Journal of Privacy and Confidentiality
Vol. 10 (1) 2020
TPDP 2018
Submitted Feb 2019
Published Jan 2020

THE BOUNDED LAPLACE MECHANISM IN DIFFERENTIAL PRIVACY

NAOISE HOLOHAN, SPIROS ANTONATOS, STEFANO BRAGHIN, AND PÓL MAC AONGHUSA

IBM Research, Dublin, Ireland
e-mail address: naoise@ibm.com

IBM Research, Dublin, Ireland
e-mail address: santonat@ie.ibm.com

IBM Research, Dublin, Ireland
e-mail address: stefanob@ie.ibm.com

IBM Research, Dublin, Ireland
e-mail address: aonghusa@ie.ibm.com

Abstract. The Laplace mechanism is the workhorse of differential privacy, applied to many instances where numerical data is processed. However, the Laplace mechanism can return semantically impossible values, such as negative counts, due to its infinite support. There are two popular solutions to this: (i) bounding/capping the output values and (ii) bounding the mechanism support. In this paper, we show that bounding the mechanism support, while using the parameters of the standard Laplace mechanism, does not typically preserve differential privacy. We also present a robust method to compute the optimal mechanism parameters to achieve differential privacy in such a setting.

1. Introduction

Data privacy is an important factor that data owners must take into consideration when collecting, storing, sharing and publishing user data. This extends to publishing statistics on user data. In recent years, differential privacy has emerged as a popular privacy framework, thanks to its robust mathematical privacy guarantees.

The Laplace mechanism is the workhorse of differential privacy, frequently utilised in applications on numerical data. Its strength lies in its mathematical and computational simplicity, in contrast to other mechanisms such as the exponential mechanism. In spite of its popularity however, the Laplace mechanism lacks *consistency* in its outputs. Consider, for example, adding noise from the Laplace mechanism to a count query; negative results hold no meaning, yet are a valid output of the mechanism, occurring especially frequently for low-numbered counts.

Key words and phrases: Differential privacy, approximate differential privacy, Laplace mechanism, consistency, bounds, bounded mechanism, truncation, truncated mechanism, resampling, rejection sampling.

JPrivacy Confidentiality www.journalprivacyconfidentiality.org
DOI:10.29012/jpc.175

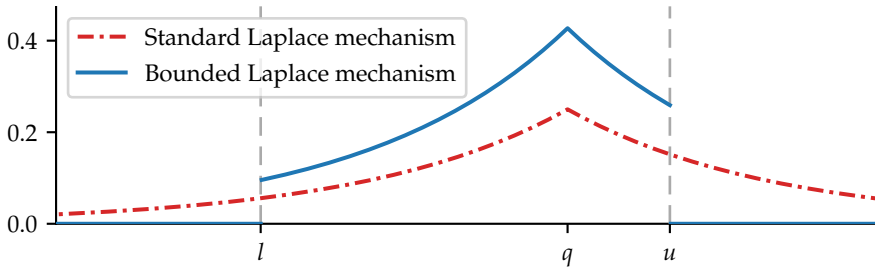

FIGURE 1. Comparison of the probability density functions (PDFs) of the standard and bounded Laplace mechanisms for a given domain $[l, u] \in \mathbb{R}$ and a mechanism input $q \in [l, u]$.

Example 1.1. Suppose we are querying a census dataset and seeking to learn the number of people born on Mars. Adding noise from a Laplace mechanism with variance $\frac{2}{\epsilon^2}$ will satisfy differential privacy. Although the real answer to the query is 0 (for now at least!), we must introduce uncertainty to protect the privacy of future human Martians. Successive outputs from the Laplace mechanism could be: -1.71, 2.31, -1.20, 0.652.

However bizarre the query, negative outputs are patently illogical and inconsistent. By the symmetry of the Laplace distribution, on average 50% of the outputs will be negative.

Currently there are two solutions to this drawback, both involving the selection of an appropriate output domain. The first solution, *truncation*, is to project values outside the domain to the closest value within the domain (see Appendix A.2). The second solution, *bounding* (shown in Figure 1), is to continue to sample independently from the mechanism until a value within the domain is returned. We refer the reader to Liu (2016) for a study of the statistical properties of the truncated and bounded Laplace mechanisms.

Example 1.2. Using the same set-up as Example 1.1, if the Laplace mechanism returns a value -1.71, the truncation method projects the output to 0 (the lower bound of a count query). If the bounding method is used, the value is simply re-sampled, meaning the second value 2.31 is returned (an analyst may subsequently wish to round this to 2).

By design, the truncated Laplace mechanism has a (possibly large) non-zero probability of returning values at the domain bounds. There are instances where this may be unsuitable and/or undesirable, such as when the domain bounds coincide with singularities or other qualitative changes in behaviour (*e.g.*, bifurcation points). In such instances, using the bounded Laplace mechanism may be more appropriate.

Example 1.3 (Truncation vs. bounding)**.** The naïve Bayes classifier is a probabilistic classifier that learns the means and variances of each feature (assumed independent) for each label, allowing Bayes' theorem to be applied to classify unseen examples. Differential privacy can be achieved by adding appropriately-scaled noise to these means and variances (Vaidya et al. (2013)). While the standard Laplace mechanism can be used to perturb the means, a mechanism with bounded outputs is required to perturb each variance, as variance must be non-negative.

