SUPPLEMENTARY MATERIAL A
PROOF DETAIL OF UPPER BOUNDS

This part provides more details on proving the privacy amplification bounds in Theorem 1, Corollary 1, and Corollary 2. The first step interprets the output distribution of $\mathcal{K}(c)$ as a mixture of $\mathcal{K}(a)$ and $\mathcal{K}(b)$ for any $a, b, c \in \mathcal{X}$; The second step follows previous works on tight privacy amplification bounds for the classical local privacy [28], [51], reduces the distinguishability analyses on anonymous messages into several binomial counts. The third step gives the final bound based on tail bounds of binomial counts. The last step provides numerically tighter bounds and simplified bounds from the main result.

Step 1 (Mixture interpretation of outputs). Let $R_{a,b}(c)$ denote a parameter that satisfies $R_{a,b}(c) \geq \max_{t\in\mathcal{D}_K} \frac{\mathbb{P}[\mathcal{K}(a)=t]+\mathbb{P}[\mathcal{K}(b)=t]}{\mathbb{P}[\mathcal{K}(c)=t]}$, the Proposition 1 indicates that every $\mathcal{K}(c)$ corresponds to a mixture distribution on $\mathcal{K}(a)$ and $\mathcal{K}(b)$.

Observe that $\max_{t\in\mathcal{D}_K} \frac{\mathbb{P}[\mathcal{K}(a)=t]+\mathbb{P}[\mathcal{K}(b)=t]}{\mathbb{P}[\mathcal{K}(c)=t]}$ is never greater than $\exp(\mathsf{D}(a,c)) + \exp(\mathsf{D}(b,c))$, now set the parameter $R_{a,b}(c) = \exp(\mathsf{D}(a,c)) + \exp(\mathsf{D}(b,c))$, we have the following simple corollary.

***Corollary** 3:* [Mixture of Metric Private Outputs] For any mechanism $\mathcal{K}$ that satisfies local metric D-differential privacy, the distribution of $\mathcal{K}(c)$ is a finite mixture:

$$\mathcal{K}(c) = \begin{cases} \mathcal{K}(a), & \text{with probability } \frac{1}{\exp(\mathsf{D}(a,c))+\exp(\mathsf{D}(b,c))}; \\ \mathcal{K}(b), & \text{with probability } \frac{1}{\exp(\mathsf{D}(a,c))+\exp(\mathsf{D}(b,c))}; \\ \mathcal{W}'_{a,b}(c), & \text{else.} \end{cases} \tag{5}$$

where $\mathcal{W}'_{a,b}(c)$ is some variable with a valid probability distribution.

As for a special case that $c = b$, according to Corollary 3 with $R_{a,b}(b) = \exp(\mathsf{D}(a,c)) + 1)$, we have:

$$\mathcal{K}(b) = \begin{cases} \mathcal{K}(a), & \text{with probability } \frac{1}{\exp(\mathsf{D}(a,b))+1)}; \\ \mathcal{K}(b), & \text{with probability } \frac{1}{\exp(\mathsf{D}(a,b))+1}; \\ \mathcal{W}"_a(b), & \text{else.} \end{cases} \tag{6}$$

Step 2 (Reduction to indistinguishability on binomial counts). Without loss of generality, in the dataset $\{x_i\}_{i=1}^n$ of all users, we assume the value of x_1 changes from a to b, and aim to derive the indistinguishable level between $S = \{x_1 = a, ..., x_n\}$ and $S' = \{x_1 = b, ..., x_n\}$.

According to Corollary 3, every private view of x_i ($i \in [2:n]$) can be seen as a clone of $x_1 = a$ or $x_1 = b$ with probability $\frac{2}{\exp(\mathsf{D}(a,c))+\exp(\mathsf{D}(b,c))}$, hence the true value of x_1 hides among these clones and is less distinguishable. For any $x_i \in \mathcal{X}$, the clone probability is at least $\frac{2}{\max_{c'\in\mathcal{X}} R_{a,b}(c')}$.

