tag{149}$$

$$= \frac{1}{\binom{K}{2m-1}}\sum_{i=m}^{2m-1}\mathbb{E}_{X_{(K-1)}}[\binom{X}{i}\binom{K-X-1}{2m-1-i} + \binom{X}{i-1}\binom{K-X-1}{2m-1-i}] \tag{150}$$

$$\text{(Since } \binom{X+1}{i} = \binom{X}{i} + \binom{X}{i-1}\mathbb{I}\{i \geq 1\}\text{)}$$

$$= \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \left(\mathbb{E}_{X_{(K-1)}}[\binom{X}{i}\binom{K-1-X}{2m-1-i}] + \mathbb{E}_{X_{(K-1)}}[\binom{X}{i-1}\binom{K-1-X}{(2m-2)-(i-1)}] \right) \tag{151}$$

$$= \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \left(\binom{K-1}{2m-1}\binom{2m-1}{i} p^i (1-p)^{2m-1-i} + \binom{K-1}{2m-2}\binom{2m-2}{i-1} p^{i-1}(1-p)^{2m-1-i} \right) \tag{152}$$

(By Lemma B.9)

$$\mathbb{E}_{X_{(K-1)}}[h(X)] = \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \mathbb{E}_{X_{(K-1)}}[\binom{X}{i}\binom{K-X}{2m-1-i}] \tag{153}$$

(Since $\binom{K-X}{2m-1-i} = \binom{K-1-X}{2m-1-i} + \binom{K-1-X}{2m-2-i}$)

$$= \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \left(\mathbb{E}_{X_{(K-1)}}[\binom{X}{i}\binom{K-1-X}{2m-1-i}] + \mathbb{E}_{X_{(K-1)}}[\binom{X}{i}\binom{K-1-X}{2m-2-i}] \mathbb{I}\{i \leq 2m-2\} \right) \tag{154}$$

$$= \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \left(\binom{K-1}{2m-1}\binom{2m-1}{i} p^i (1-p)^{2m-1-i} + \binom{K-1}{2m-2}\binom{2m-2}{i} p^i (1-p)^{2m-2-i} \mathbb{I}\{i \leq 2m-2\} \right) \tag{155}$$

(By Lemma B.9)

Hence, following Eq. 155 and Eq. 152,

$$\mathbb{E}_{X_{(K-1)}}[h(X+1) - h(X)] \tag{156}$$

$$= \frac{1}{\binom{K}{2m-1}} \left(\sum_{i=m}^{2m-1} \binom{K-1}{2m-2}\binom{2m-2}{i-1} p^{i-1}(1-p)^{2m-1-i} - \sum_{i=m}^{2m-2} \binom{K-1}{2m-2}\binom{2m-2}{i} p^i (1-p)^{2m-2-i} \right) \tag{157}$$

$$= \frac{1}{\binom{K}{2m-1}} \left(\sum_{i=m-1}^{2m-2} \binom{K-1}{2m-2}\binom{2m-2}{i} p^i (1-p)^{2m-2-i} - \sum_{i=m}^{2m-2} \binom{K-1}{2m-2}\binom{2m-2}{i} p^i (1-p)^{2m-2-i} \right) \tag{158}$$

$$= \frac{2m-1}{K}\binom{2m-2}{m-1} p^{m-1}(1-p)^{m-1} \tag{159}$$

Similarly,

$$\mathbb{E}_{Y_{(K-1)}}[h(Y+1) - h(Y)] = \frac{2m-1}{K}\binom{2m-2}{m-1} (e^\epsilon p)^{m-1}(1-e^\epsilon p)^{m-1} \tag{160}$$

Since $p \in [0, \frac{1}{1+e^\epsilon}]$, there is $p(1-p) \geq e^{-\epsilon} e^\epsilon p(1-e^\epsilon p)$. Hence,

$$e^{(m-1)\epsilon}\mathbb{E}_{X_{(K-1)}}[h(X+1) - h(X)] = \frac{2m-1}{K}\binom{2m-2}{m-1} e^{(m-1)\epsilon} p^{m-1}(1-p)^{m-1} \tag{161}$$

$$\geq \frac{2m-1}{K}\binom{2m-2}{m-1} e^{(m-1)\epsilon} (e^{-\epsilon} e^\epsilon p(1-e^\epsilon p))^{m-1} \tag{162}$$

$$= \frac{2m-1}{K}\binom{2m-2}{m-1} (e^\epsilon p)^{m-1}(1-e^\epsilon p)^{m-1} \tag{163}$$

$$= \mathbb{E}_{Y_{(K-1)}}[h(Y+1) - h(Y)] \tag{164}$$

implying

$$e^{m\epsilon}\mathbb{E}_{X_{(K-1)}}[h(X+1) - h(X)] \geq e^{\epsilon}\mathbb{E}_{Y_{(K-1)}}[h(Y+1) - h(Y)] \tag{165}$$

and the first condition is satisfied.

To show the second condition is satisfied, let $\hat{X}_{(K-1)} \sim \text{Binom}(K-1, 1-e^{\epsilon}(1-p))$ and $\hat{Y}_{(K-1)} \sim \text{Binom}(K-1, p)$, and consider $p \in [\frac{1}{1+e^{-\epsilon}}, 1]$.

$$\mathbb{E}_{\hat{X}_{(K-1)}}[h(\hat{X}+1)] = \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \left(\mathbb{E}_{\hat{X}_{(K-1)}}[\binom{\hat{X}}{i}\binom{K-1-\hat{X}}{2m-1-i}] + \mathbb{E}_{\hat{X}_{(K-1)}}[\binom{\hat{X}}{i-1}\binom{K-1-\hat{X}}{(2m-2)-(i-1)}] \right) \tag{166}$$

$$= \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \Big(\binom{K-1}{2m-1}\binom{2m-1}{i}(1-e^{\epsilon}(1-p))^{i}(e^{\epsilon}(1-p))^{2m-1-i} \tag{167}$$

$$+ \binom{K-1}{2m-2}\binom{2m-2}{i-1}(1-e^{\epsilon}(1-p))^{i-1}(e^{\epsilon}(1-p))^{2m-1-i} \Big)$$

By Lemma B.9

and

$$\mathbb{E}_{\hat{X}_{(K-1)}}[h(\hat{X})] = \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \left(\mathbb{E}_{\hat{X}_{(K-1)}}[\binom{\hat{X}}{i}\binom{K-1-\hat{X}}{2m-1-i}] + \mathbb{E}_{\hat{X}_{(K-1)}}[\binom{\hat{X}}{i}\binom{K-1-\hat{X}}{2m-2-i}]\mathbb{I}\{i \leq 2m-2\} \right) \tag{168}$$

$$= \frac{1}{\binom{K}{2m-1}} \sum_{i=m}^{2m-1} \Big(\binom{K-1}{2m-1}\binom{2m-1}{i}(1-e^{\epsilon}(1-p))^{i}(e^{\epsilon}(1-p))^{2m-1-i} \tag{169}$$

$$+ \binom{K-1}{2m-2}\binom{2m-2}{i}(1-e^{\epsilon}(1-p))^{i}(e^{\epsilon}(1-p))^{2m-2-i}\mathbb{I}\{i \leq 2m-2\} \Big)$$

By Lemma B.9

Hence, following Eq. 167 and Eq. 169,

$$\mathbb{E}_{\hat{X}_{(K-1)}}[h(\hat{X}+1) - h(\hat{X})] \tag{170}$$

$$= \frac{1}{\binom{K}{2m-1}} \Big(\sum_{i=m}^{2m-1} \binom{K-1}{2m-2}\binom{2m-2}{i-1}(1-e^{\epsilon}(1-p))^{i-1}(e^{\epsilon}(1-p))^{2m-1-i} \tag{171}$$

$$- \sum_{i=m}^{2m-2} \binom{K-1}{2m-2}\binom{2m-2}{i}(1-e^{\epsilon}(1-p))^{i}(e^{\epsilon}(1-p))^{2m-2-i} \Big)$$

$$= \frac{1}{\binom{K}{2m-1}} \Big(\sum_{i=m-1}^{2m-2} \binom{K-1}{2m-2}\binom{2m-2}{i}(1-e^{\epsilon}(1-p))^{i}(e^{\epsilon}(1-p))^{2m-2-i} \tag{172}$$

$$- \sum_{i=m}^{2m-2} \binom{K-1}{2m-2}\binom{2m-2}{i}(1-e^{\epsilon}(1-p))^{i}(e^{\epsilon}(1-p))^{2m-2-i} \Big)$$

$$= \frac{2m-1}{K}\binom{2m-2}{m-1}(1-e^{\epsilon}(1-p))^{m-1}(e^{\epsilon}(1-p))^{m-1} \tag{173}$$

Similarly,

$$\mathbb{E}_{\hat{Y}_{(K-1)}}[h(\hat{Y}+1) - h(\hat{Y})] = \frac{2m-1}{K}\binom{2m-2}{m-1}p^{m-1}(1-p)^{m-1} \tag{174}$$

Hence,

$$e^{(m+1)\epsilon}\mathbb{E}_{\hat{X}_{(K-1)}}[h(\hat{X}+1)-h(\hat{X})] = e^{(m+1)\epsilon}\frac{2m-1}{K}\binom{2m-2}{m-1}(1-e^{\epsilon}(1-p))^{m-1}(e^{\epsilon}(1-p))^{m-1} \tag{175}$$

$$\geq \frac{2m-1}{K}\binom{2m-2}{m-1}(1-e^{\epsilon}(1-p))^{m-1}e^{(m-1)\epsilon}(1-p)^{m-1} \tag{176}$$

$$= \frac{2m-1}{K}\binom{2m-2}{m-1}(e^{\epsilon}-e^{2\epsilon}(1-p))^{m-1}(1-p)^{m-1} \tag{177}$$

Note that

$$e^{\epsilon}-e^{2\epsilon}(1-p) = e^{\epsilon}-e^{2\epsilon}+e^{2\epsilon}p \geq p \tag{178}$$

$$\iff (e^{\epsilon}+1)(e^{\epsilon}-1)p \geq e^{\epsilon}(e^{\epsilon}-1) \tag{179}$$

$$\iff p \geq \frac{e^{\epsilon}}{e^{\epsilon}+1} = \frac{1}{1+e^{-\epsilon}} \tag{180}$$

and the condition needs to hold for $p \in [\frac{1}{1+e^{-\epsilon}}, 1]$.