In Figure 2 we evaluate the truncated and bounded Laplace mechanisms in perturbing each feature's variance, using the ϵ-differentially private bounded mechanism given in Section 3.2. We ran experiments on the iris flower dataset (Anderson (1936); Fisher (1936)), a dataset of 4 features, 3 labels and 150 observations. We split the privacy budget ϵ

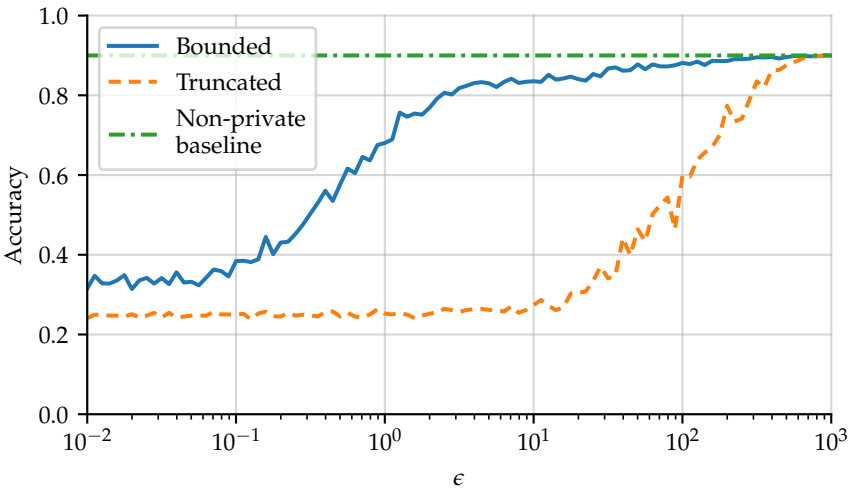

FIGURE 2. Comparison of accuracy versus ϵ for a differentially private naïve Bayes classifier on the Iris dataset using the bounded and truncated Laplace mechanisms for perturbation of each feature's variance. For each ϵ, the average accuracy over 100 simulations is shown.

evenly across each feature, and also across the perturbation of the mean and variance, giving a budget of $\frac{\epsilon}{8}$ for each perturbation. The standard Laplace mechanism was used to perturb each feature's mean, while each feature's variance was perturbed by the bounded and truncated Laplace mechanisms (with an output domain of $[0, 10^{10}]$). We adopted an 80%/20% train/test split, and, for each value of ϵ, took the mean accuracy of 100 simulations. The baseline (non-private) accuracy for the train/test split used in the simulations was 90%.

As shown in Figure 2, bounding outperforms truncation over a large range of ϵ. This can be attributed to the singularity produced in the Gaussian distribution at zero variance, the lower bound of the output domain. Because of the singularity, which occurs especially frequently for small ϵ, naïve Bayes is unable to classify correctly. We have observed similar behaviour for synthetic datasets, leading us to conclude that this behaviour is data-independent.

In this paper we show that the bounded Laplace mechanism *does not* typically satisfy differential privacy when inheriting parameters from the standard Laplace mechanism (Section 3). In fact, in almost all cases (except when the query sensitivity spans the domain, $\Delta Q = u - l$), the variance of the Laplace distribution must be increased for the bounded Laplace mechanism to satisfy the same differential privacy constraints. With a simple algorithm, we also show how to calculate the optimal noise scale for the bounded Laplace mechanism (Section 4).

Complete proofs to lemmas and theorems that are omitted in the main text are presented in the Appendix, alongside other additional material.

2. PRELIMINARIES

We first detail the notation to be used in this paper. Extended preliminaries are given in Appendix A.1.

We are interested in queries $Q : \mathcal{S}^n \to D$ on databases $\mathbf{d} \in \mathcal{S}^n$ of n rows mapping to a finite domain $D = [l, u]$ ($l < u$, both finite). The sensitivity of Q is defined in the usual way, $\Delta Q = \max_{\mathbf{d} \sim \mathbf{d}' \in \mathcal{S}^n} |Q(\mathbf{d}) - Q(\mathbf{d}')|$, where $\mathbf{d} \sim \mathbf{d}' \in \mathcal{S}^n$ denotes two datasets $\mathbf{d}, \mathbf{d}' \in \mathcal{S}^n$ that differ in exactly one row.

In this paper we are only concerned with *output perturbation* mechanisms, so we need only consider response mechanisms on the output domain of Q (*i.e.*, a random variable $Y_q : \Omega \to \mathbb{R}$ for each $q \in D$). Given $\epsilon \geq 0$ and $0 \leq \delta \leq 1$, the mechanism $\{Y_q \mid q \in D\}$ satisfies (ϵ,δ)-differential privacy when

$$\mathbb{P}(Y_q \in A) \leq e^{\epsilon} \mathbb{P}(Y_{q'} \in A) + \delta,$$

for all measurable $A \subseteq \mathbb{R}$ and whenever $|q - q'| \leq \Delta Q$.

We denote by $\text{Lap}(\mu, b)$ a Laplace distribution with mean μ and variance $2b^2$. The standard Laplace mechanism is therefore given by

$$Y_q = q + \text{Lap}(0, b) = \text{Lap}(q, b), \tag{2.1}$$

and satisfies (ϵ,δ)-differential privacy when $b \geq \frac{\Delta Q}{\epsilon - \log(1-\delta)}$ (Holohan et al. (2015)).

3. Bounded Laplace Mechanism

Truncation and bounding are two common approaches to overcoming out-of-domain outputs from the Laplace mechanism, as covered in Section 1. Details of the truncated Laplace mechanism are given in Appendix A.2.

Another approach is to bound the support of the response mechanism, and then sample directly from the output domain (*e.g.* by inverse transform sampling). This can also be achieved through rejection sampling, by continually redrawing from the unbounded distribution until an output falls within the domain. We will refer to this process as *bounding*, as the immediate outputs of the mechanism are bounded by design.

Definition 3.1 (Bounded Laplace Mechanism)**.** Given $b > 0$ and $D \subset \mathbb{R}$, the *bounded Laplace mechanism* $W_q : \Omega \to D$, for each $q \in D$, is given by its probability density function f_{W_q}:

$$f_{W_q}(x) = \begin{cases} 0, & \text{if } x \notin D, \\ \frac{1}{C_q} \frac{1}{2b} e^{-\frac{|x-q|}{b}}, & \text{if } x \in D, \end{cases}$$

where $C_q = \int_D \frac{1}{2b} e^{-\frac{|x-q|}{b}} dx$ is a normalisation factor.

Remark 1: It follows that $\mathbb{P}(W_q \in D) = 1$, and, conversely, that $\mathbb{P}(W_q \in \mathbb{R} \setminus D) = 0$.

