Statistical Properties of Sanitized Results from Differentially Private Laplace Mechanism with Univariate Bounding Constraints

Fang Liu*

* Department of Applied and Computational Mathematics and Statistics

University of Notre Dame, Notre Dame, IN 46556

E-mail: fang.liu.131@nd.edu

Abstract. Protection of individual privacy is a common concern when releasing and sharing data and information. Differential privacy (DP) formalizes privacy in probabilistic terms without making assumptions about the background knowledge of data intruders, and thus provides a robust concept for privacy protection. Practical applications of DP involve development of differentially private mechanisms to generate sanitized results at a pre-specified privacy budget. For the sanitization of statistics with publicly known bounds such as proportions and correlation coefficients, the bounding constraints will need to be incorporated in the differentially private mechanisms. There has been little work on examining the consequences of the bounding constraints on the accuracy of sanitized results and the statistical inferences of the population parameters based on the sanitized results. In this paper, we formalize the differentially private truncated and boundary inflated truncated (BIT) procedures for releasing statistics with publicly known bounding constraints. The impacts of the truncated and BIT Laplace procedures on the statistical accuracy and validity of sanitized statistics are evaluated both theoretically and empirically via simulation studies.

Keywords. truncated mechanism, boundary inflated truncated (BIT) mechanism, bias, consistency, mean squared error; global and data-invariant

1 Introduction

Protection of individual privacy is always a concern when releasing and sharing information. A data release mechanism aims to provide useful information to the public without compromising individual privacy. Differential privacy (DP) is a concept developed in theoretical computer science [1–3] and has gained great popularity in recent years in both theoretical research and practical applications. DP formalizes privacy in mathematical terms without making assumptions about the background knowledge of data intruders and thus provides a robust concept for privacy protection. Practical applications of DP involve development of differentially private mechanisms, also referred to as sanitizers, through which original results are processed and converted to results that do not reveal individual information at a pre-specified privacy budget. There are general differentially

private mechanisms such as the Laplace mechanism [1], the Exponential mechanism [4, 5], and more recently, the staircase mechanism [6], the generalized Gaussian mechanism [7], and the adaptive mechanisms such as the multiplicative weighting mechanism [8] and the median mechanism [9] for sanitizing multiple correlated queries. There are also differentially private mechanisms targeting specifically at certain statistical analyses such as robust and efficient point estimators [10, 3], principle component analysis [11], linear and penalized regression [12, 13], Bayesian inferences of probabilistic graphical models [14], machine learning, data mining, and big data analytics in genomics, healthcare, biometrics [15–19], among others.

In the context of DP, it is sometimes assumed that data and statistics (numerical query results) are bounded. The bounding assumption is supported from a technical perspective as well as justified from a practical point of view. First, some statistics are naturally bounded per definition, such as proportions (bounded by $[0, 1]$) and correlation coefficients (bounded by $[-1, 1]$). Second, some statistics are required be bounded in order to be sanitized by some of the common differentially private mechanisms while ensuring some degree of usefulness of the sanitized results. For example, the scale parameter of the Laplace mechanism is proportional to the global sensitivity of a statistic. Suppose the statistic is the sample mean of a numerical attribute; then the value of the attribute needs to be bounded on both ends to have a finite global sensitivity for the mean so that it can be sanitized by the Laplace mechanism in a meaningful way. Third, real-life data in general support the assumption of bounded data when the assumption is needed. Though bounded numerical attributes in statistical parametric modelling are often modelled via distributions with unbounded domains (e.g., Gaussian or Poisson assumptions), those distributional assumptions are in many cases only approximate and the probabilities of out-of-bounds values are often small enough to be ignorable under these distributional assumptions. For instance, it is safe to say human height is bounded within $(0, 300)$cm. Though it is often modelled by Gaussian distributions with support $[-\infty, \infty]$, $\Pr(\text{height} < 0 \text{ cm or } > 300 \text{ cm}) \approx 0$ under the Gaussian assumption. If all numerical attributes in a data set are bounded, descriptive or inferential statistics based on the data in general are also bounded. For example, if a numerical attribute is bounded within $[c_0, c_1]$, then its sample mean is bounded within $[c_0, c_1]$, and its variance is bounded within $[0, n(c_1 - c_0)^2/(4(n-1))]$ for a given sample size n [20]. Besides the above examples on the univariate bounding constraint per statistic, there are also many types of multivariable bounding constraints. For example, a linear combination of multiple statistics is bounded or the sum of a proportion vector is 1. In the following discussion, we focus our theoretical and empirical analysis on the univariate bounding and discuss a couple examples of multivariable constraints briefly; more in-depth investigation on the latter topic will be conducted in the future.

Modifications are often needed in some commonly used differentially private mechanisms in order to accommodate the release of bounded statistics without compromising the pre-specified DP. For example, both the Laplace and Gaussian mechanisms release sanitized results from the real line $(-\infty, \infty)$ and do not automatically deal with bounding constraints. Some practitioners choose to ignore the bounds and release the raw sanitized results as is. We would not recommend this approach since the out-of-bounds values carry no practical meaning and eventually will be discarded by data users. Another approach is to formulate the problem as a constrained optimization problem with inequality constraints in the general framework of constrained inference in DP. The constrained inference approach concerns about finding a set of sanitized results that are optimal estimators of the unconstrained sanitized results by some criteria, such as the l_2 distance, subject to a set of predefined and known constraints. Therefore, "inferences" in this context is not the same as

the classical statistical inferences, which is about inferring population parameters via point and interval estimations and hypothesis testing given finite sample data. The constrained inference can deal with both inequality and equality constraints. For instance, proportions $\mathbf{p} = (p_1, \ldots, p_K)^T$ are subject to both the equality constraint $\sum_{j=1}^{K} p_j = 1$ and the inequality constraint $p_j \in [0, 1]$ for $j = 1, \ldots, K$. Barak et al. [21] employed linear programming (and the Fourier transformation) to obtain a non-negative and consistent sanitized contingency tables. Hay et al. [22] boosted the accuracy of sanitized histograms (measured by the mean squared error between the sanitized and original results) by incorporating the prior rank constraint in unattributed histograms and the equality constrained in universal histograms to increase the accuracy of lower-order marginals. Qardaji et al. [23] showed that combining the choice of a good branching factor with constrained inference can further boost the accuracy of a sanitized histogram. Li et al. [24] investigated an extension to the matrix mechanism they proposed that incorporates non-negativity constraints when realizing count queries. While the constrained inference approach provides a general framework to deal with constraints in DP, all the above mentioned work deals with count data. In addition, the approach can be analytically and computationally demanding, depending on the statistics and the objective functions employed. We will save the in-depth investigation of this approach as a future research topic.

In this paper, we focus on examining two straightforward methods that appeal to practitioners for sanitizing statistics with bounding constraints. One approach, referred to as the *truncation* procedure, throws the out-of-bounds sanitized values away and re-sanitizes until a within-bound value is obtained. This procedure can also be realized by sampling sanitized results directly from a truncated distribution that satisfies DP. The other approach legitimizes the out-of-bounds sanitized results by setting the out-of-bounds values at the boundaries, referred as the *boundary inflated truncation/BIT* procedure. This procedure can also be realized by sampling sanitized results directly from a piece-wise distribution with probability mass at the boundaries. We assess their impact on the statistical accuracy of sanitized results vs the original results both theoretically and empirically in the context of the Laplace mechanism. To the best of our knowledge, this is the first work on the statistical properties of the BIT and truncation bounding procedures in the context of DP.

The rest of the paper is organized as follows. Section 2 overviews some of the key concepts in DP. Section 3 presents the truncation and BIT bounding procedures and investigates the impact of the BIT and truncation bounding procedures on the utility of sanitized results in terms of bias, mean squared error, and consistency. Section 4 illustrates the applications of the truncated and BIT Laplace procedures and examines the statistical properties of the sanitized results in two simulation studies. The paper concludes in Section 5 with some final remarks and plans for future works.

2   Preliminaries

2.1   Definition of Differential Privacy

Let $\Delta(\mathbf{x}, \mathbf{x}') = 1$ denote that data $\mathbf{x}'$ differs from $\mathbf{x}$ by only one individual. There are two definitions on "differing by one". Briefly, one refers to the case that the $\mathbf{x}'$ and $\mathbf{x}$ have the same sample size n, but one and only one record differs in at least one attributes (a substitution would make the two data sets identical). In the second definition, $\mathbf{x}'$ has one less record that $\mathbf{x}$, so the sample sizes for $\mathbf{x}$ is n and $n-1$ for $\mathbf{x}'$) (a deletion or a insertion would make the two data sets identical) ϵ-differential privacy (DP) is defined as follows.

Definition 1 (ϵ-Differential privacy [1]). *A randomized mechanism $\mathcal{R}$ satisfies ϵ-differential*

privacy if for all all data sets $(\mathbf{x}, \mathbf{x}')$ that is $\Delta(\mathbf{x}, \mathbf{x}') = 1$ and all result subsets Q to query $\mathbf{q}$,
$\left| \log \left(\frac{\Pr(\mathcal{R}(q(\mathbf{x})))\in Q)}{\Pr(\mathcal{R}(q(\mathbf{x}'))\in Q)} \right) \right| \leq \epsilon$, *where $\epsilon > 0$ is the privacy budget parameter.*

Under DP, the ratio in the probabilities of obtaining the same query results from $\mathbf{x}$ and $\mathbf{x}'$ after sanitization via $\mathcal{R}$ is bounded within $(e^{-\epsilon}, e^{\epsilon})$ – a neighborhood around 1. For a small ϵ value, the ratio is close to 1, meaning that the query result is similar with or without that individual in the data set, and the chance that a participant in the data will be identified based on the query result sanitized via $\mathcal{R}$ would be low. DP provides a robust and powerful model against privacy attacks in the sense that it does not make assumptions on the background knowledge or the behavior on data intruders. ϵ can be used as a tuning parameter – the smaller ϵ is, the more protection there is on the released information via $\mathcal{R}$.

There also exist softer versions to the pure ϵ-DP that include the (ϵ, δ)-approximate differential privacy (aDP) [25], the (ϵ, δ)-probabilistic DP (pDP) [26], the (ϵ, δ) random differential privacy (rDP) [27], and the (ϵ, τ)-concentrated differential privacy (cDP) [28]. In all the relaxed versions, one additional parameter is employed to characterize the amount of relaxation on top of the privacy budget ϵ. In (ϵ, δ)-aDP, $\Pr(\mathcal{R}(q(\mathbf{x})) \in Q) \leq e^{\epsilon} \Pr(\mathcal{R}(q(\mathbf{x})) \in Q) + \delta$. A sanitization algorithm satisfies (ϵ, δ)-probabilistic differential privacy (pDP) if the probability of generating an output belonging to the disclosure set is bounded below δ, where the disclosure set contains all the possible outputs that leak information for a given privacy tolerance ϵ. The (ϵ, δ)-rDP is also a probabilistic relaxation of DP; but it differs from (ϵ, δ)-pDP in that the probabilistic relaxation is with respect to the generation of the data while it is with respect to the sanitizer in the (ϵ, δ)-pDP. The (ϵ, τ)-cDP, similar to the (ϵ, δ)-pDP, relaxes the satisfaction of DP with respect to the sanitizer, and ensures that the expected privacy cost is ϵ and (Prob(the actual cost $> \epsilon$)$> a$) is bounded by $e^{-(a/\tau)^2/2}$.

2.2   Laplace Mechanism and Global Sensitivity

The Laplace mechanism is a popular sanitizer to release statistics with ϵ-DP [1]. Liu [7] introduces the generalized Gaussian mechanism (ϵ, δ)-pDP that includes the Laplace mechanism as a special case (when $p = 1$ and $\delta = 0$). Denote the statistics of interest by $\mathbf{s}_{r \times 1}$. The Laplace and the generalized Gaussian mechanisms are based on the l_p global sensitivity, which is defined as

$$\delta_p = \max_{\substack{\mathbf{x}, \mathbf{x}' \\ \Delta(\mathbf{x}, \mathbf{x}')=1}} \|\mathbf{s}(\mathbf{x}) - \mathbf{s}(\mathbf{x}')\|_p = \max_{\substack{\mathbf{x}, \mathbf{x}' \\ \Delta(\mathbf{x}, \mathbf{x}')=1}} \left(\textstyle\sum_{j=1}^{r} |s_j(\mathbf{x}) - s_j(\mathbf{x}')|^p \right)^{1/p} \tag{1}$$

for all pairs of data sets $(\mathbf{x}, \mathbf{x}')$ that are $\Delta(\mathbf{x}, \mathbf{x}') = 1$. δ_p is the maximum difference in $\mathbf{s}$ in the l_p distance between any pair of data sets $\mathbf{x}, \mathbf{x}'$ with $\Delta(\mathbf{x}, \mathbf{x}') = 1$. The sensitivity is "global" since it is defined for all possible data sets and all possible ways of two data sets differing by one record. The larger the global sensitivity for $\mathbf{s}$, the larger the disclosure risk is from releasing the original $\mathbf{s}$, and the more perturbation is needed for $\mathbf{s}$ to offset the risk. Specifically, the Laplace mechanism of ϵ-DP sanitizes $\mathbf{s}_{r \times 1}$ as in

$$\mathbf{s}^*_{r \times 1} = (s_1, \ldots, s_r) \sim \textstyle\prod_{j=1}^{r} \mathrm{Lap}\left(s_j, \delta_1 \epsilon^{-1}\right), \tag{2}$$

where $\delta_1 = \sum_{j=1}^{r} \delta_{j1}$ is the l_1-global sensitivity of $\mathbf{s}$ and δ_{j1} is the global sensitivity of the j-th statistic in $\mathbf{s}$. For integer $p \geq 2$, the generalized Gaussian mechanism of order p sanitizes $\mathbf{s}$ with (ϵ, δ)-pDP by drawing sanitized $\mathbf{s}^*$ from the generalized Gaussian distribution

$$f(\mathbf{s}^*) = \textstyle\prod_{j=1}^{r} \frac{p}{2b\Gamma(p^{-1})} \exp\{(|s_j^* - s_j|/b)^p\} = \textstyle\prod_{j=1}^{r} \mathrm{GG}(s_j, b, p), \tag{3}$$

where b satisfies $\Pr\left(\sum_{k=1}^{r} a_j > b^p - \epsilon\delta_p^p\right) < \delta$ with $a_j = \sum_{j=1}^{p-1}\binom{p}{j}|s_j^* - s_j|^{p-i}\delta_{1,k}^j$, $\delta_{1,k} = $ the l_1-global sensitivity of s_j, and $\delta_p = $ the l_p-global sensitivity of $\mathbf{s}$. When $p = 2$, the generalized Gaussian mechanism becomes the Gaussian mechanism of (ϵ, δ)-pDP that generates sanitized s_j^* from $N(s_j, \sigma^2 = b^2/2)$ for $k = 1, \ldots, r$. The Laplace and generalized Gaussian mechanisms produce unbound sanitized results from the real line $(-\infty, \infty)$; therefore, some bounding procedures will be needed if the mechanisms are to be applied to sanitize bounded statistics.