Therefore, following Eq. 177,

$$e^{(m+1)\epsilon}\mathbb{E}_{\hat{X}_{(K-1)}}[h(\hat{X}+1)-h(\hat{X})] \geq \frac{2m-1}{K}\binom{2m-2}{m-1}p^{m-1}(1-p)^{m-1} \tag{181}$$

$$= \mathbb{E}_{\hat{Y}_{(K-1)}}[h(\hat{Y}+1)-h(\hat{Y})] \tag{182}$$

implying the second condition is satisfied.

Therefore, by Lemma B.7, $\mathsf{DaRRM}_{\gamma_{DSub}}$ is $m\epsilon$-differentially private.

□

B.3 Comparing the Utility of Subsampling Approaches

Intuitively, if we subsample $2m-1$ mechanisms, the utility is higher than that of the naïve subsampling approach which outputs the majority based on only m mechanisms. To complete the story, we formally compare the utility of outputting the majority of $2m-1$ subsampled mechanisms (Theorem 4.1) and outputting the majority of m subsampled mechanisms (simple composition, Theorem 2.2) in the i.i.d. mechanisms and pure differential privacy setting, fixing the output privacy loss to be $m\epsilon$.

Lemma B.11. *Consider Problem 1.1 with i.i.d. mechanisms $\{M_i\}_{i=1}^K$, i.e., $p = p_i = \Pr[M_i(\mathcal{D})=1], p' = p'_i = \Pr[M_i(\mathcal{D}')=1], \forall i \in [K]$. Let $\gamma_1 : \{0,1,\ldots,K\} \to [0,1], \gamma_2 : \{0,1,\ldots,K\} \to [0,1]$ be two functions that are both symmetric around $\frac{K}{2}$. If $1 \geq \gamma_1(l) \geq \gamma_2(l) \geq 0, \forall l \in \{0,\ldots,K\}$, then $\mathcal{E}(\mathsf{DaRRM}_{\gamma_1}) \leq \mathcal{E}(\mathsf{DaRRM}_{\gamma_2})$.*

Proof. Recall $\mathcal{S} = \{S_1,\ldots,S_K\}$, where $S_i \sim M_i(\mathcal{D})$, is the set of observed outcomes from the mechanisms $\{M_i\}_{i=1}^K$. By Definition 2.4, for any γ that is symmetric around $\frac{K}{2}$, the error of DaRRM_γ is

$$\mathcal{E}(\mathsf{DaRRM}_\gamma) = \Big|\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D})=1] - \Pr[g(\mathcal{S})=1]\Big| \tag{183}$$

$$= \Big|\sum_{l=\frac{K+1}{2}}^{K}\Big(\gamma(l)+\frac{1}{2}(1-\gamma(l))\Big)\cdot\alpha_l + \sum_{l=0}^{\frac{K-1}{2}}\frac{1}{2}(1-\gamma(l))\cdot\alpha_l - \sum_{l=\frac{K+1}{2}}^{K}\alpha_l\Big| \tag{184}$$

$$= \Big|\sum_{l=\frac{K+1}{2}}^{K}\Big(\frac{1}{2}\gamma(l)-\frac{1}{2}\Big)\cdot\alpha_l + \sum_{l=0}^{\frac{K-1}{2}}\Big(\frac{1}{2}-\frac{1}{2}\gamma(l)\Big)\cdot\alpha_l\Big| \tag{185}$$

$$= \Big|\frac{1}{2}\sum_{l=\frac{K+1}{2}}^{K}(1-\gamma(l))\cdot(\alpha_l-\alpha_{K-l})\Big| \tag{186}$$

where $\alpha_l = \binom{K}{l} p^l (1-p)^{K-l}$, $\forall l \in \{0, 1, \ldots, K\}$ and recall $p = \Pr[M_i(\mathcal{D}) = 1]$, $\forall i \in [K]$.

For any $l \geq \frac{K+1}{2}$,

1. If $p = 0$ or $p = 1$, $\alpha_l = \alpha_{K-l}$.
2. Otherwise, for $p \in (0, 1)$,
   (a) If $p \geq \frac{1}{2}$,

$$\frac{\alpha_l}{\alpha_{K-l}} = \frac{p^l (1-p)^{K-l}}{p^{K-l}(1-p)^l} = p^{2l-K}(1-p)^{K-2l} = (\underbrace{\frac{p}{1-p}}_{\geq 1})^{\underbrace{2l-K}_{\geq 0}} \geq 1, \quad \Rightarrow \alpha_l \geq \alpha_{K-l} \tag{187}$$

   (b) If $p < \frac{1}{2}$,

$$\frac{\alpha_l}{\alpha_{K-l}} = (\underbrace{\frac{p}{1-p}}_{\leq 1})^{\underbrace{2l-K}_{\geq 0}} \leq 1, \quad \Rightarrow \alpha_l \leq \alpha_{K-l} \tag{188}$$

Hence, if $p \geq \frac{1}{2}$, then $\alpha_l \geq \alpha_{K-l}, \forall l \geq \frac{K+1}{2}$. Since $\gamma_1(l) \geq \gamma_2(l), \forall l \in \{0, \ldots, K\}$, $1 - \gamma_1(l) \leq 1 - \gamma_2(l)$, and so

$$\mathcal{E}(\mathsf{DaRRM}_{\gamma_1}) = \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}(1-\gamma_1(l)) \cdot (\alpha_l - \alpha_{K-l}) \leq \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}(1-\gamma_2(l)) \cdot (\alpha_l - \alpha_{K-l}) = \mathcal{E}(\mathsf{DaRRM}_{\gamma_2}) \tag{189}$$

Similarly, if $p < \frac{1}{2}$, then $\alpha_l \leq \alpha_{K-l}, \forall l \geq \frac{K+1}{2}$ and

$$\mathcal{E}(\mathsf{DaRRM}_{\gamma_1}) = \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}(1-\gamma_1(l)) \cdot (\alpha_{K-l} - \alpha_l) \leq \sum_{l=\frac{K+1}{2}}^{K} \frac{1}{2}(1-\gamma_2(l)) \cdot (\alpha_{K-l} - \alpha_l) = \mathcal{E}(\mathsf{DaRRM}_{\gamma_2})) \tag{190}$$

Therefore,

$$\mathcal{E}(\mathsf{DaRRM}_{\gamma_1}) \leq \mathcal{E}(\mathsf{DaRRM}_{\gamma_2}) \tag{191}$$

□

Since $\gamma_{DSub}(l) \geq \gamma_{Sub}(l)$, $\forall l \in \{0, 1, \ldots, K\}$, by Lemma B.11, $\mathcal{E}(\mathsf{DaRRM}_{\gamma_{DSub}}) \leq \mathcal{E}(\mathsf{DaRRM}_{\gamma_{Sub}})$ — that is, outputting $2m - 1$ mechanisms has a higher utility than outputting m mechanisms.