Formally, based on the concept of clone [28], for proving the shuffled messages T is (ϵ_c, δ)-indistinguishable (w.r.t. $x_1 = a$ or $x_1 = b$), it suffices to show that the frequencies of $\mathcal{K}(a)$ and $\mathcal{K}(b)$ are (ϵ_c, δ)-indistinguishable, which can be analogously induced as in Lemma 1. The key idea is that the final shuffled messages T is a post-processing information of the frequencies of $\mathcal{K}(a)$ and $\mathcal{K}(b)$, and the privacy guarantee comes from the post-processing property of differential privacy.

The lemma also uses the observation that the $\mathcal{K}(b)$ itself is a clone of $\mathcal{K}(a)$ with probability $\frac{1}{e^{\mathsf{D}(a,b)}}$ (see Equation 6), the Bernoulli variable B records the event of being a clone of $\mathcal{K}(b)$ (or $\mathcal{K}(a)$) for the user 1 with $x_1 = a$ (or $x_1 = b$). The variable C records the total number of clones from users $[2:n]$, and the variable A (or $C - A$) records the number of clones of $\mathcal{K}(a)$ (or $\mathcal{K}(b)$) separately.

Step 3 (Tail bounds on ratio of counts). Recall that the number of clones is $C \sim Binomial(n-1, p)$, according to the Chernoff bound on binomial counts [68], with probability at most $\delta/2$, we have:

$$|C - (n-1)p| < \sqrt{3(n-1)p\log(4/\delta)}.$$

Applying Hoeffding's inequality on the binomial variable A, with probability at most $\delta/2$, we have:

$$|A - C/2| > \sqrt{C\log(4/\delta)/2}.$$

Then, according to the union bound of probabilities, with probability at least $1-\delta$, both $C \geq (n-1)p - \sqrt{3(n-1)p\log(4/\delta)}$ and $|A - C/2| \leq \sqrt{C\log(4/\delta)/2}$ hold.

Define $P' = (A+1, C-A)$ and $Q' = (A, C-A+1)$, since the following equality always hold [28]: $\mathbb{P}[P' = (a,b)] = \mathbb{P}[C = a+b+1] \cdot \mathbb{P}[A = a-1 \mid C = a+b+1] = \mathbb{P}[C = a+b+1] \cdot \mathbb{P}[A = a \mid C = a+b+1] \cdot \frac{a}{b} = \mathbb{P}[Q' = (a,b)] \cdot \frac{a}{b}$, we then have (with probability $1-\delta$):

$$\begin{aligned} \frac{\mathbb{P}[P=(a,b)]}{\mathbb{P}[Q=(a,b)]} &= \frac{a}{b} = \frac{A+1}{C-A} \\ &\leq \frac{C/2 + \sqrt{C\log(4/\delta)/2} + 1}{C/2 - \sqrt{C\log(4/\delta)/2}}. \end{aligned} \tag{7}$$

Similarly, we have $\frac{\mathbb{P}[Q=(a,b)]}{\mathbb{P}[P=(a,b)]} = \frac{C-A}{A+1} \leq \frac{C/2+\sqrt{C\log(4/\delta)/2}}{C/2-\sqrt{C\log(4/\delta)/2}+1} < \frac{C/2+\sqrt{C\log(4/\delta)/2}+1}{C/2-\sqrt{C\log(4/\delta)/2}}$. Because $\frac{C/2+\sqrt{C/2\log(4/\delta)}+1}{C/2-\sqrt{C/2\log(4/\delta)}}$ decreases with C when $C \geq 1 + \frac{1+\sqrt{1+2\log(4/\delta)}}{\log(4/\delta)}$, we get the bound by replacing C as the lower $\delta/2$-tail $\Omega = (n-1)p - \sqrt{3(n-1)p\log(4/\delta)}$. Hence P' and Q' are $(\log(1 + \frac{2\sqrt{\Omega/2\log(4/\delta)}+1}{\Omega/2-\sqrt{\Omega/2\log(4/\delta)}}), \delta)$-indistinguishable. Now consider $P = (A+B, C-A+1-B)$ and $Q = (A+1-B, C-A+B)$, since $B \sim Bernoulli(q)$, according to privacy accountant with sub-sampling [69], we have P and Q are $(\log(1 + q \cdot \frac{2\sqrt{\Omega/2\log(4/\delta)}+1}{\Omega/2-\sqrt{\Omega/2\log(4/\delta)}}), \delta)$-indistinguishable.

