

Tailoring Differentially Private Bayesian Inference to Distance Between Distributions

Jiawen Liu^{*}, Mark Bun^{**}, Gian Pietro Farina^{*}, and Marco Gaboardi^{*}

^{*}Department of Computer Science and Engineering, University at Buffalo, SUNY. {jliu223,gianpiet,gaboardi}@buffalo.edu

^{**}Department of Computer Science, Princeton University.
mbun@cs.princeton.edu

1 Preliminaries

Bayesian Inference.

Given a prior belief $\Pr(\theta)$ on some parameter θ , and an observation \mathbf{x} , the posterior distribution on θ given \mathbf{x} is computed as:

$$\Pr(\theta|\mathbf{x}) = \frac{\Pr(\mathbf{x}|\theta) \cdot \Pr(\theta)}{\Pr(\mathbf{x})}$$

where the expression $\Pr(\mathbf{x}|\theta)$ denotes the *likelihood* of observing \mathbf{x} under a value of θ . Since we consider \mathbf{x} to be fixed, the likelihood is a function of θ . For the same reason $\Pr(\mathbf{x})$ is a constant independent of θ . Usually in statistics the prior distribution $\Pr(\theta)$ is chosen so that it represents the initial belief on θ , that is, when no data has been observed. In practice though, prior distributions and likelihood functions are usually chosen so that the posterior belongs to the same *family* of distributions. In this case we say that the prior is conjugate to the likelihood function. Use of a conjugate prior simplifies calculations and allows for inference to be performed in a recursive fashion over the data.

Beta-binomial System.

In this work we will consider a specific instance of Bayesian inference and one of its generalizations. specifically, a Beta-binomial mode. We will consider the situation the underlying data is binomial distribution ($\sim \text{binomial}(\theta)$), where θ represents the parameter –informally called *bias*– of a Bernoulli distributed random variable. The prior distribution over $\theta \in [0, 1]$ is going to be a beta distribution, $\text{beta}(\alpha, \beta)$, with parameters $\alpha, \beta \in \mathbb{R}^+$, and with p.d.f:

$$\Pr(\theta) \equiv \frac{\theta^\alpha (1 - \theta)^\beta}{B(\alpha, \beta)}$$

where $B(\cdot, \cdot)$ is the beta function. The data \mathbf{x} will be a sequence of $n \in \mathbb{N}$ binary values, that is $\mathbf{x} = (x_1, \dots, x_n)$, $x_i \in \{0, 1\}$, and the likelihood function is:

$$\Pr(\mathbf{x}|\theta) \equiv \theta^{\Delta\alpha} (1 - \theta)^{n - \Delta\alpha}$$

where $\Delta\alpha = \sum_{i=1}^n x_i$. From this it can easily be derived that the posterior distribution is:

$$\Pr(\theta|\mathbf{x}) = \text{beta}(\alpha + \Delta\alpha, \beta + n - \Delta\alpha)$$

Dirichlet-multinomial Systems.

The beta-binomial model can be immediately generalized to Dirichlet-multinomial, with underlying data multinomially distributed. The *bias* is represented by parameter $\boldsymbol{\theta}$, the vector of parameters of a categorically distributed random variable. The prior distribution over $\boldsymbol{\theta} \in [0, 1]^k$ is given by a Dirichlet distribution, $\text{DL}(\boldsymbol{\alpha})$, for $k \in \mathbb{N}$, and $\boldsymbol{\alpha} \in (\mathbb{R}^+)^k$, with p.d.f:

$$\Pr(\boldsymbol{\theta}) \equiv \frac{1}{B(\boldsymbol{\alpha})} \cdot \prod_{i=1}^k \theta_i^{\alpha_i - 1}$$

where $B(\cdot)$ is the generalized beta function. The data \mathbf{x} will be a sequence of $n \in \mathbb{N}$ values coming from a universe \mathcal{X} , such that $|\mathcal{X}| = k$. The likelihood function will be:

$$\Pr(\mathbf{x}|\boldsymbol{\theta}) \equiv \prod_{a_i \in \mathcal{X}} \theta_i^{\Delta\alpha_i},$$

with $\Delta\alpha_i = \sum_{j=1}^n [x_j = a_i]$, where $[\cdot]$ represents Iverson bracket notation. Denoting by $\Delta\boldsymbol{\alpha}$ the vector $(\Delta\alpha_1, \dots, \Delta\alpha_k)$ the posterior distribution over $\boldsymbol{\theta}$ turns out to be

$$\Pr(\boldsymbol{\theta}|\mathbf{x}) = \text{DL}(\boldsymbol{\alpha} + \Delta\boldsymbol{\alpha}).$$

where $+$ denotes the componentwise sum of vectors of reals.