C Details of Section 5: Optimizing the Noise Function γ in DaRRM

C.1 Deriving the Optimization Objective

For any γ function that is symmetric around $\frac{K}{2}$, we can write the optimization objective as

$$\mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}[\mathcal{E}(\mathsf{DaRRM}_\gamma)] \tag{192}$$
$$= \mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}[|\Pr[\mathsf{DaRRM}_\gamma(\mathcal{D})=1]-\Pr[g(\mathcal{S})=1]|] \tag{193}$$
$$= \mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[\left|\sum_{l=\frac{K+1}{2}}^{K}\left(\alpha_l\cdot(\gamma(l)+\frac{1}{2}(1-\gamma(l)))-\alpha_l\right)+\sum_{l=0}^{\frac{K-1}{2}}\alpha_l\cdot\frac{1}{2}(1-\gamma(l))\right|\right] \tag{194}$$
$$= \mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[\left|\sum_{l=0}^{\frac{K-1}{2}}\alpha_l(\frac{1}{2}\gamma(l)-\frac{1}{2})+\sum_{l=\frac{K+1}{2}}^{K}\alpha_l(\frac{1}{2}-\frac{1}{2}\gamma(l))\right|\right] \tag{195}$$

The above follows by conditioning on $\mathcal{L}=l\in\{0,1,\ldots,K\}$, i.e. the sum of observed outcomes in $\mathcal{S}$

$$= \mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[\left|\frac{1}{2}\sum_{l=\frac{K+1}{2}}^{K}(\alpha_l-\alpha_{K-l})\,(1-\gamma(l))\right|\right] \tag{196}$$

The above follows by symmetry of γ

Furthermore, notice the objective is symmetric around 0, and can be written as

$$\mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[\frac{1}{2}\sum_{l=\frac{K+1}{2}}^{K}(\alpha_l-\alpha_{K-l})\,(1-\gamma(l))\right] \tag{197}$$
$$= \frac{1}{2}\mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[\sum_{l=\frac{K+1}{2}}^{K}\Big((\alpha_l-\alpha_{K-l})-(\alpha_l-\alpha_{K-l})\gamma(l)\Big)\right] \tag{198}$$
$$= \underbrace{\frac{1}{2}\mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[\sum_{l=\frac{K+1}{2}}^{K}(\alpha_l-\alpha_{K-l})\right]}_{:=A}\underbrace{-\frac{1}{2}\mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[\sum_{l=\frac{K+1}{2}}^{K}(\alpha_l-\alpha_{K-l})\gamma(l)\right]}_{:=B} \tag{199}$$

Since expression A in Eq. 199 does not involve γ, we only need to optimize expression B in Eq. 199. That is,

$$-\frac{1}{2}\mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[\sum_{l=\frac{K+1}{2}}^{K}(\alpha_l-\alpha_{K-l})\gamma(l)\right] \tag{200}$$
$$= -\frac{1}{2}\sum_{l=\frac{K+1}{2}}^{K}\mathbb{E}_{p_1,p_2,\ldots,p_K\sim\mathcal{T}}\left[(\alpha_l-\alpha_{K-l})\right]\cdot\gamma(l) \tag{201}$$

Eq. 201 is the optimization objective we use in the experiments. We see the optimization objective is linear in γ.

Note in the general setting, $\mathcal{L}(\mathcal{D})\sim\text{PoissonBinomial}(p_1,p_2,\ldots,p_K)$, where recall $\mathcal{L}(\mathcal{D})$ is the sum of observed outcomes on dataset $\mathcal{D}$, and hence, $\alpha_l=\Pr[\mathcal{L}(\mathcal{D})=l]$ is the pmf of the Poisson Binomial distribution at $l\in\{0,1,\ldots,K\}$.

C.2 Practical Approximation of the Objective

Since the optimization objective in Eq. 200 requires taking an expectation over $p_1,\ldots,p_K$, and this invovles integrating over K variables, which can be slow in practice, we propose the following approximation to

efficiently compute the objective. We start with a simple idea to compute the objective, by sampling p_i's from $[0, 1]$ and take an empirical average of the objective value over all subsampled sets of $p_1, \ldots, p_K$ as the approximation of the expectation in Section C.2.1. However, we found this approach is less numerically stable. We then propose the second approach to approximate the objective in Section C.2.2, which approximates the integration over p_i's using the rectangular rule instead of directly approximating the objective value. We use the second approximation approach in our experiments and empirically demonstrates its effectiveness. **Note approximating the optimization objective does not affect the privacy guarantee.**

C.2.1 Approximation via Direct Sampling of p_i's

One straightforward way of efficiently computing an approximation to the optimization objective is as follows:

Algorithm 4 Straightforward Approximation of the Optimization Objective

Input: # mechanisms $K \in \mathbb{N}$, # iterations $T \in \mathbb{N}$, noise function $\gamma : \{0, 1, \ldots, K\} \to [0, 1]$
for $t = 1, 2, \ldots, T$ **do**
  Sample $\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_K \sim \mathcal{T}$
  $\widehat{\mathcal{L}} \leftarrow \text{PoissonBinomail}(\hat{p}_1, \ldots, \hat{p}_K)$
  $\hat{\alpha}_l \leftarrow \Pr[\widehat{\mathcal{L}} = l], \forall l \{0, \ldots, K\}$
  $g_t \leftarrow -\frac{1}{2} \sum_{l=\frac{K+1}{2}}^{K} (\hat{\alpha}_l - \hat{\alpha}_{K-l}) \cdot \gamma(l)$
end for
Return $\frac{1}{T} \sum_{t=1}^{T} g_t$

However, we found this approximation is not very numerically stable even for $T = 10000$ in the experiments and so we propose to adopt the second approximation as follows.

C.2.2 Approximating the Integration Over p_i's

Consider the following surrogate objective:

$$-\frac{1}{2} \sum_{l=\frac{K+1}{2}}^{K} \int_{0.5}^{1} \int_{0.5}^{1} \cdots \int_{0.5}^{1} (\alpha_l - \alpha_{K-l}) dp_1 dp_2 \ldots dp_K \cdot \gamma(l) \tag{202}$$

where we approximate the integration instead of directly approximating the objective value. The approximation of the integration is based on the rectangular rule and that the Poisson Binomial distribution is invariant to the order of its probability parameters.

First, we discretize the integration over p_i's: pick $\tau = 50$ points representing probabilities between $[0.5, 1)$ with equal distance $\theta = \frac{0.5}{\tau}$. Denote this set of points as $\mathcal{W}$. We pick only $\tau = 50$ samples to ensure the distance between each sample, i.e., θ, is not too small; or this can cause numerical instability. For each $l \in \{\frac{K+1}{2}, \frac{K+1}{2} + 1, \ldots, K\}$, we want to compute an approximated coefficient for $\gamma(l)$ as follows:

$$\int_{0.5}^{1} \int_{0.5}^{1} \cdots \int_{0.5}^{1} (\alpha_l - \alpha_{K-l}) dp_1 dp_2 \ldots dp_K \approx \sum_{p_1 \in \mathcal{W}} \sum_{p_2 \in \mathcal{W}} \cdots \sum_{p_K \in \mathcal{W}} (\alpha_l - \alpha_{K-l}) \tag{203}$$

which approximates integration over a K-dimensional grid $\mathcal{W}^K$.

The idea is then to sample points from this K-dimensional grid $\mathcal{W}^K$ and compute an empirical mean of the integration based on the sample probabilities for $p_1, \ldots, p_K$ from $\mathcal{W}^K$ as the approximation of the integration in the objective.

Let $(s_1, s_2, \ldots, s_K)$ be randomly sampled probability values from $\mathcal{W}^K$ and we want to compute $(\alpha_l - \alpha_{K-l})$ for all l based on $(p_1, \ldots, p_K) = (s_1, \ldots, s_K)$. To apply the rectangular rule, since the grid of probabilities is K-dimensional, the weight of $(\alpha_l - \alpha_{K-l})$ in the approximate integration is θ^K. Furthermore, observe

that α_l is the pmf at l from a Poison Binomial distribution in our case, and PoissonBinomial$(p_1, \ldots, p_K) \overset{dist.}{\sim}$ PoissonBinomial$(\pi(p_1, \ldots, p_K))$, where π denotes a permutation of $p_1, \ldots, p_K$ and $\overset{dist.}{\sim}$ denotes "the same distribution". Hence, with a single probability sample $(s_1, \ldots, s_K)$, we can indeed compute $\alpha_l - \alpha_{K-l}$ for each l at $K!$ points from the grid $\mathcal{W}^K$, since they all have the same value. Therefore, we should set the weight of $\alpha_l - \alpha_{K-l}$ in the approximate integration as $w = \theta^K \cdot K!$. Furthermore, since the order of $(p_1, \ldots, p_K)$ does not affect the objective value, there is a total of (τ choose K with replacement) $= \binom{\tau+K-1}{K} := P$ different points in the grid $\mathcal{W}^K$.