The global sensitivity for a query or statistics in general needs to be determined analytically though its value might not be tight. Numerical computation of global sensitivity is not feasible as it is impossible to enumerate all possible data $\mathbf{x}$ and all possible ways of $\Delta(\mathbf{x}, \mathbf{x}') = 1$ especially when $\mathbf{x}$ contains continuous attributes or when sample size n is large. We have obtained the l_1 global sensitivity of some common statistics, including proportions, means, variances, and covariances (refer to the online supplementary materials). The global sensitivity values are calculated for both definitions of two data sets differing by one record and the global sensitivity is the same on most of the examined statistics regardless of which definition is used. For example, the global sensitivity of the sample mean of a variable the value of which is bounded within $[c_0, c_1]$ is $n^{-1}(c_1 - c_0)$, and that of the sample variance is $n^{-1}(c_1 - c_0)^2$. In all calculations, we assume the sample size n of the original data is a known and carries no privacy concern, which is often the case in statistical analysis. It should be noted that the global sensitivity of a function of a statistic s is not equal to the function of the global sensitivity of s in general. For example, δ_1 of a sample variance is $(c_1 - c_0)^2 n^{-1}$, but δ_1 of the sample standard deviation cannot be simply calculated as $\sqrt{(c_1 - c_0)^2 n^{-1}}$. In fact, the global sensitivity of the standard deviation is more difficult to calculate analytically compared to that of the variance. When the global sensitivity of s is not easy to calculate, but a data-independent function of s, say $t = f(s)$, is, we can instead sanitize t to obtain t^* and then obtain sanitized s^* via the back-transformation $s^* = f^{-1}(t^*)$.

3 Truncated and BIT Laplace Mechanisms

In this section, we first formalize two commonly used bounding procedures and then examine the statistical properties of sanitized outcomes processed by the twp procedures, respectively. Both procedures are intuitive and straightforward to apply and can be coupled with any differentially private mechanisms. We focus on their applications in the context of the Laplace mechanism given its popularity, and will look into their applications in other mechanisms as the generalized Gaussian mechanism in the future.

3.1 Definitions

Definition 2 (truncated and boundary-inflated-truncated Laplace mechanisms). *Denote the bounded statistics by $\mathbf{s}_{r \times 1} = (s_1, \ldots, s_r) \in [c_{10}, c_{11}] \times \cdots \times [c_{r0}, c_{r1}]$, where $[c_{j0}, c_{j1}]$ are the bounds for i^{th} element in $\mathbf{s}$ $(i = 1, \ldots, r)$, the privacy budget by ϵ, the l_1-global sensitivity of $\mathbf{s}$ by δ_1. Let $\lambda = \delta_1 \epsilon^{-1}$.*

(a) *The **truncated Laplace mechanism** of ϵ-DP sanitizes $\mathbf{s}$ by drawing $\mathbf{s}^*$ from the truncated Laplace distribution*

$$f(\mathbf{s}^*) = \prod_{j=1}^{r} Lap(s_j, \lambda | c_{j0} \le s^* \le c_{j1}) = \prod_{i=1}^{r} \frac{\exp\left(-\frac{|s_j^* - s_j|}{\lambda}\right)}{2\lambda\left(1 - \frac{1}{2}\exp(-\frac{c_{j1} - s_j}{\lambda}) - \frac{1}{2}\exp(\frac{c_{j0} - s_j}{\lambda})\right)}. \quad (4)$$

(b) *The **boundary-inflated-truncated (BIT) Laplace mechanism** of ϵ-DP sanitizes $\mathbf{s}$ by draw-*

ing $\mathbf{s}^*$ *from the BIT Laplace distribution* $f(\mathbf{s}^*) = \prod_{i=1}^{r} f(s_j^*)$, *where* $f(s_j^*)$ *is*

$$f(s_j^*) = \begin{cases} \frac{1}{2} \exp(-(s_j - c_{j0})/\lambda) & \textit{if } s_j^* = c_{j0} \\ Lap(s_j, \lambda) & \textit{if } c_{j0} < s_j^* < c_{j1} \\ 1 - \frac{1}{2} \exp(-(c_{j1} - s_j)/\lambda) & \textit{if } s_j^* = c_{j1} \end{cases} \qquad (5)$$

Rather than sampling directly from Eqn (4), the truncated Laplace mechanism can also be realized via in a post-hoc manner by throwing away out-of-bounds sanitized results from the regular Laplace mechanism until catching an in-bound value. Similarly, the BIT bounding procedure can be realized by post-hoc setting out-of-bounds sanitized results from the regular Laplace mechanism at the corresponding boundaries. If the scale parameter $\lambda \to \infty$ in the Laplace distribution when $\epsilon \to 0$ or $\delta_1 \to \infty$, it can be easily established that $f(s_j)$ in the truncated Laplace mechanism in Eq. (4) converges to an uniform distribution unif(c_{0i}, c_{1i}), and that in the BIT Laplace distribution in Eq. (5) converges to a Bernoulli distribution with probability mass at c_{0i} and c_{1i}, respectively. In both of the asymptotic cases, the sanitized results preserve little original information. Both the Laplace and BIT Laplace mechanisms, as in the regular Laplace mechanism, require calculation of the l_1-global sensitivity for the targeted $\mathbf{s}$ to be sanitized.

Remark 1 (multivariable-function constraints). *Definition 2 focuses on the case where each statistic s_j in $\mathbf{s}$ has its own lower and upper bounds $[c_{0j}, c_{1j}]$ $(j = 1, \ldots, r)$ that are publicly known and fixed constants. In some practical cases, besides the bounding constraints for each statistic, a subset of $\mathbf{s}$ or all of its elements are also required to satisfy multivariable constraints. For example, $\mathbf{s}$ is a set of proportions under the linear constraint $\sum_{j=1}^{r} s_j = 1$; or a subset $\mathcal{S}$ of $\mathbf{s}$ satisfies $c_0 \leq \sum_{k \in \mathcal{S}} a_k s_k \leq c_1$. In the presence of the additional multivariable constraints, one would first check whether the multivariable constraints are compatible with the bounding constraint on each s_j, which may lead to a set of new constraints that incorporate all the constraints – univariate or multivariate. For example, say $\mathbf{s} = (s_1, s_2)$ and the linear constraint is $s_1 + s_2 = c$. The final constraints on s_1 and s_2 would then be $[\max(c - c_{12}, c_{01}), \min(c - c_{02}, c_{11})]$ and $[\max(c - c_{11}, c_{02}), \min(c - c_{01}, c_{12})]$, respectively. After applying either the truncated or the BIT Laplace mechanism in Definition 2 to sanitize, say s_1 to obtain s_1^*, that satisfies the new bounding constraint, the other can be calculated as $s_2^* = c - s_1^*$. In some special cases, such as when $\mathbf{s}$ is a set of proportions under the linear constraint $\sum_{j=1}^{r} s_j = 1$, one can also apply a simple rescaling procedure to the sanitized $\mathbf{s}$ to satisfy the multivariable linear constraint (see simulation study 2 in Section 4 for an example).*

There are many types of multivariable constraints. Depending on the nature of the additional multivariable constraints, there may exist different approaches to satisfy all the constraints simultaneously – some might lead to better original information preservation than others for the same privacy budget. In addition, the computational complexity to obtain a set of sanitized $\mathbf{s}^*$ that satisfy all the constraints may reach a level beyond practicality. The diversity and complexity of satisfying multivariable constraints also indicates challenges for the theoretical analysis on the statistical properties of sanitized $\mathbf{s}^*$ in this setting. For these reasons, in the theoretical analysis in Section 3.2, we focus on learning the statistical properties of sanitized $\mathbf{s}^*$ where each of its element has only the univariate bounding constraint with no additional multivariable constraints. Incorporating multivariable-function constraints on top of the univariate bounding constraints deserves more in-depth investigation and will be a topic of future research.

3.2 Statistical Properties of Sanitized s^* with Univariate Bounding Constraint

In Definition 2, the bounds $[c_{10}, c_{11}] \times \cdots [c_{r0}, c_{r1}]$ are assumed to be data invariant and "global" to guarantee ϵ-DP, from the data release perspective. If the bounds are "local" and data-specific, meaning they are functions of data $\mathbf{x}$ at hand, then the bounds themselves, would need to be sanitized and would otherwise leak information about the original data. On the other hand, ignoring the local properties of data $\mathbf{x}$ in a bounding procedure could have a negative impact on the statistical properties of sanitized results $\mathbf{s}^*$. In this section, we investigate the statistical behaviors of $\mathbf{s}^*$ produced by a sanitizer with bounding constraints.

The rationale for studying the statistical properties of the sanitized statics is as follows. First, duo to randomness of the injected noise to satisfy DP, the sanitized statistic is a random variable (even without accounting for the sampling variability of the sample data after conditional on the original static). Though the privacy budget is often only enough for querying one or a few statistics, the users would still be interested in understanding and gaining theoretical insights on the statistical properties of the sanitized statistic in expectation over the distribution of the injected noise, such as bias and MSE of the sanitized statistic. Furthermore, one of the main goals for data collection is to draw inferences about the population from which the sample data come from. For non-private inferences, one would often examine the statistical properties of an estimator given sample data to see how good it is for a population parameter, such as unbiasedness and consistency. In the context of private inferences, the users won't have access to the original non-private estimate, but rather its sanitized version. If the sanitized statistic does not possess desirable statistical properties, say it is biased, for the original statistic, and even if the original statistic is unbiased for the population parameter, then users know the statistical inference based on the sanitized statistic would be biased as well, per the law of total expectation.

We start by defining the desired statistical properties of sanitized $\mathbf{s}^*$ (Definition 3), then examine the bias of $\mathbf{s}^*$ sanitized by the truncated and BIT Laplace mechanisms relative to the original $\mathbf{s}$ (Proposition 1), and present an upper bound for the mean squared error and examine the consistency of $\mathbf{s}^*$ for $\mathbf{s}$ and the convergence rate in Proposition 2. In addition, we list in Proposition 3 a set of sufficient conditions for the sanitized $\mathbf{s}^*$ to be asymptotically unbiased and consistent for the population parameter $\boldsymbol{\theta}$ in the case where the original $\mathbf{s}$ is an estimator for $\boldsymbol{\theta}$.

Definition 3 (unbiasedness and consistency of sanitized statistic for original statistic).
Sanitized $\mathbf{s}^*$ is unbiased for the original $\mathbf{s}$ if $E_{\mathbf{s}^*}(\mathbf{s}^*|\mathbf{s}) = \mathbf{s}$; $\mathbf{s}^*$ is asymptotically unbiased for $\mathbf{s}$ if $E_{\mathbf{s}^*}(\mathbf{s}^*|\mathbf{s}) \to \mathbf{s}$ as $n \to \infty$, where n is the sample size of original data $\mathbf{x}$; $\mathbf{s}^*$ is consistent for $\mathbf{s}$ if $\mathbf{s}^* \xrightarrow{p} \mathbf{s}$ as $n \to \infty$.

When $\mathbf{s}$ is boundless and sanitized by the regular Laplace mechanism, then $\mathbf{s}^*$ is unbiased for $\mathbf{s}$ since $E(\mathbf{s}^*) = \mathbf{s}$ per the definition of the Laplace distribution. If $\delta_1 \propto n^{-k}$, where $k > 0$, then $\mathbf{s}^*$ is also consistent for $\mathbf{s}$. When $\mathbf{s}$ is bounded and sanitized via the truncated or BIT Laplace mechanism, we will have biased $\mathbf{s}^*$ unless the bounds are symmetric around the original $\mathbf{s}$. Proposition 1 states the formal conclusions and presents the magnitude of the bias of s^* in the truncated and BIT Laplace mechanisms. The proof of Proposition 1 is provided in Appendix A.