Replace p with the lower bound $\frac{2}{\max_{c'\in\mathcal{X}}\exp(\mathsf{D}(a,c))+\exp(\mathsf{D}(b,c))}$ of the clone probability $\frac{2}{\max_{c'\in\mathcal{X}} R_{a,b}(c')}$, then replace q with the mutual clone probability $\frac{\exp(\mathsf{D}(a,b))}{\exp(\mathsf{D}(a,b))+1}$, we arrive the Equation 1 in Theorem 1. Combining the fact that $\max_{c'\in\mathcal{X}} R_{a,b}(c') \leq 2 \cdot e^{\mathsf{D}_{max}}$, we get the Equation 2 in Corollary 1.

Step 4 (Simplification and Tighter Bounds). The previous formula can be simplified if we focus on asymptotic behaviours. When $n > 8\log(4/\delta)\max_{c'\in\mathcal{X}} R_{a,b}(c')$, we

Algorithm 2: Binary search amplification lower bound of shuffle metric differential privacy

Input: A local mechanism $\mathcal{K}$ that satisfies metric D, global privacy parameter δ, number of users n, number of iterations T.

Output: A lower bound of the privacy-amplified $\hat{\mathsf{D}}_n(a,b)$

$\epsilon^L \leftarrow 0$

$\epsilon^R \leftarrow \mathsf{D}(a,b)$

for $t \in [T]$ **do**

  $\epsilon_t \leftarrow \frac{\epsilon^L + \epsilon^R}{2}$

  $\delta_t^0 \leftarrow 0,\ \delta_t^1 \leftarrow 0$

  for $A \in [n],\quad B \in [n-1-A],$ $C \in [n-1-A-B]$ **do**

    $ratio \leftarrow \frac{\mathbb{P}[\#a=A,\#b=B,\#c=C]}{\mathbb{P}[\#a'=A,\#b'=B,\#c'=C]}$

    if $ratio > \exp(\epsilon_t)$ **then**

      $\delta_t^0 \leftarrow \delta_t^0 + \mathbb{P}[\#a=A,\#b=B,\#c=C \mid S] - \exp(\epsilon_t)\mathbb{P}[\#a=A,\#b=B,\#c=C \mid S']$

    if $ratio < \exp(-\epsilon_t)$ **then**

      $\delta_t^1 \leftarrow \delta_t^1 + \mathbb{P}[\#a=A,\#b=B,\#c=C \mid S'] - \exp(\epsilon_t)\mathbb{P}[\#a=A,\#b=B,\#c=C \mid S]$

  if $\max(\delta_t^0, \delta_t^1) < \delta$ **then**

    $\epsilon^R \leftarrow \epsilon_t$

  else

    $\epsilon^L \leftarrow \epsilon_t$

return ϵ^L

have $\Omega/2 - \sqrt{\Omega/2\log(4/\delta)} \geq \Omega/4 \geq (n-1)p/8$, thus the Equation 7 can be simplified to $\log(1 + 8\frac{e^{\mathsf{D}(a,b)}-1}{e^{\mathsf{D}(a,b)}+1} \cdot (\sqrt{\frac{\max_{c'\in\mathcal{X}} R_{a,b}(c')\log(4/\delta)}{2(n-1)}} + \frac{\max_{c'\in\mathcal{X}} R_{a,b}(c')}{2(n-1)}))$ and we get the Equation 3 in Corollary 2. For tighter bounds, one may substitute the lower $\delta/2$-tail bound $\Omega = (n-1)p - \sqrt{3(n-1)p\log(4/\delta)}$ with the exact tail of $Binomial(n-1,p)$.