Remark 2: Given $A \subseteq \mathbb{R}$, $\mathbb{P}(W_q \in A) = \frac{1}{C_q} \mathbb{P}(Y_q \in A \cap D)$, where Y_q is given in (2.1).

As the output distribution is now a function of the query answer $Q(\mathbf{d}) = q$, the normalisation factor C_q is no longer constant across $q \in D$. It is therefore no longer guaranteed that the mechanism W_q satisfies differential privacy using parameters from the standard Laplace mechanism.

3.1. **Preliminary Results.** We first establish an algebraic representation for C_q.

Lemma 3.2. *For C_q as given in Definition 3.1, and for $q \in D = [l, u]$,*

$$C_q = 1 - \frac{1}{2}\left(e^{-\frac{q-l}{b}} + e^{-\frac{u-q}{b}}\right).$$

We next consider the following lemma concerning C_q.

Lemma 3.3. *Let C_q be given by Definition 3.1. Then,*

$$\max_{\substack{q,q' \in D \\ |q'-q| \le \Delta Q}} \frac{C_{q'}}{C_q} e^{\frac{|q'-q|}{b}} = \frac{C_{l+\Delta Q}}{C_l} e^{\frac{\Delta Q}{b}}.$$

Proof. The following is an outline of the full proof given in Section A.3. By the symmetry of C_q about $\frac{u+l}{2}$, we can assume that $q' \ge q$. Showing that $\frac{\partial}{\partial z}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) \ge 0$ and $\frac{\partial}{\partial q}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) \le 0$ completes the proof. □

This leads us to the following definition of $\Delta C(b)$ for later use.

Definition 3.4. Given C_q from Definition 3.1, and noting that $C_q = C_q(b)$ is a function of b, we define $\Delta C(b)$ as follows:

$$\Delta C(b) = \frac{C_{l+\Delta Q}(b)}{C_l(b)}.$$

3.2. **Main Result.** We now proceed to the main result of this paper, which defines the scale parameter required for the bounded Laplace mechanism.

Theorem 3.5. *Let W_q be the bounded Laplace mechanism given in Definition 3.1 and let $\epsilon \ge 0$ and $0 \le \delta \le 1$ be given. Then $\{W_q \mid q \in D\}$ satisfies (ϵ,δ)-differential privacy whenever*

$$b \ge \frac{\Delta Q}{\epsilon - \log \Delta C(b) - \log(1-\delta)}. \tag{3.1}$$

Proof. The following is an outline of the full proof given in Section A.4. We are seeking to show that

$$\mathbb{P}(W_q \in A) \le e^{\epsilon}\mathbb{P}(W_{q'} \in A) + \delta,$$

for any measurable $A \subseteq D$ and where $q, q' \in D$, $|q - q'| \le \Delta Q$. For this to hold it is sufficient to show that $1 \le e^{\epsilon - \frac{|q'-q|}{b}} \frac{C_d}{C_{q'}} + \delta$. Furthermore by Lemma 3.3, it is sufficient to show that

$$1 \le \frac{1}{\Delta C(b)} e^{\epsilon - \frac{\Delta Q}{b}} + \delta,$$

which can be solved implicitly for b to complete the proof. □

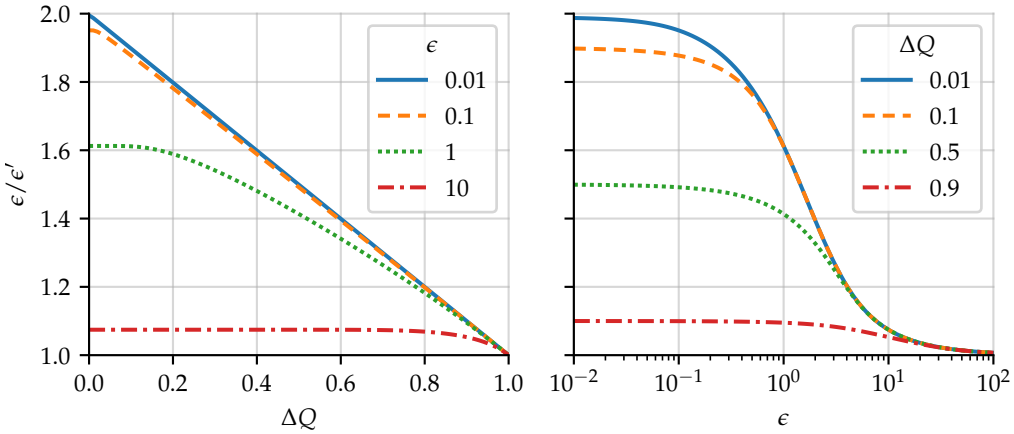

FIGURE 3. Relationship of $\frac{\epsilon}{\epsilon'}$ to ΔQ and ϵ, where $u - l = 1$ and $\delta = 0$ are fixed.

Discussion: To satisfy (ϵ,δ)-differential privacy using the bounded Laplace mechanism, its variance will never be less than that of the standard Laplace mechanism (since $\Delta C(b) \geq 1$). In the case of achieving ϵ-differential privacy (*i.e.* $\delta = 0$), the underlying Laplace distribution must be one which satisfies ϵ'-differential privacy, where $\epsilon' = \epsilon - \log \Delta C(b)$ (*i.e.* for a target ϵ, we require an effective ϵ'). As shown in Figure 3, the impact on ϵ' is most pronounced when ΔQ and ϵ are small, and that $2\epsilon' = \epsilon$ in the limiting case.

However, finding the optimal value for b is non-trivial since the relationship given in Theorem 3.5 is implicitly defined. This problem is examined in Section 4. The simpler task of determining a value of ϵ (or a relationship between ϵ and δ) for a given value of b can be achieved with (3.1).

4. CALCULATING b

From the conclusion of Theorem 3.5, the following fixed point operator can be defined for b.