Differential Privacy.

Definition 1. ϵ -differential privacy.

A randomized mechanism $\mathcal{M} : \mathcal{X} \rightarrow \mathcal{Y}$ is differential privacy, iff for any adjacent¹ input $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, a metric H over \mathcal{Y} and a $B \subseteq H(\mathcal{Y})$, \mathcal{M} satisfies:

$$\mathbb{P}[H(\mathcal{M}(\mathbf{x})) \in B] = e^\epsilon \mathbb{P}[H(\mathcal{M}(\mathbf{x}')) \in B].$$

Definition 2. (ϵ, δ) -differential privacy.

A randomized mechanism $\mathcal{M} : \mathcal{X} \rightarrow \mathcal{Y}$ is differential privacy, iff for any $\mathbf{adj}(\mathbf{x}, \mathbf{x}') \in \mathcal{X}$, a metric H over \mathcal{Y} and a $B \subseteq H(\mathcal{Y})$, \mathcal{M} satisfies:

$$\mathbb{P}[H(\mathcal{M}(\mathbf{x})) \in B] = e^\epsilon \mathbb{P}[H(\mathcal{M}(\mathbf{x}')) \in B] + \delta,$$

¹Given \mathbf{x}, \mathbf{x}' we say that \mathbf{x} and \mathbf{x}' are adjacent and we write, $\mathbf{adj}(\mathbf{x}, \mathbf{x}')$, iff $\sum_i^n [x_i = x'_i] \leq 1$.

where $\mathbf{x} = (x_i)_{i=1}^n$ and $\mathbf{x}' = (x'_i)_{i=1}^n$ is adjacent if there is only one j that $x_j \neq x'_j$ and $x_i = x'_i$ for $i = 1, 2, \dots, n; i \neq j$.

2 Technical Problem Statement and Motivations

We are interested in designing a mechanism for privately releasing the full posterior distributions derived in section 1, as opposed to just sampling from them. It's worth noticing that the posterior distributions are fully characterized by their parameters, and the family (beta, Dirichlet) they belong to. Hence, in case of the Beta-Binomial model we are interested in releasing a private version of the pair of parameters $(\alpha', \beta') = (\alpha + \Delta\alpha, \beta + n - \Delta\alpha)$, and in the case of the Dirichlet-multinomial model we are interested in a private version of $\boldsymbol{\alpha}' = (\boldsymbol{\alpha} + \Delta\boldsymbol{\alpha})$. [5] and [4] have already attacked this problem by adding independent Laplacian noise to the parameters of the posteriors. That is, in the case of the Beta-Binomial system, the value released would be: $(\tilde{\alpha}, \tilde{\beta}) = (\alpha + \Delta\alpha, \beta + n - \Delta\alpha)$ where $\Delta\alpha \sim \mathcal{L}(\Delta\alpha, \frac{2}{\epsilon})$, and where $\mathcal{L}(\mu, \nu)$ denotes a Laplace random variable with mean μ and scale ν . This mechanism is ϵ -differentially private, and the noise is calibrated w.r.t. to a sensitivity of 2 which is derived by using ℓ_1 norm over the pair of parameters. Indeed, considering two adjacent data observations \mathbf{x}, \mathbf{x}' , that, from a unique prior, give rise to two posterior distributions, characterized by the pairs (α', β') and (α'', β'') then $|\alpha' - \alpha''| + |\beta' - \beta''| \leq 2$. This argument extends similarly to the Dirichlet-Multinomial system. Details are introduced in Sec. ??.

However, in previous works, the accuracy of the posterior was measured again with respect to ℓ_1 norm. That is, an upper bound was given on

$$\Pr[|\alpha - \tilde{\alpha}| + |\beta - \tilde{\beta}| \geq \gamma]$$

where $(\alpha, \beta), (\tilde{\alpha}, \tilde{\beta})$ are as defined above. This accuracy metric is meaningless when the results released are distributions rather than numerical values. In contrast, distribution metrics such as f -divergence, Hellinger distance, etc. come into mind overtly when we are measuring distance between distributions. This gives us motivation on using a different norm (a distribution metric) to compute the sensitivity and provide guarantees on the accuracy.