Supplementary Material B
Numerical Amplification Lower Bounds

This part provides details about the privacy amplification lower bounds we used in Section IV-D. Specifically, fixing δ, we let $\mathsf{Epsilon}_{a,b}(\mathcal{K})$ denote the minimum value ϵ such that $\{\mathcal{K}(x_1 = a), ..., \mathcal{K}(x_n)\}$ and $\{\mathcal{K}(x_1 = b), ..., \mathcal{K}(x_n)\}$ are (ϵ,δ)-indistinguishable for all possible $x_2, ..., x_n \in \mathcal{X}$. We here devise a D-private mechanism $\mathcal{K}_{PEM}$ (padded exponential mechanism [46]) and numerically compute $\mathsf{Epsilon}_{a,b}(\mathcal{K}_{PEM})$. Given an edge (a,b), we pick the element $c \in \mathcal{X}$ that maximizes $\exp(\mathsf{D}(a,c)) + \exp(\mathsf{D}(b,c))$. We define the concrete transition probability matrix of $\mathcal{K}_{PEM}$ in Table I, where the last column lists the probability of outputting a special padded symbol $\perp$. We let ab, ac, bc denote $\mathsf{D}(a,b), \mathsf{D}(a,c), \mathsf{D}(b,c)$ respectively, let $\sum_a$ denote $1 + e^{-ab} + e^{-ac}$, $\sum_b$ denote $1 + e^{-ab} + e^{-bc}$, and $\sum_c$ denote $1 + e^{-ac} + e^{-bc}$. To ensure the probability design satisfies D-privacy, the probability normalization factor N_{abc} of $\mathcal{K}_{PEM}$ is the maximum one among the following three values: $\frac{e^{ab}\max\{\sum_a,\sum_b\}-\min\{\sum_a,\sum_b\}}{e^{ab}-1}$, $\frac{e^{ac}\max\{\sum_a,\sum_c\}-\min\{\sum_a,\sum_c\}}{e^{ac}-1}$, and $\frac{e^{bc}\max\{\sum_b,\sum_c\}-\min\{\sum_b,\sum_c\}}{e^{bc}-1}$. The reason for choosing PEM over exponential mechanism [5], [55] is that the exponential mechanism poses an extra factor $1/2$ on distances and thus loosely exploits local MDP constraints.

	$z=a$	$z=b$	$z=c$	$z=\perp$
$x=a$	$\frac{1}{N_{abc}}$	$\frac{e^{-ab}}{N_{abc}}$	$\frac{e^{-ac}}{N_{abc}}$	$\frac{N_{abc}-1-e^{-ab}-e^{-ac}}{N_{abc}}$
$x=b$	$\frac{e^{-ab}}{N_{abc}}$	$\frac{1}{N_{abc}}$	$\frac{e^{-bc}}{N_{abc}}$	$\frac{N_{abc}-1-e^{-ab}-e^{-bc}}{N_{abc}}$
$x=c$	$\frac{e^{-ac}}{N_{abc}}$	$\frac{e^{-bc}}{N_{abc}}$	$\frac{1}{N_{abc}}$	$\frac{N_{abc}-1-e^{-ac}-e^{-bc}}{N_{abc}}$

TABLE I
PROBABILITY DESIGN OF THE PEM MECHANISM.

Consider $\mathcal{K}_{PEM}$ on two neighboring n-size datasets: $S = [x_1 = a, x_2 = c, ..., x_n = c]$, and $S' = [x_1 = b, x_2 = c, ..., x_n = c]$, we now numerically compute a lower bound of $\mathsf{Epsilon}_{a,b}(\mathcal{K}_{PEM})$. We represent the multiset $\{\mathcal{K}_{PEM}(x_1), ..., \mathcal{K}_{PEM}(x_n)\}$ as a tuple $(\#a, \#b, \#c, \#\perp)$ that denotes the occurrences of $a, b, c, \perp$ in the multiset respectively. Then, we exhaustively compute probabilities $\mathbb{P}[\#a, \#b, \#c, \#\perp \mid S]$ and $\mathbb{P}[\#a, \#b, \#c, \#\perp \mid S']$ for all possible $\#a, \#b, \#c, \#\perp \in [0, n]$, and get the indistinguishable level between $[\#a, \#b, \#c, \#\perp \mid S]$ and $[\#a, \#b, \#c, \#\perp \mid S']$ as (ϵ_{PEM}, δ).