Proposition 1 (bias of sanitized statistic from truncated and BIT Laplace mechanisms).
Let $[c_0, c_1]$ be the global bounds on a singular s, λ be the scale parameter, s be the location parameter

of the Laplace distribution, μ_1 be the expected mean of the truncated Laplace distribution $f(s^*|s^* \in [c_0, c_1])$, and μ_2 be the expected mean of the BIT Laplace distribution defined in Eqn (5). Then

$$\mu_1 = s + \frac{\frac{\lambda-c_0+s}{2}\exp\left(\frac{c_0-s}{\lambda}\right) - \frac{\lambda+c_1-s}{2}\exp\left(\frac{s-c_1}{\lambda}\right)}{1 - \frac{1}{2}\exp(\frac{c_0-s}{\lambda}) - \frac{1}{2}\exp(\frac{s-c_1}{\lambda})} \tag{6}$$

$$\mu_2 = s + \frac{\lambda}{2}\left[\exp\left(\frac{c_0-s}{\lambda}\right) - \exp\left(\frac{s-c_1}{\lambda}\right)\right] \tag{7}$$

(a) $\mu_1 = \mu_2 = s$ (s^* is unbiased for s) if and only if $c_0 + c_1 = 2s$ (c_0 and c_1 are symmetric around s).
(b) $(\mu_1 - s)(\mu_2 - s) \geq 0$ (the direction of the bias is the same between the two). Specifically, $\mu_1 - s \geq 0$ and $\mu_2 - s \geq 0$ if $s - c_0 \leq c_1 - s$; $\mu_1 - s \leq 0$ and $\mu_2 - s \leq 0$ if $s - c_0 \geq c_1 - s$.
(c) $|\mu_1 - s| \geq |\mu_2 - s|$ (s^* sanitized via the BIT Laplace sanitizer is no more biased than that via the truncated Laplace sanitizer).

For the global and data invariant bounds $[c_0, c_1]$, it is unlikely the sanitized results would be unbiased via the truncated or the BIT Laplace mechanism per part (a) of Proposition 1 as $[c_0, c_1]$ are fixed while the original statistic s changes from data to data. To achieve unbiasedness for s^*, local bounds that depend on specific data sets can be constructed at additional privacy cost. For example, bounds $[s - \min(s - c_0, c_1 - s), s + \min(s - c_0, c_1 - s)]$, which are symmetric around s, can be used to bound sanitized results in the truncated and BIT Laplace mechanism that leads to unbiased s^*. However, since the bounds are functions of the original s, they will leak information about s, which has to be counted for in the total privacy cost.

Remark 2 (Extension of Proposition 1 to multidimensional s). *Proposition 1 examines a scalar query s. The conclusions can be easily extended to $\mathbf{s}_r = (s_1, \ldots, s_r)$ of arbitrary dimension $r \geq 1$ with univariate bounding constraint per statistic. The only difference lies in how the scale parameters λ is defined. As given in Eqn (2), the Laplace mechanism for sanitizing $\mathbf{s}_{r \times 1}$ is $\prod_{j=1}^{r} Lap(s_j, \delta_1 \epsilon^{-1})$ with $\delta_1 = \sum_{j=1}^{r} \delta_{j1}$, where δ_{j1} is the global sensitivity of s_j, the j-th statistic in $\mathbf{s}$. We can rewrite The joint distribution of sanitized $\mathbf{s}^*$ as*

$$\prod_{j=1}^{r} Lap\left(s_j, \delta_{1j}\left(\epsilon\frac{\delta_{1j}}{\delta_1}\right)^{-1}\right) = \prod_{j=1}^{r} Lap(s_j, \delta_{1j}(w_j\epsilon)^{-1}), \text{ where } w_j = \delta_{1j}/\delta \tag{8}$$

In other words, the Laplace mechanism can be regarded as a special application of the sequential composition [5] by allocating a portion $w_j = \delta_{1j}/\delta$ of total ϵ to the sanitization of s_j (also refer to Liu [29]). For the scenario examined in this paper, that is, the bounds for $\mathbf{s}$ are publicly known and fixed, the conclusions from Proposition 1 apply to each element s_j in $\mathbf{s}_{r \times 1}$ with different bounds $[c_{0j}, c_{1j}]$ and scale parameter $\lambda_j = \delta_j(w_j\epsilon)^{-1}$ for different s_j in Eqns (6) and (7).

Remark 3 (Proposition 1 for one-sided bounding constraints). *The conclusions in Proposition 1 are also applicable in the special cases when the bounds are one-sided $[c_0, \infty]$ or $[\infty, c_1]$, if such needs exist in practice. The proof and establishment of these conclusions do not impose any restrictions on the values c_0 and c_1 except for the trivial requirement $c_0 < c_1$. The bias terms actually*

have simpler expression in the case. Specifically, Eqns (6) and (7) become

$$\text{bounds } [c_0, \infty) : \mu_1 = s + \frac{\frac{\lambda - c_0 + s}{2} \exp\left(\frac{c_0 - s}{\lambda}\right)}{1 - \frac{1}{2}\exp\left(\frac{c_0 - s}{\lambda}\right)}; \quad \mu_2 = s + \frac{\lambda}{2}\exp\left(\frac{c_0 - s}{\lambda}\right) \tag{9}$$

$$\text{bounds } (-\infty, c_1] : \mu_1 = s - \frac{\frac{\lambda + c_1 - s}{2}\exp\left(\frac{s - c_1}{\lambda}\right)}{1 - \frac{1}{2}\exp\left(\frac{s - c_1}{\lambda}\right)}; \quad \mu_2 = s - \frac{\lambda}{2}\exp\left(\frac{s - c_1}{\lambda}\right). \tag{10}$$

It is easier (compared to the general case of $[c_0, c_1]$) to see that both biases from the truncation and the BIT procedures are positive when the bounds are $[c_0, \infty)$, a special case of $s - c_0 < c_1 - s$ in part b), and negative when the bounds are $(-\infty, c_1]$ a special case of $s - c_0 > c_1 - s$. The difference of the two bias terms when the bounds are $[c_0, \infty]$ is $\left[\frac{1 - (c_0 - s)/\lambda}{1 - \exp(\frac{c_0 - s}{\lambda})/2} - 1\right]\frac{\lambda}{2}\exp\left(\frac{c_0 - s}{\lambda}\right)$. Since $0 < 1 - \exp(\frac{c_0 - s}{\lambda})/2 < 1/2$ and $1 - (c_0 - s)/\lambda > 1$, thus the bias difference is > 0. Since both biases are positive, this implies that μ_1 has a larger positive bias than μ_2. Similarly, when the bounds are $[-\infty, c_1]$, the bias difference between μ_1 and μ_2 is $\left[1 - \frac{1 + (c_1 - s)/\lambda}{1 - \exp(\frac{s - c_1}{\lambda})/2}\right]\frac{\lambda}{2}\exp\left(\frac{s - c_1}{\lambda}\right)$. Since $0 < 1 - \exp(\frac{c_1 - s}{\lambda})/2 < 1/2$, and $1 + (c_1 - s)/\lambda > 1$, the bias difference is < 0. Since both biases are negative, this implies that μ_1 has a larger negative bias than μ_2. Taken together, $|\mu_1 - s| > |\mu_2 - s|$.

In addition to quantify the magnitude of bias, it is also of interest to bound the error of a sanitized statistic relative to its original value. Proposition 2 examines the upper bound for the mean squared error (MSE) for a sanitized statistic via the truncated and BIT Laplace mechanisms, and establishes the condition under which the statistic would be MSE-consistent for its original value, and the convergence rate of consistency.

Proposition 2 (consistency and convergence rate). *Let s be the location parameter and λ be the scale parameter of the Laplace distribution, and s^* be the sanitized result for s via the truncated or the BIT Laplace mechanism. The MSE $E_{s^*}(s^* - s)^2$ is upper bounded by $2\lambda^2 = 2(\delta_1/\epsilon)^2$ and s^* is MSE-consistent for s as $\lambda \to 0$. If $\delta_1 \propto n^{-k}$, where $k > 0$ and n is the sample size, then the rate of the MSE converging to 0 is $O(n^{-2k})$ for a given ϵ.*

The proof is provided in Appendix B. The proof does not have restrictions on the values of c_0 or c_1, which can be as extreme as $c_1 = \infty$ and $c_0 = -\infty$. Therefore, the conclusions in Proposition 2 also to the non-sided bounding constraints. In addition, similar to Remark 2 for Proposition 1, the established conclusions in Proposition 2 can be easily extended to $\mathbf{s}_r = (s_1, \ldots, s_r)$ of arbitrary dimension $r \geq 1$ with univariate bounding constraint. The only difference is that the statistic-specific λ_j $(j = 1, \ldots, r$ would be used. Therefore, different elements from $\mathbf{s}_r$ will have different MSE bounds and different convergence rates.

Proposition 2 shows that the sanitized result, though likely biased per Proposition 1, can still enjoy nice asymptotic properties such as asymptotic unbiasedness and consistency as $\lambda \to 0$. In the framework of the truncated and BIT Laplace mechanisms, the scale parameter of the associated Laplace distribution λ is $\delta_1\epsilon^{-1}$. For a pre-specified ϵ, to satisfy the condition $\lambda \to 0$, then δ_1 needs to be $\to 0$. Intuitively speaking, as n increases, the influence of a single individual on an aggregate measure of a data set is likely to diminish, and the individual is less prone to be identified from releasing the aggregate measure. Translated to the global sensitivity of the aggregate measure, it means that δ_1 decreases with n. δ_1 of some commonly used statistics are $\propto n^{-1}$, such as proportions, means, variances and covariances (refer to the online supplemental materials), and the sanitized copies of

these statistics via either the truncated or BIT Laplace mechanism are consistent for their original values by part c) of Proposition 1. The results in Proposition 2 imply that the MSE of a sanitized statistic from the truncated or the BIT Laplace mechanism is comparable to the MSE from the regular Laplace mechanism without bounding, which is also $2\lambda^2$. In other words, the bounding does not seem to affect the MSE of a sanitized statistic despite the loss of unbiasedness in general.

Propositions 1 and 2 examine the statistical properties of the sanitized s^* relative to the original s. In many statistical analysis, the ultimate goal is to infer unknown population parameters θ based on the sample data. Suppose s is an estimator for parameter θ in a statistical model. Proposition 3 lists the sufficient conditions for s^*, the sanitized version of s, to be asymptotically unbiased and consistent for $\boldsymbol{\theta}$.

Proposition 3 (asymptotic unbiasedness and consistency of sanitized statistic for population parameter).
(a) If $E(\mathbf{s}|\boldsymbol{\theta}) = \boldsymbol{\theta}$ or $E(\mathbf{s}|\boldsymbol{\theta}) \rightarrow \boldsymbol{\theta}$, and if $E_{\mathbf{s}^*}(\mathbf{s}^*|\mathbf{s}) \rightarrow \mathbf{s}$ as $n \rightarrow \infty$, then $\mathbf{s}^*$ is asymptotically unbiased for θ; that is, $E(\mathbf{s}^*|\boldsymbol{\theta}) \rightarrow \boldsymbol{\theta}$.
(b) If $\mathbf{s}^* \xrightarrow{p} \mathbf{s}$ and $\mathbf{s} \xrightarrow{p} \boldsymbol{\theta}$ as $n \rightarrow \infty$, then $\mathbf{s}^*$ is consistent for θ; that is, $\mathbf{s}^* \xrightarrow{p} \boldsymbol{\theta}$.

The proof of Proposition 3 is given in Appendix C. Note that Proposition 3 does not list the conditions for obtaining an unbiasedness s^* for $\boldsymbol{\theta}$ as it is meaningless given the low likelihood of obtaining an unbiased s^* for s per Proposition 1. Proposition 3 implies that asymptotic unbiasedness and consistency of the sanitized s^* for $\boldsymbol{\theta}$ can be achieved in two steps. In step one, we will choose an estimator s that is asymptotically unbiased or consistent for $\boldsymbol{\theta}$, which should be relatively straightforward given that these types of estimators are well studied and widely applied in statistics; in the second step, we will employ an appropriate differentially private mechanism to generate s^* that is asymptotically unbiased or consistent for s, such as the BIT and truncated Laplace mechanisms, which yield MSE-consistent s^* for s under the conditions listed in Propositions 2.

4   Simulation Studies

We conducted two simulation studies to demonstrate the applications of the truncated and BIT bounding mechanisms and examine the statistical properties of the sanitized results. In the first simulation, we sanitized a variance-covariance matrix, and focused on the comparisons between the sanitized results and the original results and between the truncated and BIT truncated Laplace mechanisms on the effects on the sanitized results. In the second simulation, we sanitized proportions and focused on the inferential properties of the sanitized proportions by examining the bias, root mean squared error (RMSE) and coverage probability (CP) for the true proportions based on the sanitized results.

4.1   Simulation Study 1

In this simulation, we applied the truncated and BIT Laplace mechanisms to sanitize a variance-covariance matrix $\mathbf{S}$ in a data set of size n. The variance-covariance matrix is a widely used statistic for examining the dependency structure among multiple continuous variables. It is also an ideal statistic to examine the effects of the two mechanisms given that every element in the matrix has to satisfy some type of bounding constraints. The bounding constraints of a covariance matrix of any dimension include that the marginal variances are positive and the correlations are bounded between [-1, 1]. Additionally, the marginal variances are right-bounded in bounded data from which $\mathbf{S}$ is calculated. Table 1 summarizes the bounds and global sensitivity of the components in $\mathbf{S}$.