In summary, the integration based approximation of the objective proceeds as follows:

Algorithm 5 Integration Based Approximation of the Optimization Objective

Input: # mechanisms $K \in \mathbb{N}$, # iterations $T = 10000 \in \mathbb{N}$, noise function $\gamma : \{0, 1, \ldots, K\} \to [0, 1]$, $\tau = 50$: # samples between $[0.5, 1)$ to form the set $\mathcal{W}$
$\theta \leftarrow 0.5/\tau$ distance between samples
$w \leftarrow \theta^K \cdot K!$
$P \leftarrow \binom{\tau+K-1}{K}$
for $t = 1, 2, \ldots, T$ **do**
  Sample probabilities $(s_1, s_2, \ldots, s_K) \sim \mathcal{W}^K$
  $\widehat{\mathcal{L}} \sim$ PoissonBinomial$(s_1, s_2, \ldots, s_K)$
  $\hat{\alpha}_l \leftarrow \Pr[\widehat{\mathcal{L}} = l], \forall l \in \{0, 1, \ldots, K\}$
  $g_t \leftarrow -\frac{1}{2} \sum_{l=\frac{K+1}{2}}^{K} w \cdot (\hat{\alpha}_l - \hat{\alpha}_{K-l}) \cdot \gamma(l)$
end for
Return $\frac{P}{N} \sum_{t=1}^{T} g_t$

C.3 Reducing # Constraints from ∞ to a Polynomial Set

Lemma C.1 (Restatement of Lemma 5.1)**.** *Consider using DaRRM (Algorithm 1) to solve Problem 1.1 and let f be the privacy cost objective as defined in Lemma 3.4. Given an arbitrary noise function γ, let the worst case probabilities be*

$$(p_1^*, \ldots, p_K^*, p_1'^*, \ldots, p_K'^*) = \underset{\{(p_i, p_i')\}_{i=1}^K}{\arg\max} f(p_1, \ldots, p_K, p_1', \ldots, p_K'; \gamma)$$

Then, each pair $(p_i^*, p_i'^*), \forall i \in [K]$ satisfies

$$(p_i^*, p_i'^*) \in \{(0,0), (1,1), (0, \Delta), (\Delta, 0), (1-\Delta, 1), (1, 1-\Delta), (\frac{e^\epsilon + \Delta}{e^\epsilon + 1}, \frac{1-\Delta}{e^\epsilon + 1}), (\frac{1-\Delta}{e^\epsilon + 1}, \frac{e^\epsilon + \Delta}{e^\epsilon + 1})\}$$

Furthermore, when $\delta > 0$, there exists a finite vector set $\mathcal{P}$ of size $O(K^7)$ such that if $\beta = \max_{\{(p_i, p_i')\}_{i=1}^K \in \mathcal{P}} f(p_1, \ldots, p_K, p_1', \ldots, p_K'; \gamma)$, then $f(p_1^*, \ldots, p_K^*, p_1'^*, \ldots, p_K'^*; \gamma) \leq \beta$. When $\delta = 0$, the size of $\mathcal{P}$ can be reduced to $O(K^3)$.

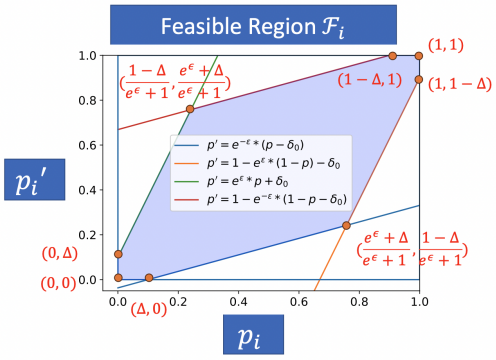

Figure 5: An illustration of the feasible region $\mathcal{F}_i$.

Proof. **Part I: Reducing # privacy constraints from ∞ to exponentially many.**

Consider (p_i, p_i') for an arbitrary $i \in [K]$ and fixing $(p_j, p_j'), \forall j \neq i$. Given any noise function γ, recall the privacy cost objective $f(p_1, \ldots, p_K, p_1', \ldots, p_K'; \gamma)$ (see Lemma 3.4), is

$$f(p_1, \ldots, p_K, p_1', \ldots, p_K'; \gamma) = \sum_{l=0}^{\frac{K-1}{2}} (e^{m\epsilon}\alpha_l' - \alpha_l) \cdot \gamma(l) + \sum_{l=\frac{K+1}{2}}^{K} (\alpha_l - e^{m\epsilon}\alpha_l') \cdot \gamma(l)$$

and the privacy constraints are of the form

$$f(p_1, \ldots, p_K, p_1', \ldots, p_K'; \gamma) \leq e^{m\epsilon} - 1 + 2\delta$$

where recall that $\alpha_l = \Pr[\mathcal{L}(\mathcal{D}) = l]$ is a function of $\{p_i\}_{i=1}^K$ and $\alpha_l' = \Pr[\mathcal{L}(\mathcal{D}') = l]$ is a function of $\{p_i'\}_{i=1}^K$, $\forall l \in \{0, 1, \ldots, K\}$ and $\mathcal{L}(\mathcal{D})$, $\mathcal{L}(\mathcal{D}')$ are the sum of observed outcomes on neighboring datasets $\mathcal{D}$ and $\mathcal{D}'$. By Lemma 3.4, γ needs to make the above privacy constraint hold for all possible $\{(p_i, p_i')\}_{i=1}^K$ to make DaRRM_γ $(m\epsilon, \delta)$-differentially private. This is equivalent to saying, γ needs to ensure $\max_{\{(p_i,p_i')\}_{i=1}^K} f(p_1, \ldots, p_K, p_1', \ldots, p_K'; \gamma) \leq e^{m\epsilon} - 1 + 2\delta$.

Notice that the sum of observed outcomes follows a Poisson Binomial distribution, i.e., $\mathcal{L}(\mathcal{D}) \sim \text{PoissonBinomial}(p_1, \ldots, p_K)$ and $\mathcal{L}(\mathcal{D}') \sim \text{PoissonBinomial}(p_1', \ldots, p_K')$. Hence, by the pmf of the Poisson Binomial distribution[6], the privacy cost objective f is linear in each p_i and p_i', fixing all (p_j, p_j'), $\forall j \neq i$. Since each mechanism M_i is (ϵ, Δ)-differentially private, by definition, (p_i, p_i') satisfies all of the following:

$$p_i \leq e^\epsilon p_i' + \Delta, \quad p_i' \leq e^\epsilon p + \Delta$$
$$1 - p_i \leq e^\epsilon (1 - p_i') + \Delta, \quad 1 - p_i' \leq e^\epsilon (1 - p_i) + \Delta$$

That is, (p_i, p_i') lies in a feasible region $\mathcal{F}_i$ (see Figure 5). Note the constraints on (p_i, p_i'), that is, the boundaries of $\mathcal{F}_i$, are linear in p_i and p_i'. And so the optimization problem $(p_i^*, p_i'^*) = \arg\max_{(p_i,p_i')} f(p_1, \ldots, p_K, p_1', \ldots, p_K'; \gamma)$, which finds the worst case probabilities in (p_i, p_i'), is a Linear Programming (LP) problem in (p_i, p_i') for $i \in [K]$. This implies $(p_i^*, p_i'^*)$ has to be on one of the eight corners of $\mathcal{F}_i$ — that is $(p_i^*, p_i'^*) \in \{(0,0), (1,1), (0,\Delta), (\Delta,0), (1-\Delta,1), (1,1-\Delta), (\frac{e^\epsilon+\Delta}{e^\epsilon+1}, \frac{1-\Delta}{e^\epsilon+1}), (\frac{1-\Delta}{e^\epsilon+1}, \frac{e^\epsilon+\Delta}{e^\epsilon+1})\} := \mathcal{C}$. Since all (p_i, p_i') and (p_j, p_j'), for $i \neq j$, are independent, we can search for the worst case probabilities by searching for $(p_i^*, p_i'^*) \in \mathcal{C}$, instead of searching for $(p_i, p_i') \in \mathcal{F}_i, \forall i \in [K]$. Therefore, the infinitely many privacy constraints are now reduced to only 8^K to optimize for the best γ function that maximizes the utility of DaRRM_γ, while ensuring the output is $m\epsilon$-differentially private.

Part II: Reducing # privacy constraints from exponentially many to a polynomial set.

To further reduce the number of privacy constraints in optimization, observe that the Poisson Binomial distribution is invariant under the permutation of its parameters. That is, $\text{PoissonBinomial}(p_1, \ldots, p_K) \overset{dist.}{\sim}$

[6]See, e.g. https://en.wikipedia.org/wiki/Poisson_binomial_distribution, for the pmf of Poisson Binomial distribution.

PoissonBinomial($\pi(p_1, \ldots, p_K)$), for some permutation π and $\overset{dist.}{\sim}$ means "follows the same distribution". Similarly, PoissonBinomial$(p'_1, \ldots, p'_K) \overset{dist.}{\sim}$ PoissonBinomial$(\pi(p'_1, \ldots, p'_K))$.

The above observation implies if we have one privacy constraint $f(p_1 = v_1, \ldots, p_K = v_K, p'_1 = v'_1, \ldots, p'_K = v'_K; \gamma) \leq e^{m\epsilon} - 1 + 2\delta$, for some $\{(v_i, v'_i)\}_{i=1}^K \in \mathcal{C}^K$, then any privacy constraint $f(p_1 = s_1, \ldots, p_K = s_K, p'_1 = s'_1, \ldots, p'_K = s'_K; \gamma) \leq e^{m\epsilon} - 1 + 2\delta$, where $(s_1, \ldots, s_K) = \pi_1(v_1, \ldots, v_K)$, $(s'_1, \ldots, s'_K) = \pi(v'_1, \ldots, v'_K)$, for permutations π_1 and π_2, is redundant.