We now focus on the formula of $\mathbb{P}[\#a = A, \#b = B, \#c = C]$. For the probability formula that takes S as the input, according to law of total probability, we get the probability formula:

$$\begin{aligned}&\mathbb{P}[\#a = A, \#b = B, \#c = C \mid S]\\ =\ & p_{a,a}P_{n-1}^{A-1,B,C} + p_{ab}P_{n-1}^{A,B-1,C}\\ &+ p_{ac}P_{n-1}^{A,B,C-1} + (1 - p_{aa} - p_{ab} - p_{ac})P_{n-1}^{A,B,C},\end{aligned}$$

where $P_{n-1}^{A,B,C}$ denote $\binom{n-1}{A}\binom{n-A-1}{B}\binom{n-A-B-1}{C}p_{ca}^A p_{cb}^B p_{cc}^C(1 - p_{ca} - -p_{cb} - p_{cc})^{n-A-B-C-1}$. Similarly, for $\mathbb{P}[\#a = A, \#b = B, \#c = C \mid S']$ that takes S' as the input, we get:

$$\begin{aligned}&\mathbb{P}[\#a = A, \#b = B, \#c = C \mid S']\\ =\ & p_{b,a}P_{n-1}^{A-1,B,C} + p_{bb}P_{n-1}^{A,B-1,C}\\ &+ p_{bc}P_{n-1}^{A,B,C-1} + (1 - p_{ba} - p_{bb} - p_{bc})P_{n-1}^{A,B,C},\end{aligned}$$

Plug the transition probabilities (see Table I) into Algorithm 2, we have the privacy amplification lower bounds. We also present complementary numerical results in Figure 13 for reference.

Supplementary Material C
Proof of Pareto optimality of Algorithm 1

Denote the derived semi-metric with diameter m in lines 7-18 as D^m, we first show every feasible semi-metric D'' with the same diameter (i.e., $\mathsf{D}''_{max} = m$) is dominated by D^m. We

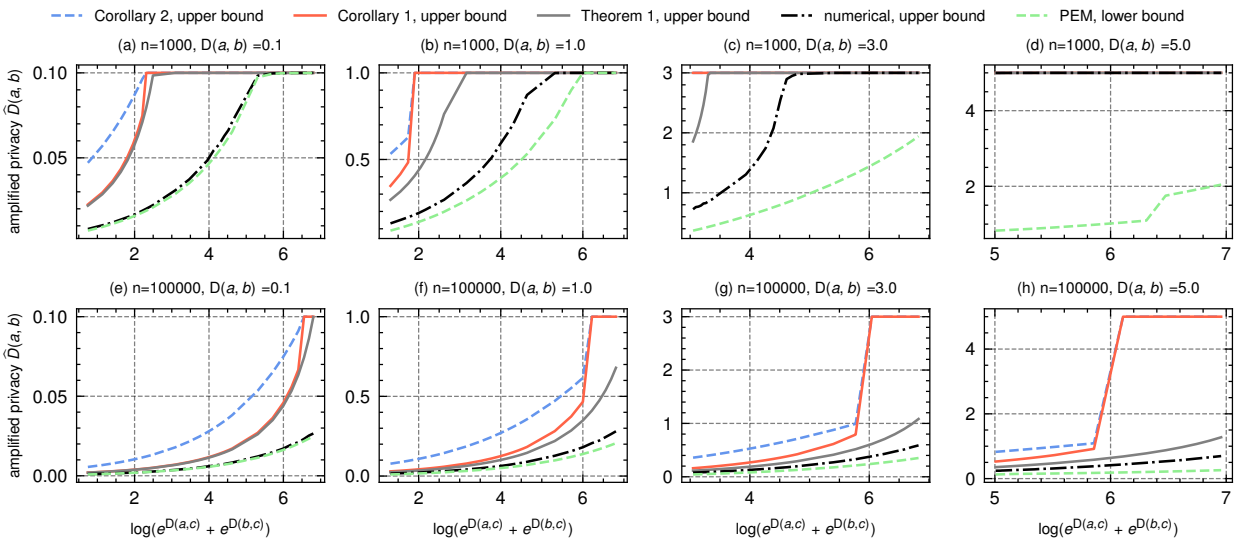

Fig. 13. Numerical comparison of upper bounds and lower bounds with $n = 10^3$ or 10^5, $\mathsf{D}(a,b) = 0.1, 1, 3,$ or 5, the $e^{\mathsf{D}(a,c)} + e^{\mathsf{D}(b,c)}$ varies from $e^{\mathsf{D}(a,b)} + 1$ to 1000, assumed $\mathsf{D}_{max} = \mathsf{D}(a,c) = \mathsf{D}(b,c) + \mathsf{D}(a,b)$ and $\delta = 0.01/n$.