Definition 4.1 (Fixed Point Operator)**.** Given $\Delta Q > 0$, $\epsilon \geq 0$ and $0 \leq \delta \leq 1$, we define the fixed point operator $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ by

$$f(b) = \frac{\Delta Q}{\epsilon - \log \Delta C(b) - \log(1 - \delta)}. \tag{4.1}$$

Any positive fixed point of f (*i.e.* $b^* = f(b^*) > 0$) will act as a differentially private scale parameter for the bounded Laplace mechanism. In advance of examining f, we first define

$$b_0 = \frac{\Delta Q}{\epsilon - \log(1 - \delta)}. \tag{4.2}$$

Note that b_0 determines the variance required for the standard Laplace mechanism to achieve (ϵ,δ)-differential privacy.

We now present a number of lemmas concerning f, namely: (i) the value of $f(b_0)$; and (ii) the monotonicity of f. Proofs are given in Appendices A.5 and A.6.

Lemma 4.2. *$f(b_0) \geq b_0$, and $f(b_0) = b_0$ if and only if $\Delta Q = u - l$.*

Lemma 4.3. *$f'(b) \leq 0$ whenever $b \neq 0$, and $f'(b) = 0$ if and only if $\Delta Q = u - l$.*

This leads us to the main result of this section, that f has a unique fixed point b^*.

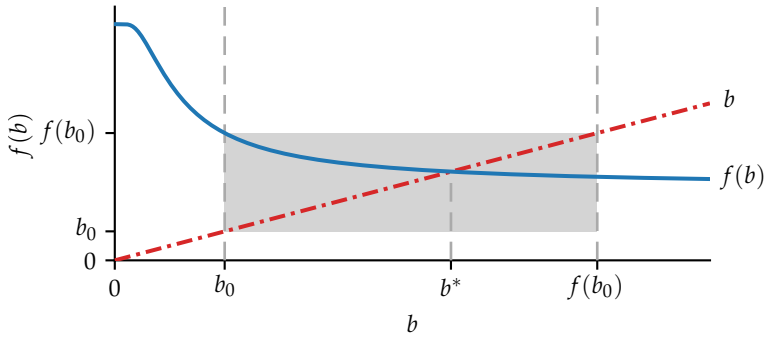

FIGURE 4. Outline of the fixed point approach of Algorithm 1 for finding $b^* \in [b_0, f(b_0)]$.

Theorem 4.4 (Fixed Point)**.** *There exists a unique* $b^* \in [b_0, f(b_0)]$ *such that* $b^* = f(b^*)$.

Proof. Since $f(b_0) \geq b_0$ (Lemma 4.2), $f' \leq 0$ (Lemma 4.3) and $f(b)$ is continuous on $b \in [b_0, \infty)$ (since it is differentiable), it follows that $f(b)$ has a unique fixed point $b^* \in [b_0, \infty)$, where uniqueness follows from the monotonicity of f.

Since $f' \leq 0$ and $f(b_0) \geq b_0$, it follows that $f(f(b_0)) \leq f(b_0)$. We must therefore have $b^* \in [b_0, f(b_0)]$. □

It follows from Theorem 4.4 that the mechanism W_q from Definition 3.1 satisfies differential privacy for $b = b^*$. Given that we have a bounded domain in which b^* lies, and since f is continuous, the bisection method gives robust convergence to b^* within machine precision, for any given $\epsilon \geq 0$, $0 \leq \delta \leq 1$, $u > l$ and $\Delta Q \leq u - l$. A sample algorithm for calculating b^* is given in Algorithm 1, while Figure 4 gives a graphical outline of the approach.

The following corollary to Theorem 4.4 shows that in the case of $\Delta Q = u - l$ (and only in this case), differential privacy is achieved when using the scale parameter b_0 from the standard Laplace mechanism. Conversely, when $\Delta Q < u - l$, we must have $b^* > b_0$, and hence the bounded Laplace mechanism requires greater noise to achieve differential privacy than the standard Laplace mechanism.

Corollary 4.5. $b^* = b_0$ *if and only if* $\Delta Q = u - l$.

Proof. By Theorem 4.4, $b^* \in [b_0, f(b_0)]$. However, by Lemma 4.2, $f(b_0) = b_0$ if and only if $\Delta Q = u - l$, hence $b^* \in [b_0, b_0]$ and the result follows. □

This final corollary confirms that any fixed point b^* is the lower bound of all values b that satisfy (ϵ,δ)-differential privacy, in line with (3.1).

Corollary 4.6. *Let* $b^* \in \mathbb{R}_{>0}$ *such that* $b^* = f(b^*)$. *Then, given any* $\xi > 0$,

$$b^* + \xi > f(b^* + \xi).$$

Proof. By Theorem 4.4, such a fixed point b^* exists. Furthermore, from Lemma 4.3 we have $f'(b) < 0$, hence

$$f(b^* + \xi) < f(b^*) = b^* < b^* + \xi.$$

□

Algorithm 1: A robust and precise method for finding b^*

input : Fixed point operator `f` (as given in (4.1)), b_0 as given (4.2)
output : Fixed point b^*, where `f`$(b^*) = b^*$

left $\leftarrow b_0$;
right $\leftarrow$ `f`(b_0);
intervalSize $\leftarrow$ (left + right) $\times$ 2;

while intervalSize $>$ right $-$ left **do**
  intervalSize $\leftarrow$ right $-$ left;
  b $= \frac{\text{left+right}}{2}$;
  if `f`(b) $\geq$ b **then**
    left $\leftarrow$ b;
  end
  if `f`(b) $\leq$ b **then**
    right $\leftarrow$ b;
  end
end

return b;

5. Related Work

In Liu (2016), the statistical properties of bounding and truncating the Laplace mechanism were explored, without examining the differential privacy properties of the bounded Laplace mechanism. The same author followed with a study on generalised Gaussian mechanisms for differential privacy, Liu (2018). The results applied to the bounded Laplace mechanism showed a doubling of the noise variance ($\epsilon = 2\epsilon'$) is required, an increase Figure 3 shows to be excessive.