Table 1: Global sensitivity of variance and covariance terms in a covariance matrix

statistic		bounds[‡]	δ_1
variance	S_{jj}	$(0, n(c_{j1} - c_{j0})^2/(4(n-1))]$	$(c_{j1} - c_{j0})^2/n$
variance	$S_{j'j'}$	$(0, n(c_{j'1} - c_{j'0})^2/(4(n-1))]$	$(c_{j'1} - c_{j'0})^2/n$
covariance	$S_{jj'}$	$(-\sqrt{S_{jj}S_{j'j'}}, \sqrt{S_{jj}S_{j'j'}})$	$(c_{j1} - c_{j0})(c_{j'1} - c_{j'0})/n$

[‡] $[c_{j0}, c_{j1}] \times [c_{j'0}, c_{j'1}]$ were the bounds of variables X_j and $X_{j'}$

When sanitizing $\mathbf{S}$ in general, we first obtain legitimate sanitized S_{jj}^* and $S_{j'j'}^*$, and then sanitize $S_{jj'}$ given S_{jj}^* and $S_{j'j'}^*$ under the constraint that $S_{jj'}^* \in [-\sqrt{S_{jj}^* S_{j'j'}^*}, \sqrt{S_{jj}^* S_{j'j'}^*}]$. Though the bounds for $S_{jj'}^*$ depend on S_{jj}^* and $S_{j'j'}^*$, the latter two are already sanitized; therefore, bounding procedures for $S_{jj'}^*$ using information S_{jj}^* and $S_{j'j'}^*$ does not incur additional privacy cost. The interval constraint on $S_{jj'}^*$ does not have to contain the original $S_{jj'}$. Comparing to containing $S_{jj'}$, which would be nice from information preservation perspective, satisfying the hardcore requirement that the Pearson correlation is within $[-1, 1]$ is necessary, when there is a conflict between the two. In addition, requiring the interval to contain the original $S_{jj'}$ would also lead to privacy leak, strictly speaking.

It is possible that the sanitized covariance matrix $\mathbf{S}^*$ is not positive definite (PD) with the element-wise sanitization approach. If a sanitized covariance matrix is not PD and has a significant number of (small) negative eigenvalues, it can be made PD with semi-definite optimization via, e.g., the alternating projections algorithm [30], the Newton methods for nearest correlation matrix [31, 32], and the spectral projected gradient method [31, 33][1]. A possible alternative is to sanitize $\mathbf{S}$ as a whole instead of element-wise, an interesting and worthwhile topic for future research.

We examined in this simulation study a 2×2 variance-covariance matrix with three different specifications of (S_{11}, S_{22}, r): $(1, 2, 0), (1, 2, -0.4)$, and $(1, 2, 0.7)$, respectively. The 3 correlation settings allow us to examine the bounding effects on the pairwise correlation when there is no correlation, moderate (negative) correlation, and strong (positive) correlation. We set the global bounds $[c_{10}, c_{11}]$ at $[-3, 3]$ and $[c_{20}, c_{21}]$ at $[-4.5, 4.5]$ [2]. The total privacy budget was $\epsilon = 1$. Since the 3 statistics were calculated on the same set of data, the sequential composition principle applied [5] when it comes to privacy budgeting. There are different ways to allocate the total budget when sanitizing multiple statistics calculated from the same data set, such as by an equal allocation across all the statistics or according to some type statistical or practical "importance" of the statistics (see Liu [29] for more discussion). Here we divided the total privacy budget ϵ equally among the 3 statistics; that is, each sanitization received $1/3$ of the total budget. We also investigated a wide range of sample size n from 50 to 800. In each examined scenarios of $(\mathbf{S}, n)$, 500 independent sanitizations were carried out to examine the distributional properties of the sanitized results.

The results are presented in Figure 1. In each plot, the average and the (2.5%, 25%, 75% and 97.5%) percentiles of the sanitized results from the 500 sanitizations are presented, benchmarked against the original results. The main findings are summarized as follows. First, when n was relatively small, there was noticeable bias in the sanitized results compared to the original results, except for S_{12} and r when $r = 0$ (the boundaries were sym-

---

[1] The R function `nearPD()` in package `Matrix` implements the alternating projections algorithm.

[2] For approximately Gaussian variables with means 0, both the bounds of $[-3, 3]$ with a standard deviation of 1 and $[-4.5, 4.5]$ with a standard deviation of $\sqrt{2}$ represent $> 99\%$ data mass though this simulation did not require the Gaussian assumption.

metric about the original results and thus there was no bias per part (a) of Proposition 1). Second, the sanitized results generated via the truncated Laplace mechanism were more biased than those via the BIT Laplace mechanism, consistent with part (b) of Proposition 1. Third, as n increased, both the bias and dispersion of the sanitized results approached 0, consistent with part (c) of Proposition 1. Lastly, when n was small, the scale parameter associated with the Laplace distribution in both mechanisms was large, therefore more sanitized values were set at the boundary values in the BIT mechanism and the distribution of the sanitized values became flatter in the truncated mechanism, especially for r.
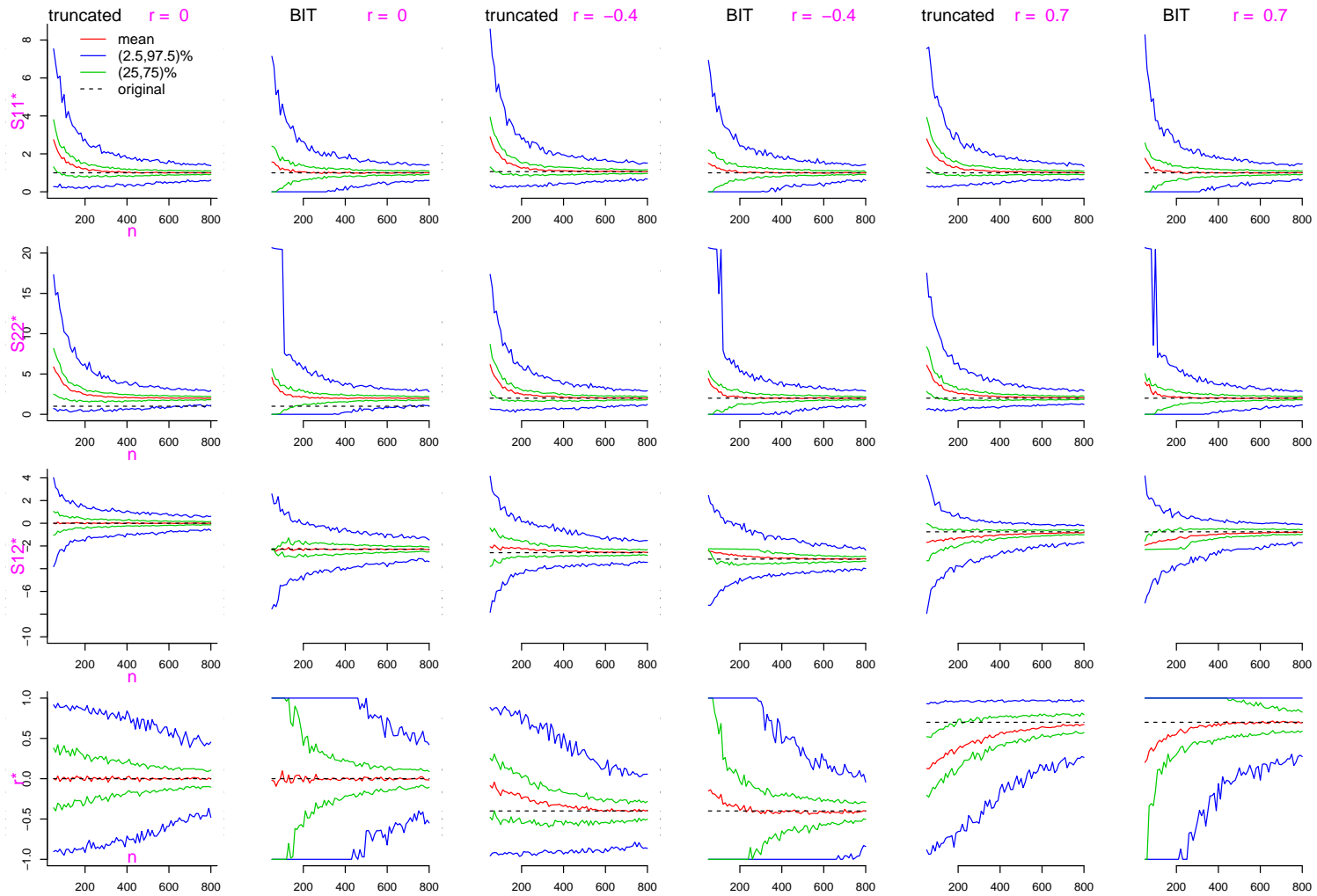
Figure 1: Sanitized components in $\mathbf{S}$ ((Q_1, Q_2)% represents the Q_1 and Q_2 percentiles of the distribution of the sanitized statistic)

4.2   Simulation Study 2

In this simulation, we aimed to release a proportion vector $\mathbf{p}$ where $\sum_{j=1}^{K} p_j = 1$. Proportions are very common statistics in public data release. For example, $\mathbf{p}$ could be the proportions of different income levels in the US population, or the cell proportions from a cross-tabulation of, say gender and race. Besides the bounding constraint $[0,1]$ on each proportion component, $\mathbf{p}$ is also subject to the equality constraint $\sum_{j=1} p_j = 1$, which has to be retained in the released sanitized results, making it an interesting problem to study.

Since the proportions are calculated from disjoint subsets of data, the addition or removal of a single observation affects the count in exactly one cell, the global sensitivity (δ_1) of releasing $\mathbf{p}$ is n^{-1} or $2n^{-1}$, depending on which definition of "differing by one record" is used on $\Delta(x, x') = 1$ (refer to the online supplementary materials). With the Laplace mechanism, sanitization of each proportion in $\mathbf{p}$ is perturbed with a noise term from $\text{Lap}(\delta_1 \epsilon^{-1})$. In this simulation, we used $\delta_1 = n^{-1}$, and examined 3 different specifications of ϵ $(0.1, 0.5, 1)$ and a range of sample size ($n = 50 \sim 500$) (the results obtained from this simulation are also applicable to $\delta_1 = 2n^{-1}$ with doubled ϵ). 500 multinomial data sets, each sized at n, were simulated from multinomial$(n, \mathbf{p})$. We examined a 4-element $\mathbf{p}$ in this simulation and set $\mathbf{p} = (0.1, 0.2, 0.3, 0.4)$ (these parameter values were chosen because they expand a wide range – some are closer to the boundaries while others are closer to the center, allowing us to examine the effects of the binding mechanisms on proportions of different magnitude).

The sample proportions $\hat{\mathbf{p}}$ were calculated in each repeat and were sanitized via the truncated and BIT Laplace mechanisms respectively. We employed 3 procedures to ensure the equality constraint $\sum_{k=1}^{4} p_k = 1$ was met (in addition to the bounding constrained on each proportion). In the first approach (re-scaling and normalization), each proportion in $\hat{p}_j$ was sanitized independently to obtain the perturb version $\hat{q}_k^*$, which was then normalized to obtain $\hat{p}_j^* = \hat{q}_k^* (\sum_{k=1}^{4} \hat{q}_k^*)^{-1}$; it is $\hat{\mathbf{p}}^*$ that was released. This re-scaling approach was intuitive and straightforward. In the second approach, referred to as the all-but-one approach, we sanitized 3 proportions out of 4, and then calculated the 4-th proportion via $1 - \sum_{k=1}^{3} \hat{q}_k^*$. The all-but-one approach obeyed the equality constraint during the sanitization without post-processing as in the re-scaling approach. In the third approach, referred to as the universal histogram/UH approach, we applied a slightly modified procedure from Hay et al. [22] to ensure the equality constraint in $\mathbf{p}$. In the original UH procedure in Hay et al. [22], the root node is sanitized and there is no inequality constraint; in our case, the root node was fixed at 1 as it was the summation of $\mathbf{p}$, and there were inequality constraints. Since the UH requires the number of children per node to be constant across the whole tree and there needs to be at least three layers in the tree (since the root node is fixed at one in our case) in order to show any level of improvement in the accuracy of some of the sanitized nodes, the four proportions in $\mathbf{p}$, represented by four leaf nodes, are thus combined in a binary fashion in 3 layers; otherwise, it would be the same as the re-scaling approach.

1. Arrange the 4 proportions in a 3-layer binary tree structure. The root node in the tree always has a value of 1, and its two child nodes h_1 and h_2 in layer 2 satisfy the constraint $h_1 + h_2 = 1$ (equality constraint 1). Similarly, the two child nodes of h_1 (h_{11} and h_{12} in layer 3) satisfy $h_{11} + h_{12} = h_1$ (equality constraint 2) and the two child nodes of h_2 (h_{21} and h_{22} in layer 3) satisfy $h_{21} + h_{22} = h_2$ (equality constraint 3). As a result, the four leaf nodes in layer 3, corresponding to the 4 proportions in $\mathbf{p}$, satisfy $h_{11} + h_{12} + h_{21} + h_{22} = 1$.

2. Sanitize h_1 and h_2 in layer 2, and $(h_{11}, h_{12}, h_{21}, h_{22})$ in layer 3 via the BIT and Laplace Laplace mechanisms with the scale parameter $2/(n\epsilon)$. The global sensitivity doubles in this procedure as there are two sets of proportions to be sanitized on the same data set.