Specifically, we will use the Hellinger distance $\mathcal{H}(\cdot, \cdot)$: Given two beta distributions $\beta_1 = \text{beta}(\alpha_1, \beta_1)$, and $\beta_2 = \text{beta}(\alpha_2, \beta_2)$ the following equality holds

$$\mathcal{H}(\beta_1, \beta_2) = \sqrt{1 - \frac{B(\frac{\alpha_1 + \alpha_2}{2}, \frac{\beta_1 + \beta_2}{2})}{\sqrt{B(\alpha_1, \beta_1)B(\alpha_2, \beta_2)}}}$$

Our choice to use Hellinger distance is motivated by three facts:

- It simplifies calculations in the case of the probabilistic models considered here.
- It also automatically yields bounds on the total variation distance, which represents also the maximum advantage an unbounded adversary can have in distinguishing two distributions.

- The accuracy can be improved by using a smooth bound on Hellinger distance’s local sensitivity. As shown in Fig. 1, taking advantage of the gap between the global sensitivity and local sensitivity, we can improve the accuracy by applying a smooth upper bound on local sensitivity instead of using global sensitivity.

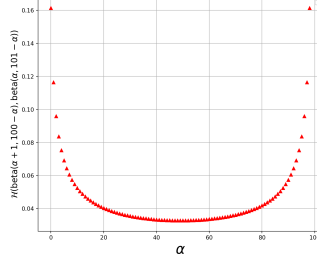


Figure 1: Sensitivity of \mathcal{H}

3 Mechanism Proposition

Given a prior distribution $\beta_{\text{prior}} = \text{beta}(\alpha, \beta)$ and a sequence of n observations $\mathbf{x} \in \{0, 1\}^n$, we define the following set:

$$\mathcal{R}_{\text{post}} \equiv \{\text{beta}(\alpha', \beta') \mid \alpha' = \alpha + \Delta\alpha, \beta' = \beta + n - \Delta\alpha\},$$

where $\Delta\alpha$ is as defined in Section 1. Notice that $\mathcal{R}_{\text{post}}$ has $n + 1$ elements, and the Bayesian Inference process will produce an element from $\mathcal{R}_{\text{post}}$ that we denote by $\text{BI}(\mathbf{x})$ – we don’t explicitly parametrize the result by the prior, which from now on we consider fixed and we denote it by β_{prior} .

3.1 $\mathcal{M}_{\mathcal{H}}$: Smoothed Hellinger Distance Based Exponential Mechanism

Definition 3. The mechanism $\mathcal{M}_{\mathcal{H}}(\mathbf{x})$ outputs a candidate $r \in \mathcal{R}_{\text{post}}$ with probability

$$\Pr_{z \sim \mathcal{M}_{\mathcal{H}}} [z = r] = \frac{\exp\left(\frac{-\epsilon \cdot \mathcal{H}(\text{BI}(\mathbf{x}), r)}{2 \cdot S(\mathbf{x})}\right)}{\sum_{r' \in \mathcal{R}_{\text{post}}} \exp\left(\frac{-\epsilon \cdot \mathcal{H}(\text{BI}(\mathbf{x}), r')}{2 \cdot S(\mathbf{x})}\right)}.$$

where $S_{\beta}(\mathbf{x})$ is the smooth sensitivity of $\mathcal{H}(\text{BI}(\mathbf{x}), -)$, calculated by:

$$S(\mathbf{x}) = \max_{\mathbf{x}' \in \{0, 1\}^n} \left\{ LS(\mathbf{x}') \cdot e^{-\gamma \cdot d(\mathbf{x}, \mathbf{x}')}\right\}, \quad (1)$$

where d is the Hamming distance between two datasets, and $\beta = \beta(\epsilon, \delta)$ is a function of ϵ and δ .

This mechanism is based on the basic exponential mechanism [2], with $\mathcal{R}_{\text{post}}$ as the range and $\mathcal{H}(\cdot, \cdot)$ as the scoring function. The difference is that in this mechanism we don't calibrate the noise w.r.t. to the global sensitivity of the scoring function but w.r.t. to the smooth sensitivity $S(\mathbf{x})$ – defined by [3]– of $\mathcal{H}(\text{Bl}(\mathbf{x}), \cdot)$.