Researchers at Google have considered a different form of a bounded Laplace mechanism in Geng et al. (2018) in order to achieve approximate (ϵ,δ)-differential privacy. The authors were investigating their use of a bounded Laplace mechanism as an alternative to the Gaussian mechanism, which relies on non-zero δ. Their solution was to bound the magnitude of noise added to each query output (rather than bounding the domain of all noisy outputs), while using the 'probability of error' δ to cover the non-overlapping output spaces of neighbouring datasets. The bound on the noise added to each value is not chosen in advance by the user, but instead is calculated as a function of ϵ and δ.

In Zhang et al. (2012), regression analysis under differential privacy was studied. The authors looked to add noise (using the Laplace mechanism) to the coefficients of an objective function to achieve differential privacy, but this can result in an unbounded objective function. Their first approach at solving this was to re-run the differential privacy mechanism until the result gives a solution to the optimisation problem. This approach has the effect of doubling the noise variance (since $\epsilon = 2\epsilon'$), which our work has shown may be excessive.

A naïve Bayes machine learning classifier was described in Vaidya et al. (2013), which achieves differential privacy by adding Laplace noise to the model parameters. For numerical data, naïve Bayes calculates the mean and standard deviation of each feature in order to classify unseen data. The authors propose re-sampling from the Laplace distribution to

ensure the differentially private standard deviations are positive, without modifying the variance of the Laplace distribution. By Theorem 4.4 and Corollary 4.5 we now know that this approach does not satisfy differential privacy, since $\Delta Q < u - l$ in this case. We make use of this work, with minor amendments to ensure differential privacy, in Example 1.3.

Consistency in differential privacy has also been studied previously. Examples include achieving consistent releases of marginals, Barak et al. (2007), and histograms, Hay et al. (2010). In Barak et al. (2007) the authors sought to release marginals consisting of non-negative integers, with consistent sums across marginals. This was achieved using Fourier transformations and linear programming. In Hay et al. (2010), the authors used *constrained inference* to ensure consistency in histogram counts through post-processing. The consistency of marginals was also studied in Zhang et al. (2018), for the case of local privacy.

6. Conclusion

In this paper, we have shown that the bounded Laplace mechanism does not typically satisfy differential privacy when inheriting parameters from the standard Laplace mechanism (Theorem 3.1). We have also presented details of calculating the required parameters for the bounded Laplace mechanism to satisfy differential privacy (Algorithm 1). We showed that only in the case of $\Delta Q = u - l$ can we use the same scale parameter from the standard Laplace mechanism for the bounded mechanism to satisfy differential privacy (Corollary 4.5).

The results of this paper highlight the dangers of re-sampling from the Laplace mechanism in applications of differential privacy to achieve valid/plausible outputs. Researchers may be inadvertently violating differential privacy in doing so, or overcompensating by increasing the privacy budget excessively. Our robust method of calculating the optimal noise variance will allow privacy researchers and practitioners to deploy the bounded Laplace mechanism with confidence and certainty.

References

E. Anderson. The species problem in iris. *Annals of the Missouri Botanical Garden*, 23(3): 457–509, 1936.

B. Barak, K. Chaudhuri, C. Dwork, S. Kale, F. McSherry, and K. Talwar. Privacy, accuracy, and consistency too: a holistic solution to contingency table release. In *Proceedings of the twenty-sixth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 273–282. ACM, 2007.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography*, pages 265–284. Springer, 2006.

R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2):179–188, 1936. doi: 10.1111/j.1469-1809.1936.tb02137.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1936.tb02137.x.

Q. Geng, W. Ding, R. Guo, and S. Kumar. Truncated Laplacian mechanism for approximate differential privacy. *cs.CR*, abs/1810.00877, 2018. URL http://arxiv.org/abs/1810.00877.

M. Hay, V. Rastogi, G. Miklau, and D. Suciu. Boosting the accuracy of differentially private histograms through consistency. *Proc. VLDB Endow.*, 3(1-2):1021–1032, Sept. 2010. ISSN 2150-8097. doi: 10.14778/1920841.1920970. URL http://dx.doi.org/10.14778/1920841.1920970.

N. Holohan, D. J. Leith, and O. Mason. Differential privacy in metric spaces: Numerical, categorical and functional data under the one roof. *Information Sciences*, 305:256–268, 2015. ISSN 0020-0255. doi: http://dx.doi.org/10.1016/j.ins.2015.01.021.

F. Liu. Statistical properties of sanitized results from differentially private Laplace mechanisms with noninformative bounding. *ArXiv e-prints*, 1607.08554 [stat.ME], July 2016.

F. Liu. Generalized Gaussian mechanism for differential privacy. *IEEE Transactions on Knowledge and Data Engineering*, page In press, 2018. ISSN 1041-4347. doi: 10.1109/TKDE.2018.2845388.

J. Vaidya, B. Shafiq, A. Basu, and Y. Hong. Differentially private naïve Bayes classification. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 01*, WI-IAT '13, pages 571–576. IEEE Computer Society, 2013. ISBN 978-0-7695-5145-6. doi: 10.1109/WI-IAT.2013.80. URL http://dx.doi.org/10.1109/WI-IAT.2013.80.

J. Zhang, Z. Zhang, X. Xiao, Y. Yang, and M. Winslett. Functional mechanism: Regression analysis under differential privacy. *Proc. VLDB Endow.*, 5(11):1364–1375, July 2012. ISSN 2150-8097. doi: 10.14778/2350229.2350253. URL http://dx.doi.org/10.14778/2350229.2350253.

Z. Zhang, T. Wang, N. Li, S. He, and J. Chen. CALM: Consistent Adaptive Local Marginal for marginal release under local differential privacy. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, CCS '18, pages 212–229, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5693-0. doi: 10.1145/3243734.3243742. URL http://doi.acm.org/10.1145/3243734.3243742.

Appendix

A.1. **Extended Preliminaries.** We begin with a database $\mathbf{d} \in \mathcal{S}^n$, where $\mathcal{S}$ is the domain of each row of the dataset. A query is a map $Q : \mathcal{S}^n \to D$, where $D = [l, u] \subset \mathbb{R}$ is a finite interval on the real line ($l < u$, both finite). Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we denote a response mechanism by $X_{Q,\mathbf{d}} : \Omega \to \mathbb{R}$ for each $\mathbf{d} \in \mathcal{S}^n$. $X_{Q,\mathbf{d}}$ is a random variable taking the query Q and the database $\mathbf{d}$ as inputs, and produces a *randomised* output in $\mathbb{R}$.