prove it by contradiction and let (a', b') denote an edge satisfying $\mathsf{D}''(a',b') > \mathsf{D}^m(a',b')$. When $2\log(4/\delta) > \Omega_m$ holds, no amplification is applied. Since $\mathsf{D}^m(a,b) = \min(m, \mathsf{D}^\#(a,b))$, then $\mathsf{D}''(a',b') > \min(m, \mathsf{D}^\#(a',b'))$ causes a contradiction. When $2\log(4/\delta) \leq \Omega_m$ holds, note that $\mathsf{D}^m(a,b)$ always equal to

$$\min(m, \max(\mathsf{D}^\#(a,b), \texttt{Solve}[\mathsf{F}(v, \Omega_m) \leq \mathsf{D}^\#(a,b),\ v])$$

for all edges, thus $\min(\mathsf{D}^m(a,b), \mathsf{F}(\mathsf{D}^m(a,b), \Omega_m)) = \mathsf{D}^\#(a,b)$. Considering the dominated edge (a', b'), since $\mathsf{F}(v, \Omega_m)$ increases with v, we have $\mathsf{F}(\mathsf{D}''(a',b'), \Omega_m)) > \mathsf{F}(\mathsf{D}^m(a',b'), \Omega_m)$, therefore the privacy amplified distance $\min(\mathsf{D}''(a',b'), \mathsf{F}(\mathsf{D}''(a',b'), \Omega_m)) > \mathsf{D}^\#(a',b')$. Hence D'' is not a feasible solution (w.r.t. Corollary 1 and the global metric $\mathsf{D}^\#(a,b)$), and causes a contradiction.

Now assume there exists a feasible local semi-metric D'' that is not dominated by any element in $\{\mathsf{D}^m\}_{m\in[low,high]}$. We separately consider three cases about the diameter of the D'': (1) $\mathsf{D}''_{max} < low$, (2) $\mathsf{D}''_{max} > high$, (3) $low \leq \mathsf{D}''_{max} \leq high$. For the first case, since $D^{low}(a,b) \equiv low$ holds for all $a, b \in \mathcal{X}$, then $\mathsf{D}''(a,b) \leq \mathsf{D}''_{max} < low \leq D^{low}(a,b)$ violates the assumption about D''. For the second case, since $\mathsf{D}''_{max} > high$, then $\min(\mathsf{D}''_{max}, \mathsf{F}(\mathsf{D}''_{max}, \Omega_{\mathsf{D}''_{max}})) > \mathsf{D}^\#_{max}$ means the edges with a distance equal to D''_{max} violate Corollary 1, and D'' is not a feasible solution. For the third case, given the conclusion from the previous paragraph, we have the D'' must be dominated by $\mathsf{D}^{\mathsf{D}''_{max}}$ and also induces a contradiction.

Therefore, all feasible local semi-metrics regarding Corollary 1 are dominated by (some elements from) $\{\mathsf{D}^m\}_{m\in[low,high]}$. Further remove dominated elements inside the $\{\mathsf{D}^m\}_{m\in[low,high]}$ as in line 20 to get the final set $\mathbf{L}$, we have the conclusion.

Supplementary Material D
Proof of WitchHat's privacy guarantee

We first show that the probability density function for WitchHat is valid. Let R denote $\frac{mF_m}{2}$, for any input x, the integral of probability density multiplied by N_m is a constant value as follows:

$$\begin{aligned}&\int_{-B-R}^{x-R} e^{-m}\,\mathrm{d}z + \int_{x-R}^{x+R} e^{-2|z-x|/F_m}\,\mathrm{d}z + \int_{x+R}^{B+R} e^{-m}\,\mathrm{d}z\\ =&F_m(1-e^{-m}) + 2Be^{-m}.\end{aligned}$$

Combining the fact that the probability density is defined over the same output domain and is always non-negative for any input x, we conclude that the WitchHat defines a valid probability distribution.