Therefore, there is a vector set $\mathcal{P}$, where each probability vector $(p_1, \ldots, p_K, p'_1, \ldots, p'_K)$ in $\mathcal{P}$ is constructed by setting $(p_1, p'_1), (p_2, p'_2), \ldots, (p_K, p'_K) = (v_1, v_2, \ldots, v_K)$, where $v_i \in \mathcal{C}, \forall i \in [K]$, such that vectors constructed by $(p_1, p'_1), (p_2, p'_2), \ldots, (p_K, p'_K) = \pi(v_1, v_2, \ldots, v_K)$ is not in $\mathcal{P}$. Note $|\mathcal{P}|$ = (8 chooses K with replacement) $= \binom{K+8-1}{K} = O(K^7)$. If we can restrict our search for the worst case probabilities to this set $\mathcal{P}$ — that is, solving for $\beta := \max_{\{(p_i, p'_i)\}_{i=1}^K \in \mathcal{P}} f(p_1, \ldots, p_K, p'_1, \ldots, p'_K; \gamma)$, then $f(p^*_1, \ldots, p^*_K, p'^*_1, \ldots, p'^*_K; \gamma) \leq \beta$. This implies we only need $O(K^7)$ privacy constraints to optimize for the best noise function γ in DaRRM, while making sure DaRRM$_\gamma$ is $m\epsilon$-differentially private.

Note if $\Delta = 0$, i.e., the mechanism M_i's are pure differentially private, the feasible region $\mathcal{F}_i$ in which (p_i, p'_i) lies has only 4 corners instead of 8. This implies $(p^*_i, p'^*_i) \in \mathcal{C} = \{(0,0), (1,1), (\frac{e^\epsilon}{e^\epsilon+1}, \frac{1}{e^\epsilon+1}), (\frac{1}{e^\epsilon+1}, \frac{e^\epsilon}{e^\epsilon+1})\}$. Hence, in this case, $|\mathcal{P}|$ = (4 choose K with replacement) $= \binom{K+4-1}{K} = O(K^3)$, which implies we only need $O(K^3)$ privacy constraints to optimize for the best noise function γ in DaRRM.

□

D Full Experiment Results

D.1 Optimized γ in Simulations

D.1.1 Comparison Using General Composition

The general composition (Theorem 2.3) indicates less total privacy loss than simple composition (Theorem 2.2) when the number of folds, m, is large, or when the failure probability δ is large. To enable meaningful comparison against general composition, we consider a larger K and a larger failure probability δ.

Consider $K = 35, \epsilon = 0.1, \Delta = 10^{-5}$. By general composition, if one outputs the majority of M subsampled mechanisms for some $M < K$, the majority output is $(\epsilon_{opt}, \delta_{opt})$-differentially private, where

$$\epsilon_{opt} = \min\left\{M\epsilon, \frac{(e^\epsilon - 1)\epsilon M}{e^\epsilon + 1} + \epsilon\sqrt{2M\log(e + \frac{\sqrt{M\epsilon^2}}{\delta'})}, \frac{(e^\epsilon - 1)\epsilon M}{e^\epsilon + 1} + \epsilon\sqrt{2M\log(\frac{1}{\delta'})}\right\}, \quad \delta_{opt} = 1 - (1-\delta)^M(1-\delta')$$

for some $\delta' \geq 0$. We set this as the privacy guarantee of all majority ensembling algorithms. That is, if we want the majority output to be $(m\epsilon, \delta)$-differentially private, we set

$$m = \frac{\epsilon_{opt}}{\epsilon} = \min\left\{M, \frac{(e^\epsilon - 1)M}{e^\epsilon + 1} + \sqrt{2M\log(e + \frac{\sqrt{M\epsilon^2}}{\delta'})}, \frac{(e^\epsilon - 1)M}{e^\epsilon + 1} + \sqrt{2M\log(\frac{1}{\delta'})}\right\}$$

and $\delta = 1 - (1-\delta)^M(1-\delta')$ accordingly. The parameters τ and λ to compute p_{const} in RR (see Section A.1) are set to be

$$\tau = \min\left\{K, \frac{(e^\epsilon - 1)K}{e^\epsilon + 1} + \sqrt{2K\log(e + \frac{\sqrt{K\epsilon^2}}{\delta'})}, \frac{(e^\epsilon - 1)K}{e^\epsilon + 1} + \sqrt{2K\log(\frac{1}{\delta'})}\right\}$$

and $\lambda = 1 - (1-\delta)^K(1-\delta')$.

In the experiments, we consider $M = \{10, 13, 15, 20\}$ and $\delta' = 0.1$; and γ_{opt} is computed using a uniform prior $\mathcal{T}$.

All values of the parameters of the private ensembling algorithms we use in the experiment are listed in the table:

# Subsampled mechanisms	M	10	13	15	20
Privacy allowance	m	6.4521	7.5742	8.2708	9.8823
Parameter of constant γ	τ	14.0328	14.0328	14.0328	14.0328
Parameter of constant γ	λ	0.1003	0.1003	0.1003	0.1003
Overall privacy loss	$m\epsilon$	0.6452	0.7574	0.8271	0.9882
Overall failure probability	δ	0.1001	0.1001	0.1001	0.1002

Table 3: All parameter values. Note that all the private ensembling algorithms we compare in the experiment is required to be $(m\epsilon, \delta)$-differentially private. Here, $K = 35, \epsilon = 0.1, \Delta = 10^{-5}$ and $\delta' = 0.1$.

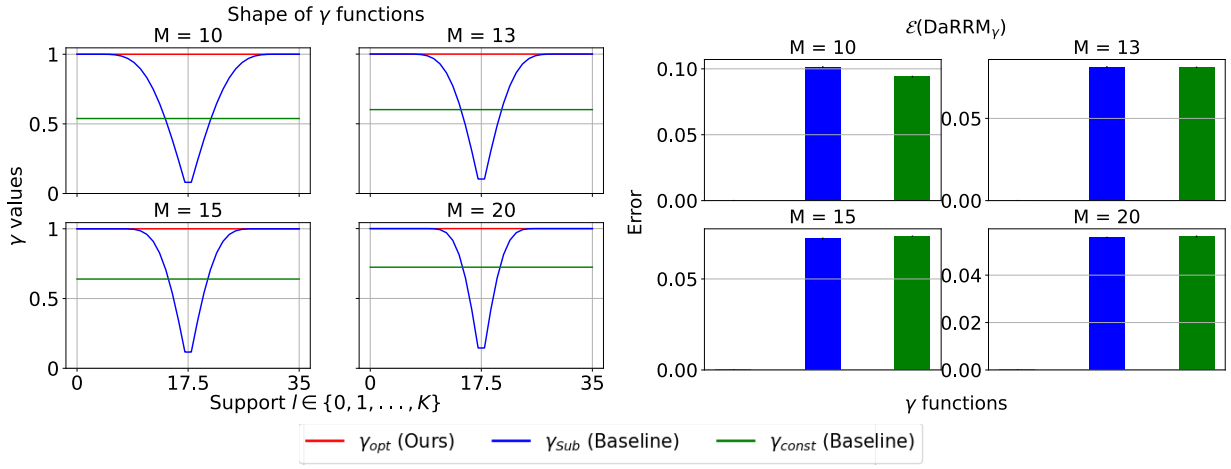

Figure 6: Plots of the shape and $\mathcal{E}(\mathsf{DaRRM}_\gamma)$ of different γ functions: the optimized γ_{Sub}, and the baselines γ_{Sub} (corresponding to subsampling) and γ_{const} (corresponding to RR). Here, $K = 35, M \in \{10, 13, 15, 20\}$, $\Delta = 10^{-5}$, $\epsilon = 0.1$, $\delta' = 0.1$.

D.1.2 Comparison in Pure Differential Privacy Settings

Consider the pure differential privacy setting, where $\Delta = \delta = 0$. Note in this setting, it is known that simple composition is tight.

To compute an optimized γ_{opt} in DaRRM, since we have shown the number of constraints is $O(K^3)$ if $\Delta = \delta = 0$ (see Lemma 5.1), we can set K to be larger. Here, we present results for $K \in \{11, 101\}$ and $\epsilon = 0.1$.

Again, we compare the shape of different γ and the corresponding $\mathcal{E}(\mathsf{DaRRM}_\gamma)$ under those γ functions, fixing the total privacy loss to be $m\epsilon$. γ_{opt} is computed using a uniform prior $\mathcal{T}$.

Since the subsampling mechanism from Section 4 with privacy amplification applies to this setting, we compare four different γ noise functions here:

1. γ_{opt} (Ours): optimized γ function using our optimization framework
2. γ_{Sub} (Baseline): the γ function that corresponds to outputting the majority of m out K subsampled mechanisms
3. γ_{DSub} (Baseline): the γ function that corresponds to outputting $2m - 1$ subsampled mechanisms from Theorem 4.1, aka., Double Subsampling (DSub)
4. γ_{const} (Baseline): the constant γ function that corresponds to the classical Randomized Response (RR) algorithm

Setting 1. $K = 11$, $m \in \{1, 3, 5, 7, 9, 11\}$.