$\gamma = \gamma(\epsilon, \delta)$ is a function of ϵ and δ to be determined later, and where $LS(\mathbf{x}')$ denotes the local sensitivity at $\text{Bl}(\mathbf{x}')$, or equivalently at \mathbf{x}' , of the scoring function used in our mechanism.

This mechanism also extends to the Dirichlet-multinomial system $\text{DL}(\boldsymbol{\alpha})$ by rewriting the Hellinger distance as:

$$\mathcal{H}(\text{DL}(\boldsymbol{\alpha}_1), \text{DL}(\boldsymbol{\alpha}_2)) = \sqrt{1 - \frac{B(\frac{\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2}{2})}{\sqrt{B(\boldsymbol{\alpha}_1)B(\boldsymbol{\alpha}_2)}}},$$

and by replacing the $\mathcal{R}_{\text{post}}$ with set of posterior Dirichlet distributions candidates. Also, the smooth sensitivity $S(\mathbf{x})$ in (1) will be computed by letting \mathbf{x}' range over all the elements in \mathcal{X}^n adjacent to \mathbf{x} . Notice that $\mathcal{R}_{\text{post}}$ has $\binom{n+1}{m-1}$ elements in this case. We will denote by $\mathcal{M}_{\mathcal{H}}^D$ the mechanism for the Dirichlet-multinomial system.

By setting the γ as $\ln(1 - \frac{\epsilon}{2 \ln(\frac{\epsilon}{\delta'})})$, $\mathcal{M}_{\mathcal{H}}$ is (ϵ, δ) –differentially private, where $\delta = \frac{\epsilon \delta'}{2}$. (or $\gamma = \ln(1 - \frac{\epsilon}{2 \ln(\frac{\epsilon}{2|\mathcal{R}_{\text{post}}|})})$ when generalized to Dirichlet-multinomial System)

4 Privacy Analysis

4.1 Privacy Analysis for Baseline Mechanisms

In baseline mechanisms, *exponential mechanism*, *Laplace mechanism*, *improved Laplace mechanism* are ϵ –differential privacy provided by [1]. The *exponential mechanism with local sensitivity* is non-differential privacy, also from [1].

4.2 Privacy Analysis for $\mathcal{M}_{\mathcal{H}}$

The differential privacy property of $\mathcal{M}_{\mathcal{H}}$ is proved based on following two lemmas: Lemma 4.1 and Lemma 4.6.

Sliding Property of $\mathcal{M}_{\mathcal{H}}$

Lemma 4.1. *Given $\mathcal{M}_{\mathcal{H}}(\mathbf{x})$ calibrated on the smooth sensitivity. Let $\lambda = f(\epsilon, \delta')$, $\epsilon \geq 0$ and $|\delta'| < 1$. Then, the following sliding property holds:*

$$\Pr_{r \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})} [u(r, \mathbf{x}) = \hat{s}] \leq e^{\frac{\epsilon}{2}} \Pr_{r \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})} [u(r, \mathbf{x}) = (\delta' + \hat{s})] + \frac{\delta'}{2},$$

Proof. In what follows, we will use a correspondence between the probability $\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r]$ of every $r \in \mathcal{R}_{\text{post}}$ and the probability $\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[\mathcal{H}(\text{BI}(\mathbf{x}), z) = \mathcal{H}(\text{BI}(\mathbf{x}), r)]$ for the utility score of $r \in \mathcal{R}_{\text{post}}$. In Beta-binomial system, because of the property of beta function, the order of parameters in Beta distribution doesn't matter when computing the Hellinegr distance.