Considering two inputs $a, b \in [-B, B]$ and any output $z \in \mathcal{D}_m$. Assuming $a > b$, by definition, the probability ratio is bounded by:

$$\begin{aligned}\frac{\mathbb{P}[z|x=a]}{\mathbb{P}[z|x=b]} &\leq \max\left\{1, \frac{\mathbb{P}[z=a|x=a]}{\mathbb{P}[z=a|x=b]}, \max_{z\geq a}\frac{\mathbb{P}[z|x=a]}{\mathbb{P}[z|x=b]}\right\}\\ &\leq e^{2\frac{a-b}{F_m}}.\end{aligned}$$

Similarly, we have $\frac{\mathbb{P}[z|x=a]}{\mathbb{P}[z|x=b]} \leq e^{2\frac{|a-b|}{F_m}}$ when $a < b$. Further since $\frac{\max_{x,z}\mathbb{P}[z|x]}{\min_{x',z'}\mathbb{P}[z'|x']} \leq e^m$ holds for all $x, x' \in [-B, B]$ and $z, z' \in \mathcal{D}_m$, we reach the conclusion.

Supplementary Material E
Proof of WitchHat's error bound

We start by showing the estimator $\hat{x}$ is unbiased, then analyze the variance of $\hat{x}$. According to the proportional probability distribution of z, the expectation $N_m \cdot \mathbb{E}[z]$ is:

$$\begin{aligned}&\int_{-B-R}^{x-R} ze^{-m}\,\mathrm{d}z + \int_{x-R}^{x+R} ze^{-2|z-x|/F_m}\,\mathrm{d}z + \int_{x+R}^{B+R} ze^{-m}\,\mathrm{d}z\\ =&0 + (N_m - 2e^{-m}(B+R))\cdot x\\ =&F_m(1-e^{-m}-me^{-m})x.\end{aligned}$$

Remove some minor negative terms, then the expectation $N_m \cdot \mathbb{E}[z^2]$ is bounded by:

$$\begin{aligned}
&\int_{-B-R}^{x-R} z^2 e^{-m}\,\mathrm{d}z + \int_{x-R}^{x+R} z^2 e^{-\frac{2|z-x|}{F_m}}\,\mathrm{d}z + \int_{x+R}^{B+R} z^2 e^{-m}\,\mathrm{d}z \\
=&\frac{2e^{-m}}{3}((B+R)^3 - R^3) - 2e^{-m}Rx^2 + \frac{F_m^3}{2} + F_m x^2 \\
&- e^{-m}F_m\frac{2x^2+2R^2+2F_mR+F_m^2}{2} \\
\leq&\frac{2B^3e^{-m}}{3} + \frac{F_m(4B^2me^{-m}+2Bm^2F_me^{-m})}{4} + \frac{F_m^3}{2} \\
&+ F_m(1-e^{-m}-me^{-m})x^2 \\
\leq&\frac{(4B^2+6mF_mB+3m^2F_m)e^{-m}B}{6} + \frac{F_m^3}{2} \\
&+ F_m(1-e^{-m}-me^{-m})x^2
\end{aligned}$$

Combining previous results together, we get the variance of the estimator as $\mathbb{E}[|\hat{x}-x|_2^2] = (\frac{N_m}{F_m(1-e^{-m}-me^{-m})})^2(\mathbb{E}[z^2] - (\mathbb{E}[z])^2) \leq \frac{N_m(e^{-m}B(8B^2+12mF_mB+6m^2F_m)+6F_m^3+3N_mx^2)}{12F_m^2(1-e^{-m}-me^{-m})^2})$.

Supplementary Material F
A local metric solver for Corollary 4.4

Follow similar procedures as Algorithm 1, we can find feasible local semi-metrics with respect to Corollary 2. We present an implementation in Algorithm 3, which uses a new amplification function according to Corollary 2 and requires a new condition for amplification: $n \geq 16\exp(m)\log(4/\delta)$. The feasibility of solutions outputted by Algorithm 3 is guaranteed in Theorem 6.