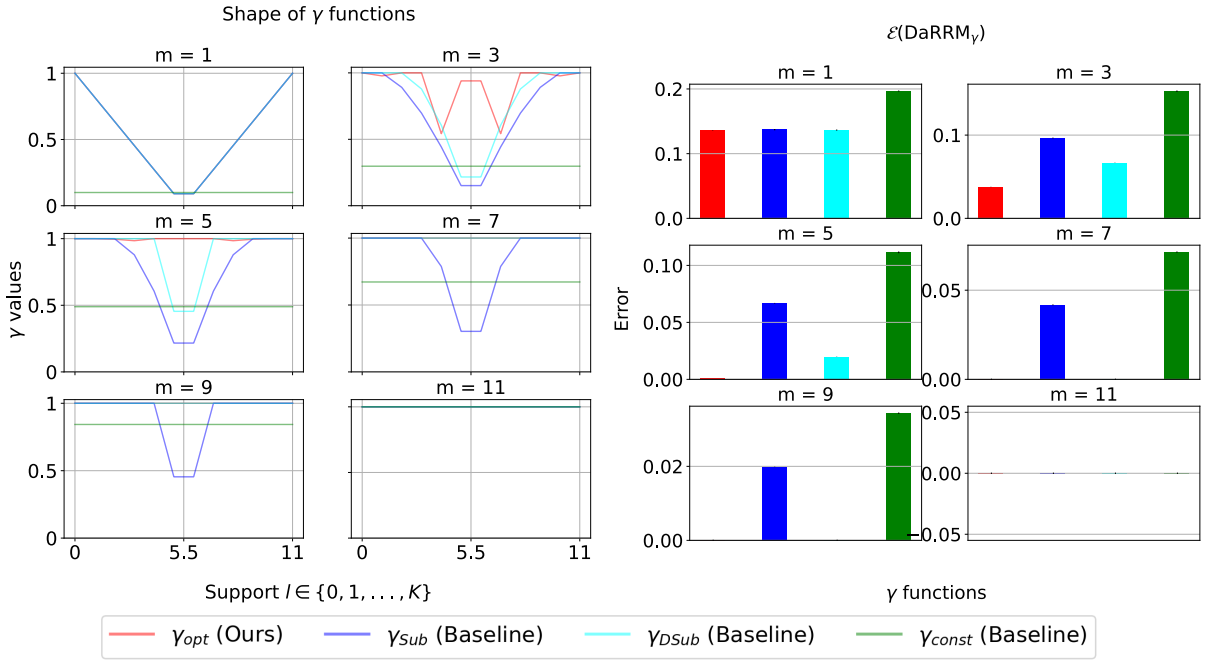

Figure 7: Plots of shape and $\mathcal{E}(\mathsf{DaRRM}_\gamma)$ of different γ functions: the optimized γ_{Opt}, the baselines γ_{Sub} and γ_{DSub} (Theorem 4.1), and the constant γ_{const} (corresponding to RR). Here, $K = 11, m \in \{1, 3, 5, 7, 9, 11\}$, $\epsilon = 0.1$ and $\delta = \Delta = 0$. Note when $m \in \{7, 9\}$, the cyan line (γ_{DSub}) and the red line (γ_{opt}) overlap. When $m = 11$, all lines overlap. Observe that when $m \geq \frac{K+1}{2}$, that is, $m \in \{7, 9, 11\}$ in this case, the above plots suggest both γ_{opt} and γ_{DSub} achieve the minimum error at 0. This is consistent with our theory.

Setting 2. $K = 101, m \in \{10, 20, 30, 40, 60, 80\}$.

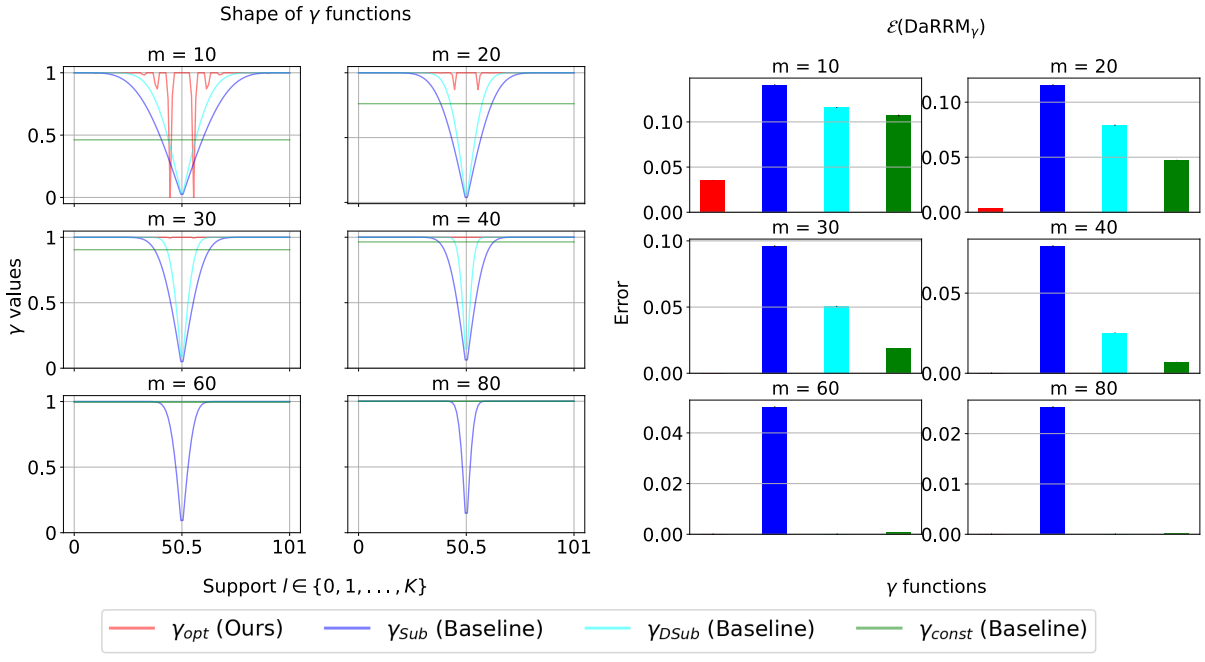

Figure 8: Plots of shape and $\mathcal{E}(\mathsf{DaRRM}_\gamma)$ of different γ functions: the optimized γ_{Opt}, the baselines γ_{Sub} and γ_{DSub} (Theorem 4.1), and the constant γ_{const} (corresponding to RR). Here, $K = 101, m \in \{10, 20, 30, 40, 60, 80\}$, $\epsilon = 0.1$ and $\delta = \Delta = 0$.

D.1.3 Comparison Using Different Prior Distributions

When optimizing γ that maximizes the utility in DaRRM, recall that the objective takes an expectation over p_i's for $p_i \sim \mathcal{T}$, where $\mathcal{T}$ is some distribution and $p_i = \Pr[M_i(\mathcal{D}) = 1]$. The previous experiments assume we do not have access to any prior knowledge about p_i's and hence $\mathcal{T}$ is the uniform distribution, i.e., Uniform($[0, 1]$). However, when one has knowledge about the mechanisms, one can set a proper prior $\mathcal{T}$ to further maximize the utility of DaRRM.

In this section, let $\mathcal{T}_U$ denote Uniform($[0, 1]$) and we present results considering a different prior distribution, which we call $\mathcal{T}_P$, as follows. Suppose our prior belief is that each mechanism M_i has a clear tendency towards voting 0 or 1, i.e., p_i is far from 0.5. Let $\mathcal{T}_P$ be Uniform($[0, 0.3] \cup [0.7, 1]$).

To optimize γ under $\mathcal{T}_P$, we change the approximate optimization objective in Eq. 202, which optimizes γ under $\mathcal{T}_U$, to be the following,

$$-\frac{1}{2} \sum_{l=\frac{K+1}{2}}^{K} \int_{0.7}^{1} \int_{0.7}^{1} \cdots \int_{0.7}^{1} (\alpha_l - \alpha_{K-l}) dp_1 dp_2 \ldots dp_K \cdot \gamma(l) \tag{204}$$

Setting. $K = 11, m \in \{3, 5\}$, $\epsilon = 0.1$, $\delta = \Delta = 0$.

We compare the shape and $\mathcal{E}(\mathsf{DaRRM}_\gamma)$ of different γ functions:

1. γ_{opt-U} denote the γ function optimized under $p_i \sim \mathcal{T}_U$
2. γ_{opt-P} denote the γ function optimized under $p_i \sim \mathcal{T}_P$
3. γ_{Sub}, corresponding to the subsampling baseline
4. γ_{const}, corresponding to the RR baseline

Note when we compute the error, we take the expectation w.r.t. the actual p_i distributions, regardless of the prior used to optimize γ. In the experiments, we consider three different actual p_i distributions:"

1. "Actual: Uniform($[0, 1]$)": $p_i \sim \mathcal{T}_U, \forall i \in [K]$
2. "Actual: $p_i = 0.5$": $p_i = 0.5, \forall i \in [K]$
   This setting implies the mechanisms do not have a clear majority
3. "Actual: Uniform($[0, 0.1]$)": $p_i \sim$ Uniform($[0, 0.1]$), $\forall i \in [K]$
   This setting implies the mechanisms have a clear majority (i.e., 0)

Since our prior $\mathcal{T}_P$ is closer to Uniform($[0, 0.1]$) (i.e., there is a clear majority), we would expect $\mathcal{E}(\mathsf{DaRRM}_{\gamma_{opt-P}})$ to be the lowest when $p_i \sim$ Uniform$[0, 0.1]$, but to be higher than $\mathcal{E}(\mathsf{DaRRM}_{\gamma_{opt-U}})$ when $p_i \sim$ Uniform($[0, 1]$) or $p_i = 0.5$. The results are presented in Figure 9.