Theorem 4.2. *for 3 different beta distributions $\text{beta}(\alpha_0, \beta_0)$, $\text{beta}(\alpha_1, \beta_1)$ and $\text{beta}(\alpha_2, \beta_2)$, $\mathcal{H}(\text{beta}(\alpha_0, \beta_0), \text{beta}(\alpha_1, \beta_1)) = \mathcal{H}(\text{beta}(\alpha_0, \beta_0), \text{beta}(\alpha_2, \beta_2))$ iff:*

$$\alpha_2 = \beta_1, \alpha_1 = \beta_2, \alpha_0 + \alpha_1 = \beta_0 + \beta_2, \alpha_0 + \alpha_2 = \beta_0 + \beta_1.$$

Theorem 4.3. *In Dirichlet-multinomial system, for 3 different Dirichlet distributions $\text{DL}(\alpha)$, $\text{DL}(\beta_1)$ and $\text{DL}(\beta_2)$, $\mathcal{H}(\text{DL}(\alpha), \text{DL}(\beta_1)) = \mathcal{H}(\text{DL}(\alpha), \text{DL}(\beta_2))$ iff:*

$$\text{set}(\beta_1) = \text{set}(\beta_2) \text{ and } \text{set}(\alpha + \beta_1) = \text{set}(\alpha + \beta_2),$$

where $\text{set}(\cdot)$ means we only consider the value of each dimension in the vector without the order.

Theorem 4.4. *There are at most two candidates in $\mathcal{R}_{\text{post}}$ can have the same score, i.e., for every $r \in \mathcal{R}_{\text{post}}$, $\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[\mathcal{H}(\text{BI}(\mathbf{x}), z) = \mathcal{H}(\text{BI}(\mathbf{x}), r)]$ either*

$$= 2 \left(\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] \right)$$

or

$$= \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r].$$

Proof. of Theorem 4.4.

Let $|z|$ be the number of candidates $z \in \mathcal{R}$ that satisfy $\mathcal{H}(\text{BI}(\mathbf{x}), z) = \mathcal{H}(\text{BI}(\mathbf{x}), r)$, i.e. $u(z, \mathbf{x}) = u(r, \mathbf{x})$, then $\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(z, \mathbf{x}) = u(r, \mathbf{x})] = |z| \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r]$.

From Theorem 4.2, It can be infer that $|z| = 2$ or $|z| = 1$, i.e., $\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(z, \mathbf{x}) = u(r, \mathbf{x})] = 2 \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r]$ or $\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r]$. \square

Theorem 4.5. *In Dirichlet-multinomial system with k -dimensional Dirichlet distribution, for every $r \in \mathcal{R}_{\text{post}}$:*

$$\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[\mathcal{H}(\text{BI}(\mathbf{x}), z) = \mathcal{H}(\text{BI}(\mathbf{x}), r)] = |z| \left(\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] \right),$$

and $|z| = \frac{k!k!}{|a_1|!|a_2|!\dots|a_{k_a}|!|b_1|!|b_2|!\dots|b_{k_b}|!}$, where $\sum |a_i| = \sum |b_i| = k$.

Proof. of Theorem 4.5.

Let $\text{BI}(\mathbf{x}) = \text{DL}(\alpha)$, and $\text{DL}(\beta)$ be a candidate from $\mathcal{R}_{\text{post}}$, a_i is the distinct value in α and b_i the same in β , $|a_i|$ denotes the number of values in α having the value a_i and $|b_i|$ the same in β . Then, there are $\frac{k!k!}{|a_1|!|a_2|!\dots|a_{k_a}|!|b_1|!|b_2|!\dots|b_{k_b}|!}$ possible permutation when the condition in Theorem 4.3 is satisfied. \square

In Beta-binomial system, let $R_1, R_2 \subset \mathcal{R}_{\text{post}}$ be a partition of $\mathcal{R}_{\text{post}}$, where every $z \in R_1$ has a distinct score, i.e., $|z| = 1$ and $z \in R_2$ has another $z' \in R_2$ with the same score, i.e., $|z| = 2$. The proofs are given by parts:

- for $r \in R_1$:

We denote the normalizer of the probability mass in $\mathcal{M}_{\mathcal{H}}(\mathbf{x})$: $\sum_{r' \in \mathcal{R}} \exp(\frac{\epsilon u(r', \mathbf{x})}{2S(\mathbf{x})})$ as $NL(\mathbf{x})$:

$$\begin{aligned} LHS &= \Pr_{r \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})} [u(r, \mathbf{x}) = \hat{s}] = \frac{\exp(\frac{\epsilon \hat{s}}{2S(\mathbf{x})})}{NL(\mathbf{x})} \\ &= \frac{\exp(\frac{\epsilon(\hat{s} + \delta' - \delta')}{2S(\mathbf{x})})}{NL(\mathbf{x})} \\ &= \frac{\exp(\frac{\epsilon(\hat{s} + \delta')}{2S(\mathbf{x})} + \frac{-\epsilon \delta'}{2S(\mathbf{x})})}{NL(\mathbf{x})} \\ &= \frac{\exp(\frac{\epsilon(\hat{s} + \delta')}{2S(\mathbf{x})})}{NL(\mathbf{x})} \cdot e^{\frac{-\epsilon \delta'}{2S(\mathbf{x})}}. \end{aligned}$$