3. Calculate the inconsistent counts $\mathbf{z}$ (due to sanitization) for the nodes in the tree. Let $\mathbf{h}^*(v) = \{h_1^*, h_2^*, h_{11}^*, h_{12}^*, h_{21}^*, h_{22}^*\}$ denote the sanitized results from step 2. $z(v) = h^*(v)$ if v is a leaf node in layer 3, and $z(v) = \frac{1}{3}\sum_{v' \in \text{children of } v} z(v') + \frac{2}{3}h^*(v)$ for v in layer 2.

4. Correct the inconsistency in the tree. For nodes in layers 2 and 3, the corrected nodes $\bar{h}(v) = z(v) + \frac{1}{2}(h(u) - \sum_{v'} z(v'))$, where u is the parent of v, and $\bar{h}(v) = 1$ for the root node[3].

5. Release $\bar{\mathbf{h}}(v)$.

Hay et al. [22] stated that the UH procedure optimizes the accuracy of the sanitized nodes closer to the root (low-order marginals) with the smallest MSE (relative to the original results) among the approaches that yield unbiased sanitized estimators for the original results while satisfying the equality/consistency constraint. However, the UH procedure decreases the accuracy of the sanitized high-order nodes and the leaf nodes. This implies the accuracy of the 4 individual proportions in $\mathbf{p}$, which were the leaf nodes in the simulation, will suffer, but some linear combinations of $\mathbf{p}$, say $p_1 + p_2$ might have higher accuracy than from the re-scaling and the all-but-one approaches. In addition, the UH procedure can be sensitive to the order of how the tree is built. We tried two different ways of grouping the 4 proportions into two nodes (h_1 and h_2) in the second layer. One way $((h_{11} = p_1 = 0.1, h_{12} = p_2 = 0.2) \in h_1; (h_{21} = p_3 = 0.3, h_{22} = p_4 = 0.4) \in h_2)$ seemed slightly better in preserving the original information than the other $((h_{11} = p_1, h_{12} = p_4) \in h_1; (h_{21} = p_2, h_{22} = p_3) \in h_2)$; therefore, we only present the results from the former.

We calculated the bias and RMSE relative to the true $\mathbf{p}$, and the coverage probability (CP) of the 95% confidence interval for the true $\mathbf{p}$ based on the sanitized $\hat{\mathbf{p}}^*$ in each of the 3 approaches. The bias, RMSE, and CP based on the sanitized results were compared to those based on the original $\mathbf{p}$. As discussed above, some of the lower-order nodes (such as h_1 and h_2) might have improved accuracy in the UH approach, we thus also compared the inferences of $p_1 + p_2$, $p_2 + p_3$ and $p_2 + p_4$ between the re-scaling and the UH approach. There is no need to examine the sum of 3 proportions, say $p_1 + p_2 + p_3$, as its accuracy would be the same as the individual proportion p_4 under the equality constraint.

In the re-scaling approach, there was minimal bias in sanitized $\hat{\mathbf{p}}^*$ when $\epsilon = 1$ and $\epsilon = 0.5$ across all n and both bounding mechanisms (truncation or BIT); and there was some bias at small n when $\epsilon = 0.1$, especially for the smallest proportion $p_1 = 0.1$ (positive bias) and the largest proportion $p_4 = 0.4$ (negative bias). Consistent with Proposition 1, the BIT mechanism yielded less bias than the truncated mechanism. The RMSE of the sanitized results was inflated compared to the original RMSE, which was expected given the noise injected during the sanitization step. The larger ϵ or the larger n was, the smaller the inflation was. When $\epsilon = 1$, the sanitized and the original RMSE values were basically the same. Though the BIT mechanism led to smaller bias compared to the truncated mechanism for small n when $\epsilon = 0.1$, the RMSE values were larger in the former. Finally, the CP was around 95% at all n when $\epsilon = 1$, decreased to $85\% \sim 92\%$ for $n \in [50, 300]$ when $\epsilon = 0.5$, and down to $50\% \sim 80\%$ for all $n \in [50, 500]$ when $\epsilon = 0.1$. The BIT mechanism had worse under-coverage than the truncated mechanism at small n for $\epsilon = 0.1$ and yielded similar CP as the truncated mechanism in other cases.

---

[3]Even if the BIT or the truncated Laplace mechanisms is employed in Step 2, negative nodes might re-appear after the correction in Step 4. If that occurs, the negative nodes are set at 0 and the nodes that share the same parent node with the negative nodes are re-normalized to sum up to their parent node. We also tried applying the regular Laplace mechanism directly in Step 2 and used the BIT or truncated Laplace mechanisms to bound the nodes after Step 4, and the results were similar or slightly worse compared to the above procedure, depending on which statistics to release.

The under-coverage observed can be resolved to some degree by using the multiple synthesis (MS) technique in DP [29, 34]. The MS takes into the variability introduced by the sanitization process by releasing multiple synthetic sets. In this case, we independently sanitized 5 sets in each simulation scenario; and the inferences were then combined over the 5 sets using the rule given in Liu [29]. In order to maintain ϵ-DP overall, each set was sanitized using $1/5$ of the total privacy budget ϵ per the sequential composition theorem. The results are given in Figure 3. The CP improved significantly from releasing multiple synthetic data sets, especially for the BIT mechanism. However, due to the decreased privacy budget per synthetic set and the bounding, the sanitized results were noisier and the biases were noticeably larger, even after being averaged the 5 sets. For example, there was noticeable bias at $\epsilon = 0.1$ and the RMSE did not approach the original RMSE within the examined range of n for both mechanisms.

Compared to the re-scaling approach, 1) the all-but-one approach (Fig 4) had similar performance in bias and RMSE for $\epsilon = 0.5$ and 1; 2) when $\epsilon = 0.1$, the bias was smaller at small n but the RMSE was noticeably large in the all-but-one approach;3) the performance on CP was similar in all proportions but was much worse for $p_1 = 0.1$, which was the proportion calculated from the other 3. Compared to the re-scaling approach, the UH approach (Fig 5) was worse in bias, RMSE and CP for all proportions at all ϵ values when n was relatively small; the under-performance was the most obvious when $\epsilon = 0.1$. The inferiority of the UH relative to the re-scaling approach are expected as the UH benefits the accuracy of the low-order marginals (the nodes closer the root) but the leaf nodes (the individual proportions in $\mathbf{p}$) actually suffer a loss of accuracy.