D.2 Private Semi-Supervised Knowledge Transfer

D.2.1 More Details about the Baseline GNMax Papernot et al. (2018)

The GNMax aggregation mechanism for majority ensembling of *non-private* teachers proceeds as follows (Section 4.1 of Papernot et al. (2018)): on input x,

$$M_\sigma(x) = \arg\max_i \{n_i(x) + \mathcal{N}(0, \sigma^2)\} \tag{205}$$

where $n_i(x)$ is # teachers who vote for class i.

How to set σ in GNMax?

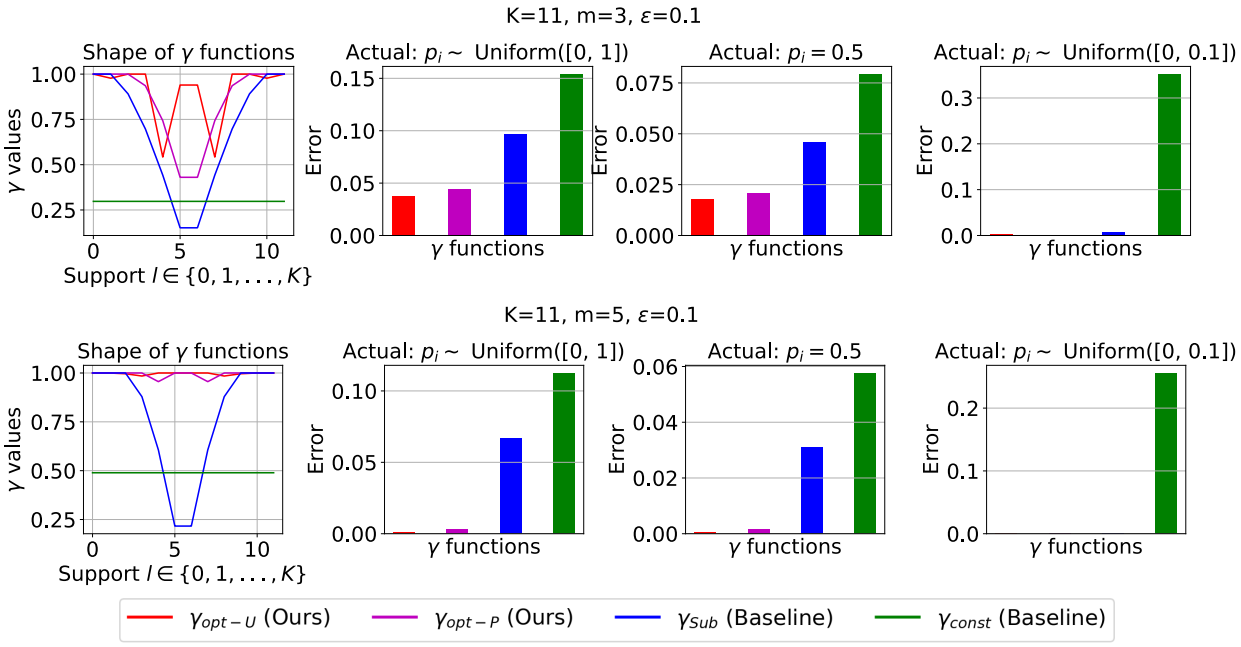

Figure 9: Comparison of the shape and $\mathcal{E}(\mathsf{DaRRM}_\gamma)$ of different γ functions: 1) γ optimized under prior $\mathcal{T}_U$, 2) γ optimized under prior $\mathcal{T}_P$, 3) γ_{Sub} (corresponding to the subsampling baseline) and 4) γ_{const} (corresponding to the RR baseline). Here, $K = 11, m \in \{3, 5\}, \epsilon = 0.1$. Observe that if the prior $\mathcal{T}_P$ used in optimizing γ is closer to the actual distribution of p_i's, there is additional utility gain (i.e., decreased error); otherwise, we slightly suffer a utility loss (i.e., increased error), compared to optimize γ under the $\mathcal{T}_U$ prior. Furthermore, regardless of the choice of the prior distribution $\mathcal{T}$ in optimizing γ, DaRRM_γ with an optimized γ achieves a lower error compared to the the baselines.

Section 4.1 of Papernot et al. (2018) states the GNMax mechanism is $(\lambda, \lambda/\sigma^2)$-Renyi differentially private (RDP), for all $\lambda \geq 1$. RDP bounds can be converted to DP bounds as follows:

Theorem D.1 (RDP to DP (Theorem 5 of Papernot et al. (2018))). *If a mechanism M guarantees (λ, ϵ)-RDP, then M guarantees $(\epsilon + \frac{\log 1/\delta}{\lambda - 1}, \delta)$-differential privacy for $\delta \in (0, 1)$.*

Therefore, GNMax with parameter σ^2 guarantees $(\frac{\lambda}{\sigma^2} + \frac{\log 1/\delta}{\lambda-1}, \delta)$-differential privacy, $\forall \lambda \geq 1$. Given m, ϵ, Δ, we want to choose λ and σ^2 here so that the output of GNMax is $(m\epsilon, m\Delta)$-differentially private. Here, $\delta = m\Delta$.

We first obtain a valid range of λ. Since $m\epsilon \geq 0$, $\frac{\lambda}{\sigma^2} + \frac{\log 1/\delta}{\lambda-1} \geq 0$ and so $\lambda \geq \frac{\log 1/\delta}{m\epsilon} + 1 := \lambda_{min}$. And $\sigma^2 = \frac{\lambda}{m\epsilon - \frac{\log 1/\delta}{\lambda-1}}$. Since the smaller σ^2 is, the higher the utility, we perform a grid search over $\lambda \in [\lambda_{min}, 500]$, with discretized λ values of equal distance 0.5, to find the minimum σ^2_{min}. For the $(m\epsilon, m\Delta)$ values used in the experiments, we observe σ^2 decreases first and then increases as λ increases, as shown in Figure 10. The λ and σ_{min} values in the RDP bound of Gaussian noise to compute the privacy loss of GNMax's output we use in the experiments are presented in Table 4.

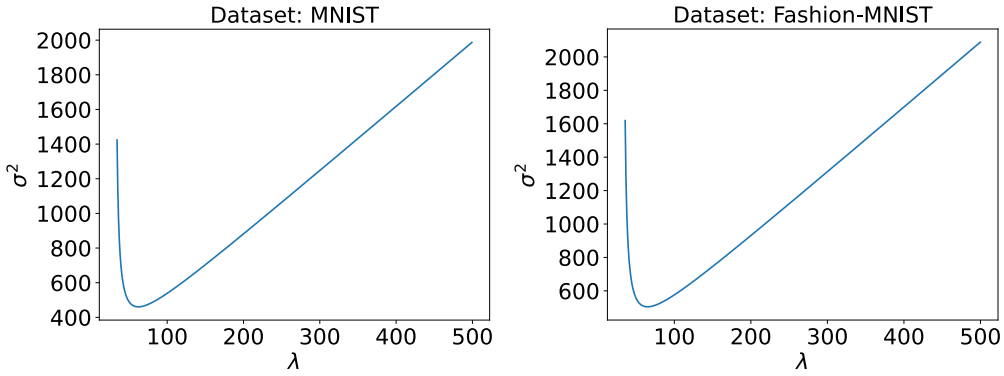

Figure 10: Plots of λ vs. σ^2 in the Gaussian RDP privacy bound. The goal is to choose a λ value that minimizes σ^2. It is not hard to see the value of σ^2 decreases at first and then increases as λ increases.

	Privacy Loss Per Query $(m\epsilon, m\Delta)$	λ	σ_{min}
MNIST	$(0.2676, 0.0003)$	34.31	21.46
Fashion-MNIST	$(0.2556, 0.0003)$	35.74	22.46

Table 4: Parameters of the RDP bound of Gaussian noise to compute the privacy loss of GNMax's output.

A Note on the Data-dependent Privacy Loss Bound

Papernot et al. (2018) gives a potentially tighter data-dependent bound on the privacy loss using GNMax to output the majority of non-private teacherss votes. We give a clean pseudo-code on computing the data-dependent privacy loss bound in Algorithm 6, based on the lemmas and theorems in Papernot et al. (2018). Given privacy parameters σ, λ and the teacher votes per class $\{n_i\}_{i=1}^C$ for C classes, the data-dependent bound can be empirically evaluated and compared against the Gaussian privacy loss bound. The smaller one is the final privacy loss. We empirically find that the condition of the data-dependent bound (line 8 in Algorithm 6) is not satisfied when K and the number of classes C are small, e.g., $K = 11, C = 2$ as in our case, even if all teachers agree on the same output. And so in the experiments, we can only apply the Gaussian privacy loss bound (line 14).

D.2.2 Additional Results for Private Semi-Supervised Knowledge Transfer

$m = 1$.