Definition A.1 (Differential Privacy)**.** Given a query $Q : \mathcal{S}^n \to D$, $\epsilon \geq 0$ and $0 \leq \delta \leq 1$, a response mechanism $\{X_{Q,\mathbf{d}} | \mathbf{d} \in \mathcal{S}^n\}$ satisfies (ϵ,δ)-differential privacy if

$$\mathbb{P}(X_{Q,\mathbf{d}} \in A) \leq e^{\epsilon}\mathbb{P}(X_{Q,\mathbf{d}'} \in A) + \delta,$$

for all $\mathbf{d} \sim \mathbf{d}' \in \mathcal{S}^n$, and all measurable $A \subseteq \mathbb{R}$, and where $\mathbf{d} \sim \mathbf{d}' \in \mathcal{S}^n$ denotes that the databases $\mathbf{d}, \mathbf{d}' \in \mathcal{S}^n$ differ in exactly one row (*i.e.* $\exists! j \in [n] : d_j \neq d'_j$).

When $\delta = 0$ we say $\{X_{Q,\mathbf{d}}\}$ satisfies ϵ-differential privacy.

Throughout this paper, we assume that ϵ and δ are not simultaneously zero, *i.e.* $\frac{e^{\epsilon}}{1-\delta} > 1$.

In this paper we focus on *output perturbation* response mechanisms, where the raw query answer is randomised. Using the notation introduced above, we are interested in mechanisms of the form

$$X_{Q,\mathbf{d}}(\omega) = Y_{Q(\mathbf{d})}(\omega),$$

where $Y_q : \Omega \to \mathbb{R}$ is defined for each $q \in D$ (producing a randomised output for each possible query answer q).

The *Laplace mechanism* was first proposed in Dwork et al. (2006) to achieve privacy using additive noise. We first define the sensitivity of Q, $\Delta Q \in \mathbb{R}_{\geq 0}$, given by

$$\Delta Q = \max_{\mathbf{d} \sim \mathbf{d}' \in \mathcal{S}^n} |Q(\mathbf{d}) - Q(\mathbf{d}')|. \tag{A.1}$$

Note that $0 \leq \Delta Q \leq u - l$. The following is a form of the Laplace mechanism which is applicable to (ϵ,δ)-differential privacy (Holohan et al., 2015, Example 5).

Definition A.2 (Laplace Mechanism)**.** The Laplace distribution with mean μ and variance $2b^2$, $\text{Lap}(\mu, b)$, has a PDF $f_{\text{Lap}(\mu,b)} : \mathbb{R} \to \mathbb{R}_{\geq 0}$ given by

$$f_{\text{Lap}(\mu,b)}(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}.$$

The response mechanism $\{Y_q \mid q \in D\}$ (where $q = Q(\mathbf{d}) \in D$ is the raw query answer) given by

$$Y_q(\omega) = q + \text{Lap}(0, b) = \text{Lap}(q, b) \tag{A.2}$$

is known as the *Laplace mechanism*, and satisfies (ϵ,δ)-differential privacy when

$$b \geq \frac{\Delta Q}{\epsilon - log(1 - \delta)}.$$

Remark: Whenever $b > 0$, $\mathbb{P}(Y_q \in \mathbb{R} \setminus D) > 0$, *i.e.* some outputs will lie outside D.

A.2. **Truncated Laplace mechanism.** Truncation is one option available to ensure the output of a response mechanism falls within a required domain. This involves a deterministic mapping to the upper/lower bounds of the output domain, if the value falls outside the domain.

Definition A.3 (Truncated Laplace Mechanism)**.** Given a Laplace mechanism $Y_q : \Omega \to \mathbb{R}$ given in (A.2), the *truncated Laplace mechanism* is given by $Z_q(\omega) = \text{trunc}(Y_q(\omega))$, where $\text{trunc} : \mathbb{R} \to D$ is defined as

$$\text{trunc}(r) = \begin{cases} r, & \text{if } r \in D, \\ l, & \text{if } r < l, \\ u & \text{if } r > u. \end{cases} \tag{A.3}$$

The truncated Laplace mechanism trivially satisfies differential privacy, as it is a deterministic post-processing of a differentially private output.

A.3. **Proof of Lemma 3.3.** In order to prove Lemma 3.3, we must first consider the following lemmas concerning C_q.

Lemma A.4. *Let $q \in D$ and $b > 0$, and let C_q be given by Definition 3.1. Then $\frac{\partial}{\partial z}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) \geq 0$, whenever $q + z \leq u$.*

Proof. We first note that

$$\frac{\partial}{\partial z} C_{q+z} = \frac{1}{2b}\left(e^{-\frac{q+z-l}{b}} - e^{-\frac{u-q-z}{b}}\right).$$

We then see that

$$\frac{\partial}{\partial z}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) = \frac{1}{C_q}\frac{1}{b}\left(1 - e^{-\frac{u-q-z}{b}}\right) e^{\frac{z}{b}}.$$

Since $b > 0$ by assumption, it follows that $\frac{\partial}{\partial z}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) \geq 0$ if and only if $q + z \leq u$. □

Lemma A.5. *Let $q \in D$ and $z \geq 0$, and let C_q be given by Definition 3.1. Then $\frac{\partial}{\partial q}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) \leq 0$.*

Proof. We first note that

$$\frac{\partial}{\partial q} C_{q+z} = \frac{1}{2b}\left(e^{-\frac{q+z-l}{b}} - e^{-\frac{u-q-z}{b}}\right).$$