Algorithm 3: Find feasible local semi-metrics regarding Corollary 2

Input: Global privacy parameter $(\mathsf{D}^{\#}, \delta)$, number of users n.
Output: A list of feasible semi-metrics.

$\mathsf{G}(v,u) \equiv \log(1 + 8\frac{e^v-1}{e^v+1}\cdot(\sqrt{\frac{e^u\log(4/\delta)}{n-1}} + \frac{e^u}{n-1}))$
$\mathbf{L} \leftarrow \Phi$
$low \leftarrow \texttt{Solve}[\mathsf{G}(v,v) \leq \mathsf{D}^{\#}_{min},\ v]$
$low \leftarrow \max(\mathsf{D}^{\#}_{min}, \min(low, \texttt{Solve}[16\log(4/\delta)\exp(v) \leq n,\ v]))$
$high \leftarrow \texttt{Solve}[\mathsf{G}(v,v) \leq \mathsf{D}^{\#}_{max},\ v]$
$high \leftarrow \max(\mathsf{D}^{\#}_{max}, \min(high, \texttt{Solve}[16\log(4/\delta)\exp(v) \leq n,\ v]))$
for $m \in [low, high]$ **do**
  Initialize a new semi-metric D
  for $a, b \in \mathcal{X}$ **do**
    $currentdist \leftarrow \mathsf{D}^{\#}(a,b)$
    if $16\log(4/\delta)\exp(m) > n$ **then**
      $dist \leftarrow currentdist$
    else
      $dist \leftarrow \texttt{Solve}[\mathsf{G}(v,m) \leq currentdist,\ v]$
    $dist \leftarrow \min(\max(dist, currentdist), m)$
    $\mathsf{D}(a,b) \leftarrow dist$
  if D *is not dominated by any metrics in* $\mathbf{L}$ **then**
    $\mathbf{L} \leftarrow (\mathbf{L} \bigcup \{\mathsf{D}\})/\{\tilde{\mathsf{D}} \mid \tilde{\mathsf{D}} \in \mathbf{L}\ and\ \tilde{\mathsf{D}} \preceq \mathsf{D}\}$
return $\mathbf{L}$

Theorem 6: The solutions from Algorithm 3 are feasible semi-metrics regarding Corollary 2.

Proof: To prove the distance measure from Algorithm 3 is a feasible semi-metric obeying Corollary 2, it is suffice to show the amplified distance $\widehat{\mathsf{D}}_n(a,b)$ is dominated by $\mathsf{D}^{\#}(a,b)$ for all $a, b \in \mathcal{X}$. We first show D is a semi-metric: the self-identity property $\mathsf{D}(x,x) = 0$ is ensured by the line 12 or the line 14 with $currentdist = 0$; the positive property is ensured by the line 16 that $dist \geq currentdist > 0$; the symmetry property is ensured by the fact that lines 10-17 involve with exactly the same variables (e.g., m, δ) for all $currentdist$. We then show $\widehat{\mathsf{D}}_n(a,b) \leq \mathsf{D}^{\#}(a,b)$. When $16\log(4/\delta)\exp(m) > n$, we have $\mathsf{D}(a,b) = \mathsf{D}^{\#}(a,b)$ except for edges $\mathsf{D}(a,b) = m$, because $m \leq high$ then $16\log(4/\delta)\exp(high) > n$ and $high = \mathsf{D}^{\#}_{max}$ by the line 5, which implies $m \leq \mathsf{D}^{\#}_{max}$ and $\mathsf{D} \preceq \mathsf{D}^{\#}$. When $16\log(4/\delta)\exp(m) \leq n$, we have $\mathsf{D}(a,b) \equiv \min(m, \max(\mathsf{D}^{\#}(a,b), \mathsf{Solve}[\mathsf{G}(v,m)) \leq \mathsf{D}^{\#}(a,b), v]))$ by the lines 14 and 16, therefore $\min(\mathsf{D}(a,b), \mathsf{G}(\mathsf{D}(a,b), m)) = \mathsf{D}^{\#}(a,b)$ and $\widehat{\mathsf{D}}_n \preceq \mathsf{D}^{\#}$ hold. ■