By bounding the $\delta' \geq -S(\mathbf{x})$, we can get:

$$\begin{aligned} \frac{\exp(\frac{\epsilon(\hat{s} + \delta')}{2S(\mathbf{x})})}{NL(\mathbf{x})} \cdot e^{\frac{-\epsilon \delta'}{2S(\mathbf{x})}} &\leq \frac{\exp(\frac{\epsilon(\hat{s} + \delta')}{2S(\mathbf{x})})}{NL(\mathbf{x})} \cdot e^{\frac{\epsilon}{2}} \\ &= e^{\frac{\epsilon}{2}} \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})} [u(z, \mathbf{x}) = (\delta' + \hat{s})] \leq RHS \end{aligned}$$

- for $r \in R_2$: Proof actually is exactly the same as above by eliminating the parameter $|z|$ in both sides.

The same in Dirichlet-multinomial distribution, partitioning the $\mathcal{R}_{\text{post}}$ by the possible value of $|z|$, the proof can be derived in the same way as above by part.

□

Dilation Property of $\mathcal{M}_{\mathcal{H}}$

Lemma 4.6. *for any exponential mechanism $\mathcal{M}_{\mathcal{H}}(\mathbf{x})$, $\lambda < |\beta|$, ϵ , $|\delta'| < 1$ and $\beta \leq \ln(1 - \frac{\epsilon}{2 \ln(\frac{\epsilon}{2|\mathcal{R}_{\text{post}}|})})$, the dilation property holds:*

$$\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})} [u(z, \mathbf{x}) = c] \leq e^{\frac{\epsilon}{2}} \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})} [u(z, \mathbf{x}) = e^{\lambda} c] + \frac{\delta'}{2},$$

where the sensitivity in mechanism is still smooth sensitivity as above.

More specifically, in Beta-binomial system it is enough to take $\beta \leq \ln(1 - \frac{\epsilon}{2 \ln(\frac{\epsilon}{2(n+1)})})$.

Proof. Based on Theorem 4.2 and 4.4, we use the same partition of $\mathcal{R}_{\text{post}}$, R_1, R_2 . It is enough to proof under just one partition. Without loss of generalization, we take R_1 .

The sensitivity is always greater than 0, and our utility function $-\mathcal{H}(\text{BI}(\mathbf{x}), z)$ is smaller than zero, i.e., $u(z, \mathbf{x}) \leq 0$, we need to consider two cases where $\lambda < 0$, and $\lambda > 0$:

We set the $h(c) = \Pr[u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) = c] = \frac{\exp(\frac{\epsilon c}{2S(\mathbf{x})})}{NL(\mathbf{x})}$.

We first consider $\lambda < 0$. In this case, $1 < e^\lambda$, so the ratio $\frac{h(c)}{h(e^\lambda c)} = \frac{\exp(\frac{\epsilon c}{2S(\mathbf{x})})}{\exp(\frac{\epsilon(c \cdot e^\lambda)}{2S(\mathbf{x})})}$ is at most $\frac{\epsilon}{2}$.

Next, we proof the dilation property for $\lambda > 0$, The ratio of $\frac{h(c)}{h(e^\lambda c)}$ is $\exp(\frac{\epsilon}{2} \cdot \frac{u(\mathcal{M}_{\mathcal{H}}(\mathbf{x}))(1-e^\lambda)}{S(\mathbf{x})})$. Consider the event $G = \{\mathcal{M}_{\mathcal{H}}(\mathbf{x}) : u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) \leq \frac{S(\mathbf{x})}{(1-e^\lambda)}\}$. Under this event, the log-ratio above is at most $\frac{\epsilon}{2}$. The probability of G under density $h(c)$ is $1 - \frac{\delta'}{2}$. Thus, the probability of a given event z is at most $\Pr[c \cap G] + \Pr[\bar{G}] \leq e^{\frac{\epsilon}{2}} \Pr[e^\lambda c \cap G] + \frac{\delta'}{2} \leq e^{\frac{\epsilon}{2}} \Pr[e^\lambda c] + \frac{\delta'}{2}$.