Algorithm 6 Compute Tighter Privacy Loss

Input: Std. of Gaussian noise σ, Privacy parameter λ, # teachers K, # classes C, # votes per class $\{n_i\}_{i=1}^{C}$
$\mathcal{B} \leftarrow \{\}$ bound candidates
for $i = 1, 2, \ldots, K$ **do**
  $q^{(i)} \leftarrow \frac{1}{2}\sum_{i \neq i^*} \text{erfc}(\frac{n_{i^*} - n_i}{2\sigma})$
  $\mu_2^{(i)} \leftarrow \sigma \cdot \sqrt{\log 1/q^{(i)}}$, $\mu_1^{(i)} \leftarrow \mu_2^{(i)} + 1$
  $\epsilon_1^{(i)} \leftarrow \frac{\mu_1^{(i)}}{\sigma^2}$, $\epsilon_2^{(i)} \leftarrow \frac{\mu_2^{(i)}}{\sigma^2}$
  $q_{ub}^{(i)} \leftarrow \exp((\mu_2^{(i)} - 1)^{\epsilon_2^{(i)}})/(\frac{\mu_1^{(i)}}{\mu_1^{(i)}-1} \cdot \frac{\mu_2^{(i)}}{\mu_2^{(i)}-1})^{\mu_2^{(i)}}$
  if $q^{(i)} < 1$ and $\mu_1^{(i)} \geq \lambda$ and $\mu_2 > 1$ and $q^{(i)} \leq q_{ub}^{(i)}$ **then**
    $A^{(i)} \leftarrow (1 - q^{(i)})/(1 - q^{(i)} \cdot \exp(\epsilon_2^{(i)})^{\frac{\mu_2^{(i)}-1}{\mu_2^{(i)}}})$
    $B^{(i)} \leftarrow \exp(\epsilon_1^{(i)})/(q^{(i)})^{\frac{1}{\mu_1^{(i)}-1}}$
    $\text{DataDependentBound} \leftarrow \frac{1}{\lambda-1} \cdot \left((1 - q^{(i)}) \cdot (A^{(i)})^{\lambda-1} + q^{(i)} \cdot (B^{(i)})^{\lambda-1}\right)$
    $\mathcal{B} \leftarrow \mathcal{B} \cup \text{DataDependentBound}$
  else
    $\text{GaussianBound} \leftarrow \frac{\lambda}{\sigma^2}$
    $\mathcal{B} \leftarrow \mathcal{B} \cup \text{GaussianBound}$
  end if
end for
Return $\min \mathcal{B}$

Dataset	# Queries	Privacy loss per query $(\epsilon_{query}, \delta_{query})$	Total privacy loss over Q queries $(\epsilon_{total}, \delta_{total})$
MNIST	$Q = 20$	$(0.0892, 0.0001)$	$(1.704, 0.002)$
	$Q = 50$		$(2.837, 0.005)$
	$Q = 100$		$(4.202, 0.010)$
Fashion MNIST	$Q = 20$	$(0.0852, 0.0001)$	$(1.620, 0.002)$
	$Q = 50$		$(2.695, 0.005)$
	$Q = 100$		$(3.988, 0.010)$

Table 5: The privacy loss per query to the teachers and the total privacy loss over Q queries. Note the total privacy loss is computed by general composition, where we set $\delta' = 0.0001$.

Dataset	MNIST			Dataset	Fashion-MNIST		
# Queries	GNMax (Baseline)	DaRRM$_{\gamma_{Sub}}$ (Baseline)	DaRRM$_{\gamma_{opt}}$ (Ours)	# Queries	GNMax (Baseline)	DaRRM$_{\gamma_{Sub}}$ (Baseline)	DaRRM$_{\gamma_{opt}}$ (Ours)
$Q = 20$	0.54 (0.11)	0.68 (0.07)	**0.74 (0.08)**	$Q = 20$	0.56 (0.10)	**0.92 (0.05)**	0.89 (0.06)
$Q = 50$	0.51 (0.07)	**0.67 (0.05)**	0.66 (0.05)	$Q = 50$	0.52 (0.05)	0.89 (0.04)	**0.92 (0.03)**
$Q = 100$	0.57 (0.03)	**0.71 (0.03)**	0.69 (0.04)	$Q = 100$	0.56 (0.04)	0.89 (0.04)	**0.91 (0.04)**

Table 6: Accuracy of the predicted labels of Q query samples on datasets MNIST (on the left) and Fashion-MNIST (on the right). We report the mean and one std. in parentheses over 10 random draws of the query samples from the test dataset. Note each prediction on the query sample is $(\epsilon_{total}, \delta_{total})$-differentially private. Note in this case where $m = 1$, by Lemma 3.2, subsampling achieves the optimal error/utility. Hence, there is not much difference in terms of accuracy between DaRRM$_{\gamma_{Sub}}$ and DaRRM$_{\gamma_{opt}}$ as expected.

$m = 5$.

Dataset	# Queries	Privacy loss per query $(\epsilon_{query}, \delta_{query})$	Total privacy loss over Q queries $(\epsilon_{total}, \delta_{total})$
MNIST	$Q = 20$	$(0.4460, 0.0005)$	$(8.920, 0.010)$
	$Q = 50$		$(18.428, 0.025)$
	$Q = 100$		$(28.926, 0.049)$
Fashion MNIST	$Q = 20$	$(0.4260, 0.0005)$	$(8.520, 0.010)$
	$Q = 50$		$(17.398, 0.025)$
	$Q = 100$		$(27.223, 0.049)$

Table 7: The privacy loss per query to the teachers and the total privacy loss over Q queries. Note the total privacy loss is computed by general composition, where we set $\delta' = 0.0001$.

Dataset	MNIST			Dataset	Fashion-MNIST		
# Queries	GNMax (Baseline)	DaRRM$_{\gamma_{Sub}}$ (Baseline)	DaRRM$_{\gamma_{opt}}$ (Ours)	# Queries	GNMax (Baseline)	DaRRM$_{\gamma_{Sub}}$ (Baseline)	DaRRM$_{\gamma_{opt}}$ (Ours)
$Q = 20$	0.73 (0.11)	0.76 (0.09)	**0.84 (0.07)**	$Q = 20$	0.72 (0.10)	0.96 (0.04)	**0.97 (0.04)**
$Q = 50$	0.75 (0.07)	0.82 (0.04)	**0.83 (0.04)**	$Q = 50$	0.72 (0.08)	0.96 (0.02)	**0.97 (0.02)**
$Q = 100$	0.72 (0.04)	0.79 (0.05)	**0.83 (0.03)**	$Q = 100$	0.72 (0.06)	**0.97 (0.01)**	**0.97 (0.01)**

Table 8: Accuracy of the predicted labels of Q query samples on datasets MNIST (on the left) and Fashion-MNIST (on the right). We report the mean and one std. in parentheses over 10 random draws of the query samples from the test dataset. Note each prediction on the query sample is $(\epsilon_{total}, \delta_{total})$-differentially private. With the same per query privacy loss (and hence the same total privacy loss over Q samples), DaRRM$_{\gamma_{opt}}$ achieves the highest accuracy compared to the other two baselines.

$m = 7$.

Dataset	# Queries	Privacy loss per query $(\epsilon_{query}, \delta_{query})$	Total privacy loss over Q queries $(\epsilon_{total}, \delta_{total})$
MNIST	$Q = 20$	$(0.6244, 0.0007)$	$(12.488, 0.014)$
	$Q = 50$		$(28.392, 0.035)$
	$Q = 100$		$(45.683, 0.068)$
Fashion MNIST	$Q = 20$	$(0.5964, 0.0007)$	$(11.928, 0.014)$
	$Q = 50$		$(26.738, 0.035)$
	$Q = 100$		$(42.873, 0.068)$

Table 9: The privacy loss per query to the teachers and the total privacy loss over Q queries. Note the total privacy loss is computed by general composition, where we set $\delta' = 0.0001$.

Dataset	MNIST			Dataset	Fashion-MNIST		
# Queries	GNMax (Baseline)	DaRRM$_{\gamma_{Sub}}$ (Baseline)	DaRRM$_{\gamma_{opt}}$ (Ours)	# Queries	GNMax (Baseline)	DaRRM$_{\gamma_{Sub}}$ (Baseline)	DaRRM$_{\gamma_{opt}}$ (Ours)
$Q = 20$	0.79 (0.07)	0.80 (0.09)	**0.85 (0.08)**	$Q = 20$	0.79 (0.07)	0.95 (0.04)	**0.96 (0.04)**
$Q = 50$	0.80 (0.05)	0.82 (0.05)	**0.85 (0.04)**	$Q = 50$	0.79 (0.05)	0.96 (0.03)	**0.97 (0.03)**
$Q = 100$	0.80 (0.04)	0.80 (0.04)	**0.83 (0.03)**	$Q = 100$	0.79 (0.03)	**0.96 (0.02)**	**0.96 (0.02)**

Table 10: Accuracy of the predicted labels of Q query samples on datasets MNIST (on the left) and Fashion-MNIST (on the right). We report the mean and one std. in parentheses over 10 random draws of the query samples from the test dataset. Note each prediction on the query sample is $(\epsilon_{total}, \delta_{total})$-differentially private. With the same per query privacy loss (and hence the same total privacy loss over Q samples), DaRRM$_{\gamma_{opt}}$ achieves the highest accuracy compared to the other two baselines.