Figures S1 and S2 in the supplementary materials display how much improvement the UH approach brought to the pairwise sums in $\mathbf{p}$ compared to the re-scaling approach. The sanitized inferences on $p_1 + p_2$ (node h_1) were better than $p_2 + p_3$ and $p_2 + p_4$ (sum of two leaf nodes from different parents) in the UH approach. However, compared to the re-scaling approach, the accuracy of sanitized inferences on $p_1 + p_2$ did not appear to be better; and those on $p_2 + p_3$ and $p_2 + p_4$ were worse. In summary, the UH approach was not efficient as the re-scaling approach in preserving the original information in the individual proportions in $\mathbf{p}$, and the supposed improvement in the low-level nodes ($p_1 + p_2$) was not obvious in this case either, perhaps due to 1) the additional inequality constraint, which is not considered in Hay et al. [22], on top of the equality constraint; and 2) the number of layers and nodes was not large enough to demonstrate the advantages of the UH approach.
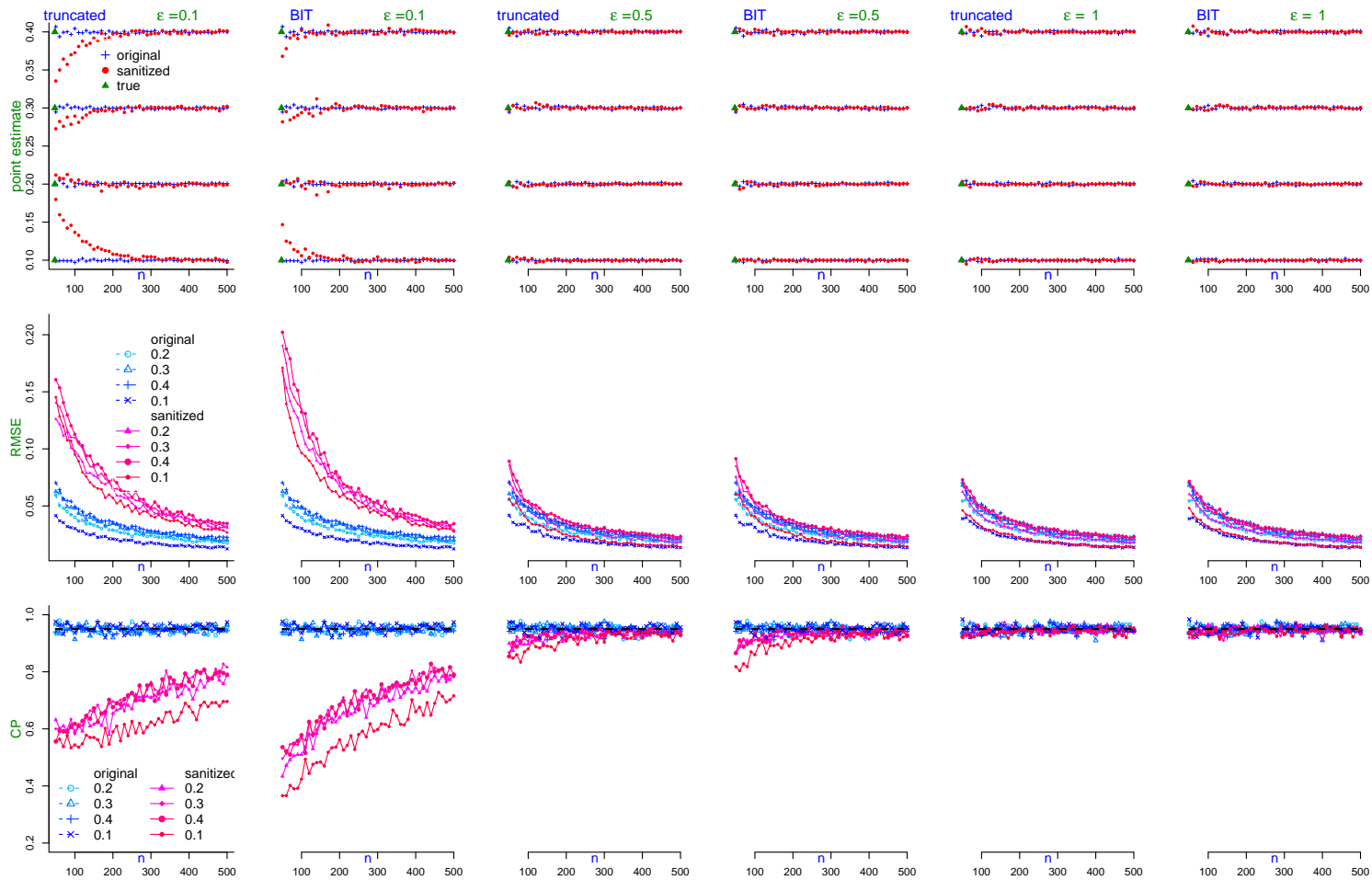
Figure 2: Bias, RMSE and CP of sanitized proportions (the re-scaling approach was applied to satisfy the equality constraint; red lines represent the 4 original proportions, and blue lines represent the 4 sanitized proportions)
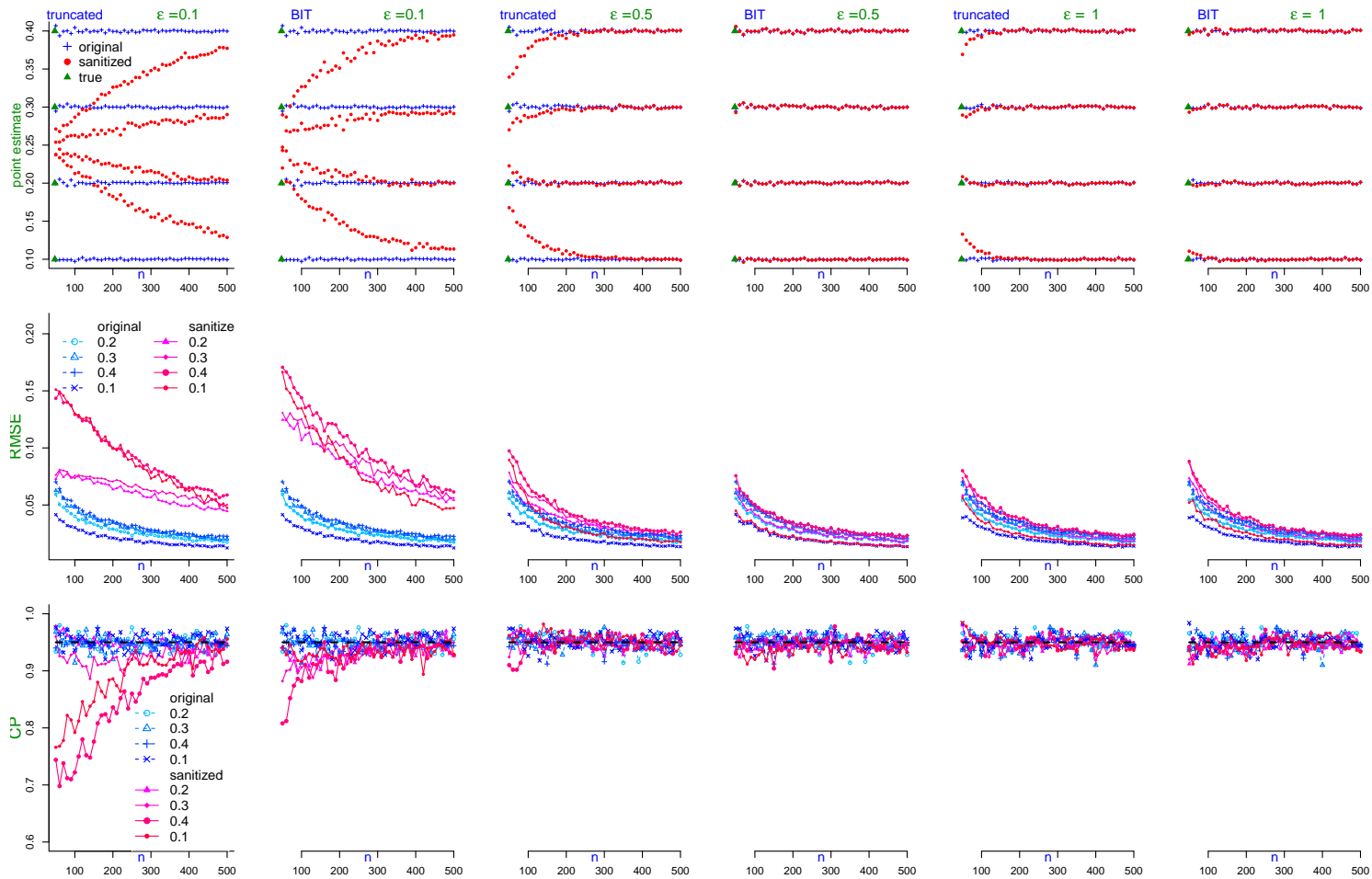
Figure 3: Bias, RMSE and CP of sanitized proportions from multiple synthesis (the re-scaling approach was applied to satisfy the equality constraint; red lines represent the 4 original proportions, and blue lines represent the 4 sanitized proportions)
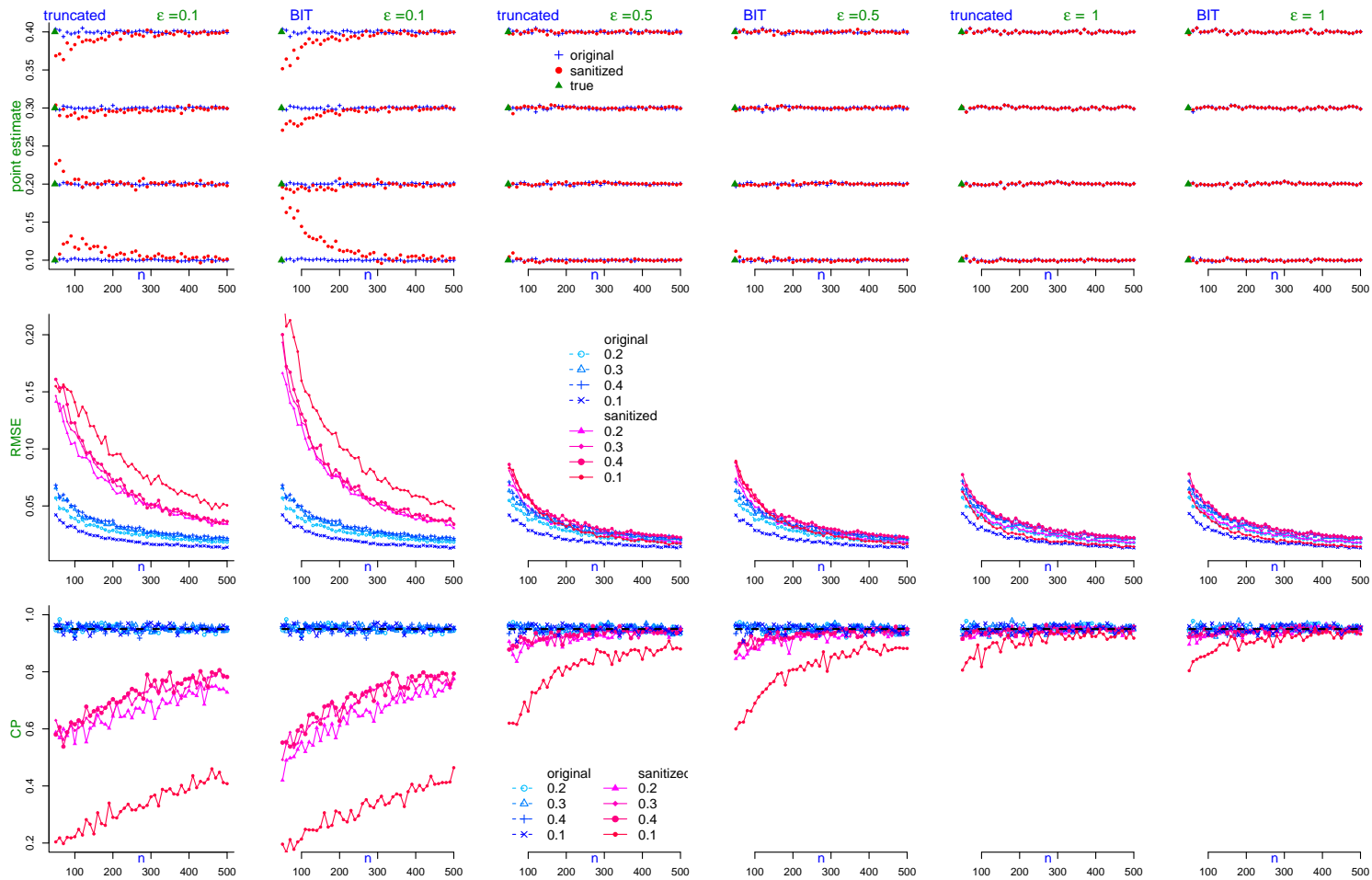
Figure 4: Bias, RMSE and CP of sanitized proportions in the all-but-one approach (red lines represent the 4 original proportions, and blue lines represent the 4 sanitized proportions)
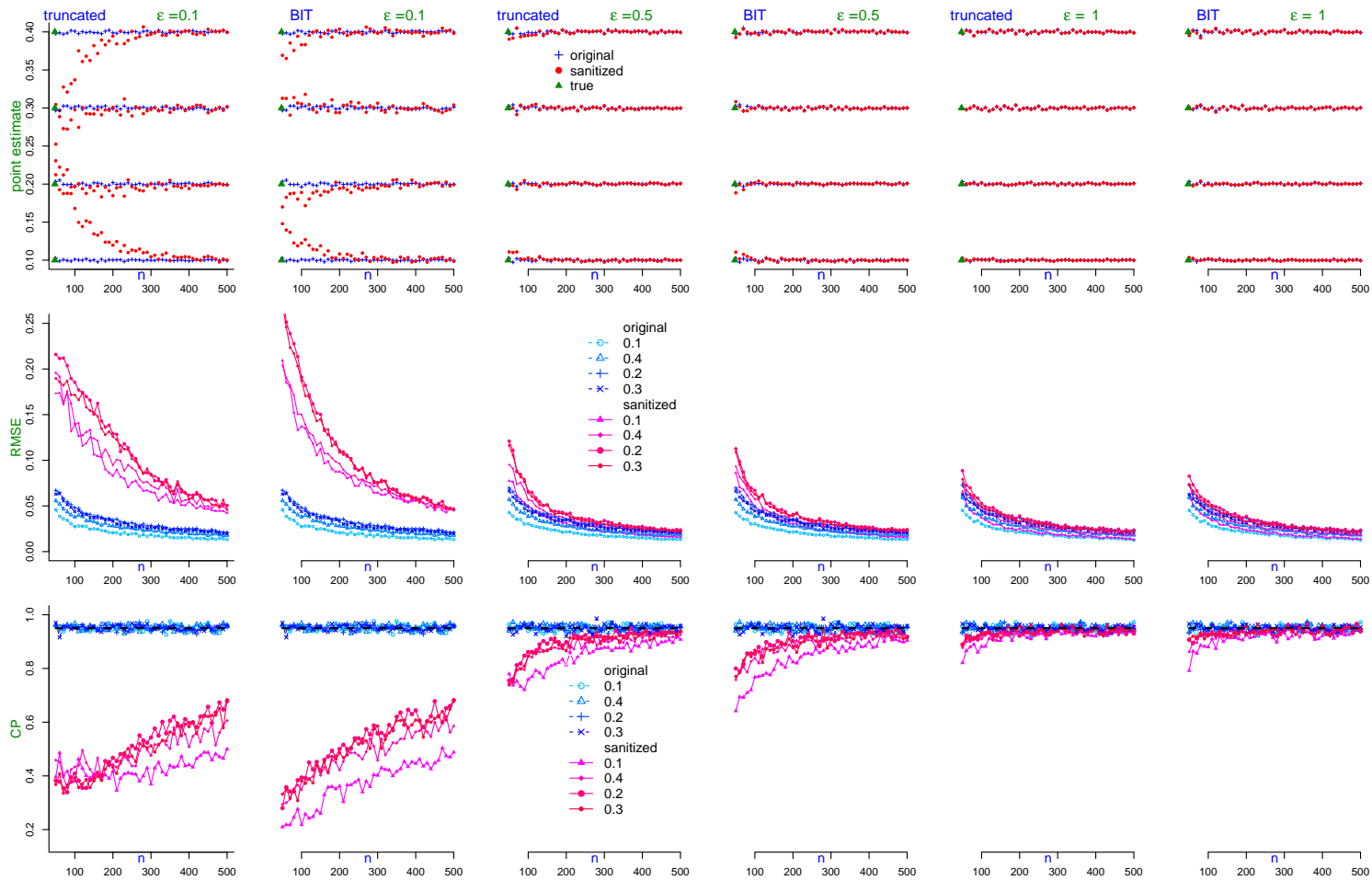
Figure 5: Bias, RMSE and CP of sanitized proportions in a modified universal histogram procedure based on Hay et. al. (2010) (red lines represent the 4 original proportions, and blue lines represent the 4 sanitized proportions)

5   Discussion

We defined two differentially private procedures for sanitizing statistics with bounding constraints and examined the statistical properties of sanitized results from the two procedures. Both the truncated and BIT Laplace procedures produce biased sanitized results relative to their original observed values unless the bounds are symmetric around the original results, which is a hard-to-satisfy condition in real life given than the original statistics change by data while the bounds are global and fixed. However, sanitized results can still be MSE-consistent for the original values if the scale parameter of the Laplace distribution associated with the two procedures approaches 0 as the sample size n approaches ∞. We also provided an upper bound for the MSE between the sanitized results relative the original for a definite n in the truncated and BIT Laplace procedures.

  Though the BIT Laplace mechanism in theory delivers less biased sanitized statistics than the truncated Laplace mechanism, the former does not seem to be more advantageous over the latter in practical applications for the following reasons. First, asymptotic unbiasedness and consistency hold in both procedures, so there is minimal difference between the two when n is large. Second, the truncated Laplace distribution is a smooth distribution while the BIT Laplace distribution is piece-wise. Though the distributional shape might be irrelevant in the release of a single sanitized statistic, it will when it comes to uncertainty quantification or making inferences about population parameters based on the sanitized data. Last, the 3-piece distributional shape of the BIT Laplace distribution requires the intervals of the outcomes to be closed on both ends so that the boundary values are exclusively defined. This is not necessary for the truncated Laplace distribution where the density function is continuous and smooth. This last point seems trivial but can be irritating in practical applications. For example, in the first simulation, closed intervals $[0, (c_1 - c_0)^2 n/(4(n-1))]$ and $[-1, 1]$ were applied to variance and correlation, respectively. As a result, some sanitized outputs were exactly 0 for variance, and exactly -1 or 1 for correlation from the BIT Laplace mechanism. In practice, these values are rare occurrences due to measurement errors and noises, and the user may choose to reject the sanitized results that are valued exactly at the boundary. If the user demands more plausible results that agree with real-life situations, what values to replace the implausible boundary values becomes an arbitrary decision and could potentially affect the statistical properties of the sanitized results. Those concerns do not exist in the truncated Laplace mechanism.

  This paper has focused on the truncation and BIT bounding procedures in the framework of the Laplace mechanism of ϵ-DP. The truncated and BIT procedures are general enough to be extended to other differentially private sanitizers that output numerical results. For example, the Gaussian mechanism of (ϵ, δ)-aDP [35] or of (ϵ, δ)-pDP [7] outputs sanitized values spanning the range $(-\infty, \infty)$, which be post-processed, if there are bounding constraints by the truncation or the BIT procedures to satisfy the constraints. The statistical properties of sanitized results in other DP mechanisms other than the the Laplace mechanism will need to examined on a case-by-case basis as the probability distribution of the injected noises are different.

  Our work is different from Ying et al. [36] though somewhat related. Ying et al. [36] also propose the concept of the unbiasedness for the true query results after the directly sanitized results are post-processed (or refinement as they refer to) based on known background knowledge and constraints. However, their theoretical work focuses on linear constraints and some of these constraints contain the original non-private information. Therefore, they examine "how refinement strategies affect differential privacy". In other words, they study DP given sanitized results that satisfy non-private linear constraints.

This is very different from our work, where the bounding and inequality constraints are public knowledge and do not contain any sensitive or private information. Therefore, DP is preserved after post-processing and our investigation focus is how satisfying the public constraints affects the utility of the sanitized results.