Detail proof: To show:

$$\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(z, \mathbf{x}) = c] \leq e^{\frac{\epsilon}{2}} \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(z, x) = e^\lambda c] + \frac{\delta'}{2},$$

for $z \in R_1$.

Let $\Pr[u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) = c] = \frac{\exp(\frac{\epsilon c}{2S(\mathbf{x})})}{NL(\mathbf{x})}$ and $\Pr[u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) = e^\lambda c] = \frac{\exp(\frac{\epsilon e^\lambda c}{2S(\mathbf{x})})}{NL(\mathbf{x})}$ by definition.

After simplification, we need to show: $u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) \leq \frac{S(\mathbf{x})}{(1-e^\lambda)}$. Because the sensitivity is always greater than 0, and our utility function $-\mathcal{H}(\text{BI}(\mathbf{x}), z)$ is smaller than zero, i.e., $u(z, x) \leq 0$, we need to consider two cases where $\lambda < 0$, and $\lambda > 0$:

- $\lambda < 0$

The left hand side will always be smaller than 0 and the right hand side greater than 0. This will always holds, i.e.

$$u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) \leq \frac{S(\mathbf{x})}{(1-e^\lambda)}$$

is always true when $\lambda < 0$

- $\lambda > 0$

Because $\hat{s} = u(r)$ where $r \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})$, we can substitute \hat{s} with $u(\mathcal{M}_{\mathcal{H}}(\mathbf{x}))$. Then, what we need to proof under the case $\lambda > 0$ is:

$$u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) \leq \frac{S(\mathbf{x})}{(1-e^\lambda)}. \quad (2)$$

Based on the accuracy property of exponential mechanism:

$$Pr[u(\mathcal{M}_E(x, u, \mathcal{R}_{\text{post}})) \leq c] \leq \frac{|\mathcal{R}| \exp(\frac{\epsilon c}{2GS})}{|\mathcal{R}_{OPT}| \exp(\frac{\epsilon OPT_{u(\mathbf{x})}}{2GS})},$$

we derived the accuracy bound for $\mathcal{M}_{\mathcal{H}}$:

$$Pr[u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) \leq c] \leq |\mathcal{R}_{\text{post}}| \exp(\frac{\epsilon c}{2S(\mathbf{x})}).$$

In Beta-binomial system, $|\mathcal{R}_{\text{post}}| = n + 1$, apply this bound to eq. 2:

$$\begin{aligned} Pr[u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) \leq \frac{S(\mathbf{x})}{(1-e^\lambda)}] &= (n+1) \exp(\frac{\epsilon S(\mathbf{x})}{(1-e^\lambda)} / 2S(\mathbf{x})) \\ &= (n+1) \exp(\frac{\epsilon}{2(1-e^\lambda)}). \end{aligned}$$

When we set $\lambda \leq \ln(1 - \frac{\epsilon}{2 \ln(\frac{\delta'}{2(n+1)})})$, it is easily to derive that $Pr[u(\mathcal{M}_{\mathcal{H}}(\mathbf{x})) \leq \frac{S(\mathbf{x})}{(1-e^\lambda)}] \leq \frac{\delta'}{2}$.

In Dirichlet-multinomial system, λ is set as $\leq \ln(1 - \frac{\epsilon}{2 \ln(\frac{\delta'}{2|\mathcal{R}_{\text{post}}|})})$ since $|\mathcal{R}_{\text{post}}| \neq n + 1$ any more.

□

(ϵ, δ) -Differential Privacy of $\mathcal{M}_{\mathcal{H}}$

Lemma 4.7. $\mathcal{M}_{\mathcal{H}}$ is (ϵ, δ) -differential privacy.