We then find

$$\begin{aligned}
\frac{\partial}{\partial q}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) &= \frac{e^{\frac{z}{b}}}{{C_q}^2}\left(C_q \frac{\partial}{\partial q} C_{q+z} - C_{q+z}\frac{\partial}{\partial q} C_q\right) \\
&= \frac{e^{\frac{z}{b}}}{2b\,{C_q}^2}\left(e^{-\frac{q-l}{b}}\left(e^{-\frac{z}{b}} - 1\right) + e^{-\frac{u-q}{b}}\left(1 - e^{\frac{z}{b}}\right) + e^{-\frac{u-l-z}{b}} - e^{-\frac{u-l+z}{b}}\right) \\
&= \frac{e^{\frac{z}{b}}\left(\left(e^{-\frac{z}{b}} - 1\right)\left(e^{\frac{u-q}{b}} - 1\right) + \left(1 - e^{\frac{z}{b}}\right)\left(e^{\frac{q-l}{b}} - 1\right)\right)}{2b\, e^{\frac{u-l}{b}} {C_q}^2}.
\end{aligned}$$

Since $b > 0$, it's clear that the denominator is positive. Furthermore, since $q \in D$, it follows that $e^{\frac{u-q}{b}}, e^{\frac{q-l}{b}} > 1$. Also, since $z \geq 0$ by assumption, we have $e^{-\frac{z}{b}} < 1$ and $e^{\frac{z}{b}} > 1$. Hence, $\frac{\partial}{\partial q}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) \leq 0$, as required. □

Using Lemmas A.4 and A.5, we can now prove Lemma 3.3.

Proof (Lemma 3.3). Since C_q is symmetric about $\frac{u+l}{2}$, we have $C_q = C_{u+l-q}$. By letting $q_0 = u+l-q$ and $q_0' = u+l-q'$, then, $\frac{C_{q'}}{C_q} e^{\frac{|q'-q|}{b}} = \frac{C_{q_0'}}{C_{q_0}} e^{\frac{|q_0'-q_0|}{b}}$, and $q' > q$ if $q_0' < q_0$. Hence, without loss of generality we can assume that $q' \geq q$, so we are examining

$$\max_{\substack{q,q' \in D \\ 0 \leq q'-q \leq \Delta Q}} \frac{C_{q'}}{C_q} e^{\frac{q'-q}{b}}.$$

Equivalently, since $q' \geq q$, we can consider $\max_{\substack{q \in D \\ 0 \leq z \leq \Delta Q}} \frac{C_{q+z}}{C_q} e^{\frac{z}{b}}$.

By Lemma A.5, $\frac{\partial}{\partial q}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) \leq 0$, hence the maximum is attained at the smallest possible q, *i.e.*

$$\max_{\substack{q \in D \\ 0 \leq z \leq \Delta Q}} \frac{C_{q+z}}{C_q} e^{\frac{z}{b}} = \max_{0 \leq z \leq \Delta Q} \frac{C_{l+z}}{C_l} e^{\frac{z}{b}}.$$

Similarly, by Lemma A.4, $\frac{\partial}{\partial z}\left(\frac{C_{q+z}}{C_q} e^{\frac{z}{b}}\right) \geq 0$, hence the maximum is attained at the largest possible z, giving $\max_{\substack{q \in D \\ 0 \leq z \leq \Delta Q}} \frac{C_{q+z}}{C_q} e^{\frac{z}{b}} = \frac{C_{l+\Delta Q}}{C_l} e^{\frac{\Delta Q}{b}}$, as required. □

A.4. **Proof of Theorem 3.5.**

Proof (Theorem 3.5). We follow a similar method of proof as used in Example 5 of Holohan et al. (2015).

Given $A \subseteq D$, and noting that $\mathbb{P}(W_q \in A) = \frac{1}{C_q}\mathbb{P}(Y_q \in A)$, where Y_q is given by (2.1), we are seeking to show that

$$\frac{1}{C_q}\mathbb{P}(Y_q \in A) \leq e^{\epsilon}\frac{1}{C_{q'}}\mathbb{P}(Y_{q'} \in A) + \delta,$$

for any measurable $A \subseteq D$ and where $q, q' \in D$ and $|q - q'| \leq \Delta Q$. Given that $\mathbb{P}(Y_q \in A) = \int_A \frac{e^{-\frac{|x-q|}{b}}}{2b} dx$, we have,

$$\frac{1}{C_q}\int_A \frac{e^{-\frac{|x-q|}{b}}}{2b} dx \leq e^{\epsilon}\frac{1}{C_{q'}}\int_A \frac{e^{-\frac{|x-q'|}{b}}}{2b} dx + \delta.$$

Using the triangle inequality, we see that $|x-q'| \leq |x-q|+|q'-q|$, so it is sufficient to show that
$\frac{1}{C_q}\int_A \frac{e^{-\frac{|x-q|}{b}}}{2b} dx \leq e^{\epsilon - \frac{|q'-q|}{b}}\frac{1}{C_{q'}}\int_A \frac{e^{-\frac{|x-q|}{b}}}{2b} dx + \delta$, or equivalently,

$$1 \leq e^{\epsilon - \frac{|q-q'|}{b}}\frac{C_q}{C_{q'}} + \frac{C_q}{\int_A \frac{e^{-\frac{|x-q|}{b}}}{2b} dx}\delta.$$

Since $A \subseteq D$ and given the definition of C_q in Definition 3.1, it follows that $C_q \geq \int_A \frac{e^{-\frac{|x-q|}{b}}}{2b} dx$, hence it is sufficient to show that $1 \leq e^{\epsilon - \frac{|q'-q|}{b}}\frac{C_d}{C_{q'}} + \delta$.

By Lemma 3.3, $\Delta C(b)\, e^{\frac{\Delta Q}{b}} \geq \frac{C_{q'}}{C_q} e^{\frac{|q'-q|}{b}}$ when $|q'-q| \leq \Delta Q$, or equivalently $\frac{1}{\Delta C(b)} e^{-\frac{\Delta Q}{b}} \leq \frac{C_q}{C_{q'}} e^{-\frac{|q'-q|}{b}}$, so it is sufficient to show that

$$1 \leq \frac{1}{\Delta C(b)} e^{\epsilon - \frac{\Delta Q}{b}} + \delta. \tag{A.4}$$

Solving (A.4) implicitly for b completes the proof. □

A.5. **Proof of Lemma 4.2.**

Proof (Lemma 4.2). We first note that $f(b) > 0$ if and only if $\epsilon - \log \Delta C(b) - \log(1-\delta) > 0$, or equivalently, if $\Delta C(b) < \frac{e^\epsilon}{1-\delta}$. We assume that $\frac{e^\epsilon}{1-\delta} > 1$ (*i.e.* that ϵ and δ are not simultaneously zero).