As briefly mentioned in the Introduction section, releasing differentially private statistics with the bounding constrains falls can be dealt with the constrained inferences approach by minimizing the loss function measuring the deviance between constrained and unconstrained sanitized results while satisfying a defined set of constraints. The simulation study implemented 3 approaches for satisfying the equality constraints, with one of them developed in the framework of the constrained inference, among a set of proportions that were also subject to bounding constraints. The simulation exercise only represents a small empirical attempt in handling constraints when releasing sanitized statistics; further research is warranted on the development of innovative and efficient approaches that release the optimal sanitized results under both inequality and equality constraints. Finally, we briefly touched on the topic of multivariable constraints throughout the discussion, including a simulation study on the multivariable linear equality constraint. More in-depth theoretical and empirical investigation on how multivariable constraints affect the statistical properties of sanitized results will be conducted in the future.

Supplementary Materials

The online supplementary materials contain the calculations of the l_1 global sensitivity of some common statistics, including the sample proportion, mean, variance, and covariance; as well as additional results from Simulation Study 2. The materials are available at arxiv `https://arxiv.org/abs/1607.08554`.

Acknowledgement

Appendix

A   Proof of Proposition 1

The mean of a truncated Laplace distribution $\text{Lap}(s, \lambda | x \in [c_0, c_1]$ is

$$
\begin{aligned}
\mu_1 =& \text{E}(x | x \in (c_0, c_1)) = (F(c_1) - F(c_0))^{-1} \int_{c_0}^{c_1} \frac{x}{2\lambda} \exp\left(-\frac{|x-s|}{\lambda}\right) dx \\
=& (F(c_1) - F(c_0))^{-1} \left[\int_{c_0}^{s} \frac{x}{2\lambda} \exp\left(\frac{x-s}{\lambda}\right) dx + \int_{s}^{c_1} \frac{x}{2\lambda} \exp\left(\frac{s-x}{\lambda}\right) dx \right] \\
=& (F(c_1) - F(c_0))^{-1} \left(s + \tfrac{1}{2}\left((\lambda - c_0) \exp\left(\tfrac{c_0-s}{\lambda}\right) - (c_1 + \lambda) \exp\left(\tfrac{s-c_1}{\lambda}\right)\right)\right) \\
=& (1 - \tfrac{1}{2} \exp(\tfrac{s-c_1}{\lambda}) - \tfrac{1}{2} \exp(\tfrac{c_0-s}{\lambda}))^{-1} \left(s + \tfrac{1}{2}\left((\lambda - c_0) \exp\left(\tfrac{c_0-s}{\lambda}\right) - (c_1 + \lambda) \exp\left(\tfrac{s-c_1}{\lambda}\right)\right)\right) \\
=& s + \left(1 - \tfrac{1}{2} \exp(\tfrac{s-c_1}{\lambda}) - \tfrac{1}{2} \exp(\tfrac{c_0-s}{\lambda})\right)^{-1} \left(\tfrac{1}{2}(\lambda - c_0 + s) \exp\left(\tfrac{c_0-s}{\lambda}\right) - \tfrac{1}{2}(c_1 + \lambda - s) \exp\left(\tfrac{s-c_1}{\lambda}\right)\right).
\end{aligned}
$$

The mean of the BIT Laplace distribution is

$$
\mu_2 = p_0 c_0 + p_1 c_1 + (1 - p_0 - p_1)\text{E}(x | x \in [c_0, c_1]).
$$

Since $p_0 = \Pr(x < c_0) = \frac{1}{2}\exp(\frac{c_0-s}{\lambda})$ and $p_1 = \Pr(x > c_1) = \frac{1}{2}\exp(\frac{s-c_1}{\lambda})$, and given the result from Part a), then μ_2 is $\frac{c_0}{2}\exp\left(\frac{c_0-s}{\lambda}\right) + \frac{c_1}{2}\exp\left(\frac{s-c_1}{\lambda}\right) + s + \frac{\lambda-c_0}{2}\exp\left(\frac{c_0-s}{\lambda}\right) - \frac{c_1+\lambda}{2}\exp\left(\frac{s-c_1}{\lambda}\right) = s + \frac{\lambda}{2}\left[\exp\left(\frac{c_0-s}{\lambda}\right) - \exp\left(\frac{s-c_1}{\lambda}\right)\right]$.

Part a): In the case of μ_1, s^* is unbiased for s if $(\lambda-c_0+s)\exp\left(\frac{c_0-s}{\lambda}\right) = (c_1+\lambda-s)\exp\left(\frac{s-c_1}{\lambda}\right)$. Let $f(x) = (\lambda + |x|)\exp(-\frac{|x|}{\lambda})$, where x is a real number. $f(x)$ is symmetric about $x = 0$. $f'(x) = -\frac{|x|}{\lambda}\exp(-\frac{|x|}{\lambda})$; therefore, $f(x)$ is a monotonic increasing function when $x < 0$ and a monotonic decreasing function when $x > 0$. Taken together, $(\lambda - c_0 + s)\exp\left(\frac{c_0-s}{\lambda}\right) = (c_1 + \lambda - s)\exp\left(\frac{s-c_1}{\lambda}\right)$ and s^* is unbiased for s iff c_0 and c_1 are symmetric about s. In the case of μ_2, s^* is unbiased for s if $\exp\left(\frac{c_0-s}{\lambda}\right) = \exp\left(\frac{s-c_1}{\lambda}\right)$. $f(x)$ is symmetric about $x = 0$ Let $f(x) = \exp(-\frac{|x|}{\lambda})$, where x is a real number. $f'(x) = -\text{sign}(x)\exp(-\frac{|x|}{\lambda})$; therefore, $f(x)$ is a monotonic increasing function when $x < 0$ and a monotonic decreasing function when $x > 0$. Taken together, $\exp\left(\frac{c_0-s}{\lambda}\right) = \exp\left(\frac{s-c_1}{\lambda}\right)$ and s^* is unbiased for s iff c_0 and c_1 are symmetric about s.

Part b): When $s - c_0 < c_1 - s$ (both > 0). Since $\exp\left(\frac{c_0-s}{\lambda}\right) \geq \exp\left(\frac{s-c_1}{\lambda}\right)$, then $\mu_2 > s$. In the case of μ_1, we have shown in Part a) that $f(x) = (|x| + \lambda)\exp(-|x|/\lambda)$ is symmetric and monotonically decreasing with $|x|$; therefore, $f(s - c_0) > f(c_1 - s)$ and the numerator in Eq. (6) is > 0. Since $\exp(\frac{s-c_1}{\lambda}) < 1$ and $\exp(\frac{c_0-s}{\lambda}) < 1$, the denominator in Eq. (6) > 0. Taken together, $\mu_1 > s$. In other words, $(\mu_1 - s)(\mu_2 - s) > 0$. When $s - c_0 > c_1 - s$, we can prove $\mu_2 < s$ and $\mu_1 < s$, and therefore $(\mu_1 - s)(\mu_2 - s) > 0$, in a similar manner as for the case of $s - c_0 < c_1 - s$.

Part c): In terms of the magnitude of the bias, we compare

$$\frac{\frac{s-c_0+\lambda}{2}\exp\left(\frac{c_0-s}{\lambda}\right) - \frac{c_1-s+\lambda}{2}\exp\left(\frac{s-c_1}{\lambda}\right)}{1 - \frac{1}{2}\exp(\frac{s-c_1}{\lambda}) - \frac{1}{2}\exp(\frac{c_0-s}{\lambda})} \quad \text{v.s.} \quad \frac{\lambda}{2}\left[\exp\left(\frac{c_0-s}{\lambda}\right) - \exp\left(\frac{s-c_1}{\lambda}\right)\right]$$

$$\Rightarrow \frac{(s-c_0)\exp\left(\frac{c_0-s}{\lambda}\right) - (c_1-s)\exp\left(\frac{s-c_1}{\lambda}\right) + \frac{\lambda}{2}\left[\exp\left(\frac{c_0-s}{\lambda}\right) - \exp\left(\frac{s-c_1}{\lambda}\right)\right]\left[\exp\left(\frac{s-c_1}{\lambda}\right) + \exp\left(\frac{c_0-s}{\lambda}\right)\right]}{1 - \frac{1}{2}\exp(\frac{s-c_1}{\lambda}) - \frac{1}{2}\exp(\frac{c_0-s}{\lambda})} \quad \text{v.s.} \quad 0$$

$$\Rightarrow \frac{(s-c_0)\exp\left(\frac{c_0-s}{\lambda}\right) - (c_1-s)\exp\left(\frac{s-c_1}{\lambda}\right) + \frac{\lambda}{2}\left[\exp\left(2\frac{c_0-s}{\lambda}\right) - \exp\left(2\frac{s-c_1}{\lambda}\right)\right]}{1 - \frac{1}{2}\exp(\frac{s-c_1}{\lambda}) - \frac{1}{2}\exp(\frac{c_0-s}{\lambda})} \quad \text{v.s.} \quad 0$$

$$\Rightarrow (s-c_0)\exp\left(\frac{c_0-s}{\lambda}\right) + \frac{\lambda}{2}\exp\left(2\frac{c_0-s}{\lambda}\right) - \left[(c_1-s)\exp\left(\frac{s-c_1}{\lambda}\right) + \frac{\lambda}{2}\exp\left(2\frac{s-c_1}{\lambda}\right)\right] \quad \text{v.s.} \quad 0 \text{ (the denominator} > 0.)$$

Let $f(x) = xe^{-x/\lambda} + \lambda e^{-2x/\lambda}/2$ and $x > 0$, then the last equation above is to compare $f(s - c_0) - f(c_1 - s)$ v.s. 0. The first derivative $f'(x) = -e^{-2x/\lambda} + e^{-x/\lambda} - xe^{-x/\lambda}/\lambda = e^{-x/\lambda}(1 - x/\lambda - e^{-x/\lambda})$. The sign of $f'(x)$ is determined by the second term $g(x) = 1 - x/\lambda - e^{-x/\lambda}$ since the first term $e^{-x/\lambda} > 0$. $g(x) = \lambda(e^{-x/\lambda} - 1)$. Since $x > 0$, then $g(x) < 0$, implying $g(x)$ decreases monotonically with increasing x, and reaches the maximum as $x \to 0$. Since $g(0) = 0$, so $g(x) < 0$ for $x > 0$. Taken together, $f'(x) = e^{-x/b}g(x) < 0$. Therefore, $f(x)$ decreases monotonically with increasing x. When $s - c_0 < c_1 - s$, $f(s - c_0) > f(c_1 - s) > 0$ or $\mu_1 - s > \mu_2 - s > 0$; when $s - c_0 > c_1 - s$, $f(s - c_0) < f(c_1 - s) < 0$ or $\mu_1 - s < \mu_2 - s < 0$. In summary, $|\mu_1 - s| > |\mu_2 - s|$.

B   Proof of Proposition 2

For the Laplace truncated mechanism, let $A = \frac{\lambda-c_0+s}{2}\exp\left(\frac{c_0-s}{\lambda}\right) - \frac{\lambda+c_1-s}{2}\exp\left(\frac{s-c_1}{\lambda}\right)$; $p_1 = \frac{1}{2}\exp(\frac{c_0-s}{\lambda})$; and $p_2 = \frac{1}{2}\exp(\frac{s-c_1}{\lambda})$, and $B = 2\lambda^2+s^2-\frac{1}{2}(2\lambda^2-2\lambda c_0 + c_0^2)\exp\left(\frac{c_0-s}{\lambda}\right) - \frac{1}{2}(2\lambda^2+2\lambda c_1 + c_1^2)\exp\left(\frac{s-c_1}{\lambda}\right)$. Since MSE is $\text{E}(s^* - s)^2 = \text{V}(s^*) + \text{bias}(s^*)^2 = \text{E}(s^{*2}) - (\text{E}(s^*))^2 + \text{bias}^2$,

thus

$$\text{bias}^2 = \left(\frac{A}{1-p_1-p_2}\right)^2; \ (\text{E}(s^*))^2 = \left(s + \frac{A}{1-p_1-p_2}\right)^2; \ \text{E}(s^{*2}) = \frac{B}{1-p_1-p_2}$$

$$\text{MSE} = \frac{B}{1-p_1-p_2} - \left(s + \frac{A}{1-p_1-p_2}\right)^2 + \left(\frac{A}{1-p_1-p_2}\right)^2 = \frac{B-2sA}{1-p_1-p_2} - s^2; \text{ where } B - 2sA =$$

$$2\lambda^2 + s^2 - \tfrac{1}{2}(2\lambda^2 - 2\lambda c_0 + c_0^2)p_1 - \tfrac{1}{2}(2\lambda^2 + 2\lambda c_1 + c_1^2)p_2 - s(\lambda - c_0 + s)p_1 + s(\lambda + c_1 - s)p_2$$

$$= -s^2 + 2(s^2 + \lambda^2)(1 - p_1 - p_2) - (-\lambda c_0 + c_0^2/2 + s(\lambda - c_0))p_1 - (\lambda c_1 + c_1^2/2 - s(\lambda + c_1))p_2$$