Proof. For all $\mathbf{adj}(\mathbf{x}, \mathbf{x}') \in \mathcal{X}$ and any beta distribution r , we need to show that:

$$Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] \leq e^\epsilon Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x}')} [z = r] + \delta.$$

Let partition $R_1, R_2 \subset \mathcal{R}_{\text{post}}$ the same as above. Then we prove by part:

- for $r \in R_1$:

Given that $Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] = Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(\mathbf{x}, z) = u(\mathbf{x}, r)]$, let $U = u(\mathbf{x}, r)$, $U_1 = u(\mathbf{x}', z) - u(\mathbf{x}, z)$, $U_2 = U + U_1$ and $U_3 = U_2 \cdot \frac{S(\mathbf{x})}{S(\mathbf{x}')} \cdot \ln(\frac{NL(\mathbf{x})}{NL(\mathbf{x}')})$. Then,

$$\begin{aligned} Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] &= Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(x, z) = U] \\ &\leq e^{\epsilon/2} \cdot Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(x, z) = U_2] \\ &\leq e^\epsilon \cdot Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(x, z) = U_3] + e^{\epsilon/2} \cdot \frac{\delta'}{2} \\ &= e^\epsilon \cdot Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x}')} [u(\mathbf{x}', z) = U] + \frac{e^{\epsilon/2} \delta'}{2} \\ &= e^\epsilon \cdot Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] + \delta, \end{aligned}$$

where $\delta = \frac{e^{\epsilon/2}\delta'}{2}$. The first inequality holds by the sliding property, since the $U_1 \geq -S(\mathbf{x})$. The second inequality holds by the dilation property, since $\frac{S(\mathbf{x})}{S(\mathbf{x}')} \cdot \ln(\frac{NL(\mathbf{x})}{NL(\mathbf{x}')}) \leq 1 - \frac{\epsilon}{2 \ln(\frac{\delta'}{2(n+1)})}$.

- for $r \in R_2$:

Given that $2 \left(\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] \right) = \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(\mathbf{x}, z) = u(\mathbf{x}, r)]$, let $U = u(\mathbf{x}, r)$, $U_1 = u(y, z) - u(\mathbf{x}, z)$, $U_2 = U + U_1$ and $U_3 = U_2 \cdot \frac{S(\mathbf{x})}{S(\mathbf{x}')} \cdot \ln(\frac{NL(\mathbf{x})}{NL(\mathbf{x}')})$. Then,

$$\begin{aligned} \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] &= \frac{1}{2} \left(\Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(x, z) = U] \right) \\ &\leq \frac{1}{2} \left(e^{\epsilon/2} \cdot \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(x, z) = U_2] \right) \\ &\leq \frac{1}{2} \left(e^{\epsilon} \cdot \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[u(x, z) = U_3] + e^{\epsilon/2} \cdot \frac{\delta'}{2} \right) \\ &= \frac{1}{2} \left(e^{\epsilon} \cdot \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x}')} [u(y, z) = U] + \frac{e^{\epsilon/2}\delta'}{2} \right) \\ &= e^{\epsilon} \cdot \Pr_{z \sim \mathcal{M}_{\mathcal{H}}(\mathbf{x})}[z = r] + \delta, \end{aligned}$$

where the $\delta = \frac{e^{\epsilon/2}\delta'}{4}$.

The first inequality holds by the sliding property, since the $U_1 \geq -S(\mathbf{x})$.

The second inequality holds by the dilation property, since $\frac{S(\mathbf{x})}{S(\mathbf{x}')} \cdot \ln(\frac{NL(\mathbf{x})}{NL(\mathbf{x}')}) \leq 1 - \frac{\epsilon}{2 \ln(\frac{\delta'}{2(n+1)})}$.

Based on two cases above, $\mathcal{M}_{\mathcal{H}}$ is (ϵ, δ) -differential privacy, where δ takes the maximum value from the two cases, i.e., $\frac{e^{\frac{\epsilon}{2}}\delta'}{2}$.

Proof in Dirichlet-multinomial system can be derived in the same way by proving in parts and taking the maximum δ value of each part. \square

References

- [1] Cynthia Dwork, Aaron Roth, et al. *The algorithmic foundations of differential privacy*. Now Publishers, Inc., 2014.
- [2] Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *FOCS, 2007*.
- [3] Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Smooth sensitivity and sampling in private data analysis. In *STOC, 2007*, pages 75–84.
- [4] Yonghui Xiao and Li Xiong. Bayesian inference under differential privacy.

- [5] Zuhe Zhang, Benjamin IP Rubinstein, Christos Dimitrakakis, et al. On the differential privacy of bayesian inference. In *AAAI, 2016*, pages 2365–2371.