Given $b_0 = \frac{\Delta Q}{\epsilon - \log(1-\delta)}$, we see that

$$\begin{aligned}
\Delta C(b_0) &= \frac{2 - e^{-\epsilon + \log(1-\delta)} - e^{-\left(\frac{u-l}{\Delta Q} - 1\right)(\epsilon - \log(1-\delta))}}{1 - e^{-\frac{u-l}{\Delta Q}(\epsilon - \log(1-\delta))}} \\
&= \frac{2 - \frac{1-\delta}{e^\epsilon} - \left(\frac{e^\epsilon}{1-\delta}\right)^{1 - \frac{u-l}{\Delta Q}}}{1 - \left(\frac{e^\epsilon}{1-\delta}\right)^{-\frac{u-l}{\Delta Q}}} \\
&= \frac{2\left(\frac{e^\epsilon}{1-\delta}\right) - 1 - \left(\frac{e^\epsilon}{1-\delta}\right)^{2 - \frac{u-l}{\Delta Q}}}{\frac{e^\epsilon}{1-\delta} - \left(\frac{e^\epsilon}{1-\delta}\right)^{1 - \frac{u-l}{\Delta Q}}}.
\end{aligned} \tag{A.5}$$

For simplicity, we relabel (A.5) by setting $\alpha = \frac{e^\epsilon}{1-\delta}$ and $\beta = \frac{u-l}{\Delta Q}$, giving

$$\Delta C(b_0) = \frac{2\alpha - 1 - \alpha^{2-\beta}}{\alpha - \alpha^{1-\beta}}.$$

We note that $\alpha > 1$ and $\beta \geq 1$.

Since $\max\left(2\alpha - \alpha^2\right) = 1$ and the maximum occurs at $\alpha = 1$, it follows that $2\alpha - \alpha^2 < 1$ when $\alpha > 1$. We can then make the following series of deductions:

$$\begin{aligned}
2\alpha - \alpha^2 &< 1, \\
2\alpha - 1 &< \alpha^2, \\
2\alpha - 1 - \alpha^{2-\beta} &< \alpha^2 - \alpha^{2-\beta}, \\
\frac{2\alpha - 1 - \alpha^{2-\beta}}{\alpha - \alpha^{1-\beta}} &< \alpha.
\end{aligned}$$

Hence,

$$\Delta C(b_0) < \alpha = \frac{e^\epsilon}{1-\delta},$$

and it follows that $f(b_0) > 0$.

We can also show that $\Delta C(b_0) \geq 1$ through the following series of deductions:

$$\begin{aligned} \alpha^{1-\beta} &\leq 1, \\ \alpha^{1-\beta}(\alpha - 1) &\leq \alpha - 1, \\ 0 \leq \alpha - \alpha^{1-\beta} &\leq 2\alpha - 1 - \alpha^{2-\beta}, \\ \frac{2\alpha - 1 - \alpha^{2-\beta}}{\alpha - \alpha^{1-\beta}} &\geq 1. \end{aligned} \tag{A.6}$$

Hence, $\log \Delta C(b_0) \geq 0$. It then follows that

$$\frac{\Delta Q}{\epsilon - \log \Delta C(b_0) - \log(1-\delta)} \geq \frac{\Delta Q}{\epsilon - \log(1-\delta)},$$

and that $f(b_0) \geq b_0$. Furthermore, from (A.6), $f(b_0) = b_0$ if and only if $\Delta Q = u - l$. □

A.6. **Proof of Lemma 4.3.**

Proof (Lemma 4.3). From (4.1), we have

$$f'(b) = \frac{f(b)^2}{\Delta Q \Delta C(b)} \frac{\partial \Delta C(b)}{\partial b},$$

hence $f' \leq 0$ if and only if $\frac{\partial \Delta C(b)}{\partial b} \leq 0$. From the definition of $\Delta C(b)$, after some simplification we have

$$\begin{aligned} \frac{\partial \Delta C(b)}{\partial b} &= -\left(\frac{1}{2b\, C_l(b)}\right)^2 \left(\Delta Q \left(e^{-\frac{\Delta Q}{b}} + e^{-\frac{2(u-l)-\Delta Q}{b}}\right) + e^{-\frac{u-l}{b}}(u - l - \Delta Q)\left(e^{\frac{\Delta Q}{b}} + e^{-\frac{\Delta Q}{b}}\right) \right. \\ &\qquad \left. - 2(u-l)e^{-\frac{u-l}{b}}\right) \\ &\leq -\left(\frac{1}{2b\, C_l(b)}\right)^2 \left(\Delta Q \left(e^{-\frac{\Delta Q}{b}} + e^{-\frac{2(u-l)-\Delta Q}{b}}\right) - 2\Delta Q e^{-\frac{u-l}{b}}\right) \qquad \text{(A.7)} \\ &= -\left(\frac{1}{2b\, C_l(b)}\right)^2 \Delta Q e^{-\frac{\Delta Q}{b}} \left(1 - e^{-\frac{u-l-\Delta Q}{b}}\right)^2 \\ &\leq 0, \end{aligned}$$

where (A.7) follows since $e^a + e^{-a} \geq 2$ for all $a \in \mathbb{R}$. Note that we have $\frac{\partial \Delta C(b)}{\partial b} = 0$ if and only if $\Delta Q = u - l$. Also note that this result holds for all $b \neq 0$, and therefore for all $b \geq b_0$.

We therefore conclude that $f'(b) \leq 0$ for all $b \geq b_0$, and furthermore that $f'(b) = 0$ if and only if $\Delta Q = u - l$. □