Therefore, MSE

$$=s^2 + 2\lambda^2 - \frac{s^2 + (-\lambda c_0 + \frac{c_0^2}{2} + s(\lambda - c_0))p_1 + (\lambda c_1 + \frac{c_1^2}{2} - s(\lambda + c_1))p_2}{1 - p_1 - p_2}$$

$$=s^2 + 2\lambda^2 - \frac{s^2 + \lambda(s - c_0)p_1 + (\frac{c_0^2}{2} - sc_0)p_1 + \lambda(c_1 - s)p_2 + (\frac{c_1^2}{2} - sc_1)p_2}{1 - p_1 - p_2}$$

$$=s^2 + 2\lambda^2 - \frac{\lambda[(s - c_0)p_1 + (c_1 - s)p_2] + [(\frac{c_0^2}{2} - sc_0)p_1 + (\frac{c_1^2}{2} - sc_1)p_2] + s^2}{1 - p_1 - p_2}$$

$$=2\lambda^2 - \frac{\lambda[(s - c_0)p_1 + (c_1 - s)p_2] + [(\frac{c_0^2}{2} - sc_0)p_1 + (\frac{c_1^2}{2} - sc_1)p_2] + s^2(p_1 + p_2)}{1 - p_1 - p_2}$$

$$=2\lambda^2 - \frac{\lambda[(s - c_0)p_1 + (c_1 - s)p_2] + [(\frac{c_0^2}{2} - sc_0 + s^2)p_1 + (\frac{c_1^2}{2} - sc_1 + s^2)p_2]}{1 - p_1 - p_2}$$

$$=2\lambda^2 - \frac{\frac{\lambda}{2}[(s - c_0)\exp(\frac{c_0 - s}{\lambda}) + (c_1 - s)\exp(\frac{s - c_1}{\lambda})]}{1 - \frac{1}{2}\exp(\frac{c_0 - s}{\lambda}) - \frac{1}{2}\exp(\frac{s - c_1}{\lambda})} +$$

$$\frac{\frac{1}{2}[(\frac{c_0^2}{2} - sc_0 + s^2)\exp(\frac{c_0 - s}{\lambda}) + (\frac{c_1^2}{2} - sc_1 + s^2)\exp(\frac{s - c_1}{\lambda})]}{1 - \frac{1}{2}\exp(\frac{c_0 - s}{\lambda}) - \frac{1}{2}\exp(\frac{s - c_1}{\lambda})}$$

$$=2\lambda^2 -$$

$$\boxed{\frac{\frac{\lambda}{2}\left[(s - c_0)\exp(\frac{c_0 - s}{\lambda}) + (c_1 - s)\exp(\frac{s - c_1}{\lambda})\right] + \frac{1}{2}[(\frac{c_0^2}{4} + (\frac{c_0}{2} - s)^2)\exp(\frac{c_0 - s}{\lambda}) + (\frac{c_1^2}{4} + (\frac{c_1}{2} - s)^2)\exp(\frac{s - c_1}{\lambda})]}{1 - \frac{1}{2}\exp(\frac{c_0 - s}{\lambda}) - \frac{1}{2}\exp(\frac{s - c_1}{\lambda})}}$$

$< 2\lambda^2$ because each term in the boxed expression is < 0 as $c_0 \le s \le c_1$.

For the BIT mechanism,

$$\text{bias}^2 = (\lambda(p_1 - p_2))^2;$$

$$(\text{E}(s^*))^2 = (s + \lambda(p_1 - p_2))^2; \ \text{E}(s^{*2}) = 2\lambda^2 + s^2 - 2(\lambda^2 - \lambda c_0)p_1 - 2(\lambda^2 + \lambda c_1)p_2$$

$$\text{MSE} = 2\lambda^2 + s^2 - 2(\lambda^2 - \lambda c_0)p_1 - 2(\lambda^2 + \lambda c_1)p_2 - (s + \lambda(p_1 - p_2))^2 + (\lambda(p_1 - p_2))^2$$

$$= 2\lambda^2(1 - p_1 - p_2) + 2\lambda((c_0 - s)p_1 + (s - c_1)p_2)$$

$$= 2\lambda^2 + \boxed{(-2\lambda^2 p_1 - 2\lambda^2 p_2 + 2\lambda((c_0 - s)p_1 + (s - c_1)p_2))}$$

$$< 2\lambda^2 \text{ because each term in the boxed expression is } < 0 \text{ as } c_0 \le s \le c_1.$$

Since $\lambda = \delta_s/\epsilon$; if $\delta_s \propto n^{-k}$ (for $k > 0$), then MSE $= O(\lambda^2) = O(n^{-2k})$ for a given ϵ in both the truncated and BIT mechanisms.

With the established upper bound $2\lambda^2$ as above, it is obvious that MSE $\to 0$ as $\lambda \to 0$.

C   Proof of Proposition 3

The proof utilizes the following lemma.

Lemma A.1: If 1) an estimator $\hat{\theta}$ is asymptotically unbiased for θ ($E(\hat{\theta}) \to \theta$ as $n \to \infty$), and 2) there exists a $k \geq 0$ such that $\int_{-} k < \infty$ and $|E(\hat{\theta}|\theta)f(\theta|\beta)| \leqq k$ for all n, where $f(\theta|\beta)$ is a probability density function, then $E(E(\hat{\theta}|\theta)|\beta) \to E(\theta|\beta)$ as $n \to \infty$.

Proof: $E(\hat{\theta}) \to \theta$ as $n \to \infty$ and $E(\hat{\theta})f(\theta|\beta) \to \theta f(\theta|\beta)$ as $n \to \infty$. With condition 2) and Theorem 2 from Cunningham (1967), we have $\int E(\hat{\theta})f(\theta|\beta)d\theta \to \int \theta f(\theta|\beta)d\theta = E(\theta|\beta)$ as $n \to \infty$.

Part a): By the the law of total expectation, $E(s^*|\boldsymbol{\theta}) = E[E(s^*|s)|\boldsymbol{\theta}]$. Since $E(s^*|s) \to s$, $E[E(s^*|s)|\boldsymbol{\theta}] \to E[s|\boldsymbol{\theta}]$ by Proposition A.1. Since $E[s|\boldsymbol{\theta}] \to \boldsymbol{\theta}$, then $E(s^*|\boldsymbol{\theta})] \to \boldsymbol{\theta}$.

Part b): By the law of total variance, $V(s^*|\boldsymbol{\theta}) = V[E(s^*|s)|\boldsymbol{\theta}] + E[V(s^*|s)|\boldsymbol{\theta}]$. Since $s^* \xrightarrow{p} s$, $V(s^*|s) \to 0$ as $n \to 0$. By Proposition A.1, $E[V(s^*|s)|\boldsymbol{\theta}] \to 0$ as $n \to 0$. Since $s^* \xrightarrow{p} s$, then $E(s^*|s) \to 0$ as $n \to 0$. By Proposition A.1, $V[E(s^*|s)|\boldsymbol{\theta}] \to 0$ as $n \to 0$. By part b), $E(s^*|s) \to s$, and $E(s|\boldsymbol{\theta}) \to \boldsymbol{\theta}$, then $E(s^*|\boldsymbol{\theta}) \to \boldsymbol{\theta}$ as $n \to 0$. All taken together, $s^* \xrightarrow{p} \boldsymbol{\theta}$.

References

[1] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adma Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography*, pages 265–284. Springer, 2006.

[2] C. Dwork. Differential privacy: A survey of results. *Theory and Applications of Models of Computation*, 4978:1–19, 2008.

[3] Cynthia Dwork. Differential privacy. In *Encyclopedia of Cryptography and Security*, pages 338–340. Springer, 2011.

[4] Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Foundations of Computer Science, 48-th Annual IEEE Symposium, FOCS'07*, pages 94–103. IEEE, 2007.

[5] Frank McSherry. Privacy integrated queries: an extensible platform for privacy-preserving data analysis. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*, pages 19–30. ACM, 2009.

[6] Q. Geng, P. Kairouz, S. Oh, and P. Viswanath. The staircase mechanism in differential privacy. *IEEE Journal of Selected Topics in Signal Processing*, 9:1176–1184, 2015.

[7] F. Liu. Generalized gaussian mechanism for differential privacy. *arXiv:1602.06028*, 2016.

[8] M. Hardt, K. Ligett, and F. McSherry. A simple and practical algorithm for differentially private data release. *Proceedings of Advances in Neural Information Processing Systems 25 (NIPS 2012)*, 2012.

[9] A. Roth and T. Roughgarden. Iterative privacy via the median mechanism. In *Proceedings of the 42nd ACM Symposium on Theory of Computing*, 2010.

[10] Cynthia Dwork and Adam Smith. Differential privacy for statistics: What we know and what we want to learn. *Journal of Privacy and Confidentiality*, 1(2):2, 2010.

[11] Kamalika Chaudhuri, Anand Sarwate, and Kaushik Sinha. Near-optimal differentially private principal components. *Proc. 26th Annual Conference on Neural Information Processing Systems (NIPS)*, 2012.

[12] Kamalika Chaudhuri, Claire Monteleoni, and Anand D. Sarwate. Differentially private empirical risk minimization. *JMLR: Workshop and Conference Proceedings*, 12:1069–1109, 2011.

[13] Daniel Kifer, Adam Smith, and Abhradeep Thakurta. Private convex empirical risk minimization and high-dimensional regression. *JMLR: Workshop and Conference Proceedings*, 23:25.1–25.40, 2012.

[14] Zuhe Zhang, Benjamin Rubinstein, and Christos Dimitrakakis. On the differential privacy of bayesian inference. *arXiv:1512.06992v1*, 2015.

[15] Avrim Blum, Katrina Ligett, and Aaron Roth. A learning theory approach to non-interactive database privacy. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 609–618. ACM, 2008.

[16] Noman Mohammed, Rui Chen, Benjamin Fung, and Philip S Yu. Differentially private data release for data mining. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 493–501. ACM, 2011.

[17] Fei Yu, Stephen E. Fienberg, Aleksandra B. Slavkovic, and Caroline Uhler. Scalable privacy-preserving data sharing methodology for genome-wide association studies. *Journal of Biomedical Informatics*, 50:133–141, 2014.

[18] C. Lin, Z.H. Song, H.B. Song, Y.H. Zhou, Y. Wang, and G.W. Wu. Differential privacy preserving in big data analytics for connected health. *Journal of Medical Systems*, 40(4):97, 2016.

[19] Debanjan Sadhya and Sanjay Kumar Singh. Privacy preservation for soft biometrics based multimodal recognition system. *Computers& Security*, 58:160–179, 2016.

[20] Ronald E Shiffler and Phillip D Harsha. Upper and lower bounds for the sample standard deviation. *Teaching Statistics*, 2(3):84–86, 1980.

[21] Boaz Barak, Kamalika Chaudhuri, Cynthia Dwork, Satyen Kale, Frank McSherry, and Kunal 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.

[22] Michael Hay, Vibhor Rastogiz, Gerome Miklauy, and Dan Suciu. Boosting the accuracy of differentially private histograms through consistency. *Proceedings of the VLDB Endowment*, 3(1):1021–1032, 2010.

[23] Wahbeh Qardaji, Weining Yang, and Ninghui Li. Understanding hierarchical methods for differentially private histograms. *Proceedings of the VLDB Endowment*, 16(14):1954–1965, 2013.

[24] Chao Li, Gerome Miklau, Michael Hay, Andrew McGregor, and Vibhor Rastogi. The matrix mechanism: optimizing linear counting queries under differential privacy. *The VLDB Journal*, 24(6):757–781, 2015.

[25] Cynthia Dwork, K. Kenthapadi, Frank McSherry, I. Mironov, and M. Naor. Our data, ourselves: privacy via distributed noise generation. In *Advances in Cyptology: Proceedings of EUROCRYPT*, pages 485–503. Springer Berlin Heidelberg, 2006.

[26] Ashwin Machanavajjhala, Daniel Kifer, John Abowd, Johannes Gehrke, and Lars Vilhuber. Privacy: Theory meets practice on the map. *IEEE ICDE IEEE 24th International Conference*, pages 277 – 286, 2008.

[27] Rob Hall, Alessandro Rinaldoy, and Larry Wasserman. Random differential privacy. *Journal of Privacy and Confidentiality*, 4(2):43–59, 2012.

[28] Cynthia Dwork and Guy N. Rothblum. Concentrated differential privacy. *arXiv:1603.01887v2*, 2016.

[29] F. Liu. Model- based differentially private data synthesis. *arXiv:1606.08052v2*, 2017.

[30] Nicholas J. Higham. Computing the nearest correlation matrixa problem from finance. *IMA Journal of Numerical Analysis*, 22:329–343, 2002.

[31] Hou-Duo Qi and Defeng Sun. A quadratically convergent newton method for computing the nearest correlation matrix. *SIAM Journal on Matrix Analysis and Applications*, 28(2), 2006.

[32] Ruediger Borsdorf and Nicholas J. Higham. A preconditioned newton algorithm for the nearest correlation matrix. *IMA Journal of Numerical Analysis*, 30:94–107, 2010.

[33] Ruediger Borsdorf, Nicholas Higham, and Marcos Raydan. Computing a nearest correlation matrix with factor structure. *SIAM Journal on Matrix Analysis and Applications*, 31(5):2603–2622, 2010.

[34] C. Bowen and F. Liu. Differentially private data synthesis methods. *arXiv:1602.01063v1*, 2016.

[35] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Theoretical Computer Science*, 9(3-4):211–407, 2013.

[36] Xiaowei Ying, Xintao Wu, and Yue Wang. On linear refinement of differential privacy-preserving query answering. In *Advances in Knowledge Discovery and Data Mining, 17th Pacific-Asia Conference Proceedings*, pages 353–364. IEEE, 2013.