Accuracy for Differentially Private Quotients by Fractional Uncertainties

Accuracy for Differentially Private Quotients by Fractional Uncertainties

Alejandro Russo
Chalmers University of Technology
and Gothenburg University
Gothenburg, Sweden
DPella AB
Gothenburg, Sweden
russo@chalmers.se

Elisabet Lobo-Vesga
DPella AB
Gothenburg, Sweden
lobo@dpella.io

Marco Gaboardi
Boston University
Boston, USA
DPella AB
Gothenburg, Sweden
gaboardi@bu.edu

Abstract

Differential Privacy (DP) is a cornerstone for ensuring privacy in data analysis by injecting carefully calibrated noise into statistical queries. While numerous DP tools focus on privacy protection, few provide accuracy information, specially for data-dependent computations like averages or quotients of DP-sums. This paper introduces a novel approach to compute confidence intervals, i.e., α-β accuracy, for these computations, leveraging principles from uncertainty propagation. Our method identifies conditions under which analytical error can be predicted, revealing two key invariants: the analytical error improves with large dataset sizes, and addition of values with higher variability require larger dataset sizes for accurate estimation. To simplify adoption, we also propose accuracy tuners to enable rapid determination of minimum dataset sizes and explore trade-offs between privacy budgets and the possibility to perform accuracy estimations. Our theoretical contributions are validated through an empirical evaluation that explores the applicability of fractional uncertainties for computing concrete α-β error across diverse scenarios.

CCS Concepts

• **Security and privacy** → **Privacy-preserving protocols**; • **Mathematics of computing** → *Nonparametric statistics*;

Keywords

Averages, Quotients, Uncertainty propagation, α-β Accuracy, Differential Privacy

ACM Reference Format:
Alejandro Russo, Elisabet Lobo-Vesga, and Marco Gaboardi. 2025. Accuracy for Differentially Private Quotients by Fractional Uncertainties. In *Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security (CCS '25), October 13–17, 2025, Taipei, Taiwan.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3719027.3744799

CCS '25, October 13–17, 2025, Taipei, Taiwan.

ACM ISBN 979-8-4007-1525-9/2025/10
https://doi.org/10.1145/3719027.3744799

1 Introduction

In today's digital era, vast quantities of individual data are collected daily for research or statistical purposes. However, privacy concerns surrounding the individuals who contribute their data restrict how this information can be utilized and disseminated. In response to these challenges, Differential Privacy (DP) [9] is increasingly recognized as an effective solution for releasing statistical information about populations while safeguarding the privacy of data subjects.

A common approach to implementing DP involves adding statistical noise to the output of a data analysis. When carefully calibrated, this noise ensures privacy protection while still allowing for meaningful insights about the population from which the data are drawn. The quantitative formulation of DP, defined by parameters ϵ and δ, provides a robust mathematical framework for rigorously assessing the privacy-accuracy trade-offs. It is worth noting that the accuracy requirement is not an inherent aspect of DP; rather, it becomes explicitly relevant when designing a differentially private analysis for a specific task.

In recent years, there has been a proliferation of DP tools. Most of them focus on providing privacy protection by the implementation of DP mechanisms (e.g., [3, 16, 23, 31, 33, 34, 36–38, 42]). Only a handful of them provide accuracy information to data analysts writing queries [13–15, 26, 32]. Such tools use *probability bounds* to report analytical error bounds for query results under DP [11]. Specifically, α-β accuracy provides a probabilistic upper bound on the error of a query's result. Concretely, for given parameters α (error threshold) and β (failure probability), a DP mechanism ensures that the probability of the query's error exceeding α is at most β. Intuitively, when noise gets added to ensure privacy, the *inverse cumulative distribution function* (iCDF) of the noise distribution can be used to determine the accuracy information as a probabilistic bound. While these probability bounds are a great tool to reason about privacy about a single query, they pose challenges when applied to the composition of multiple queries, such as additions [26] or data-dependent analysis.

Averages are inherently data-dependent in a DP setting because their sensitivity—the maximum change in the output caused by modifying a single input—depends on the size of the dataset itself, which might not always be a publicly available value. Specifically, the sensitivity of an average decreases as the dataset size increases. Since DP mechanisms rely on adding noise calibrated to the sensitivity, the amount of noise required for privacy protection changes with the dataset size. The following pseudocode snippet illustrates how to compute the average age (from 0 to 120 years old) in a

DP setting using a Laplace mechanism (i.e., drawing noise from a Laplace distribution). In this example, the sum of ages (`nSum`) is privatized with $\epsilon = 1$, and the result is divided by the dataset size (`rows`) to obtain the average. This approach assumes a bounded notion of DP [22], where the dataset size explicitly known—a critical requirement for noise calibration, note how the dataset size is directly accessed in line 6, bypassing any DP mechanism.

Listing 1: Average for bounded DP (naive)

```
1  query dataset = do
2    nSum <- dpSum Laplace{epsilon = 1}
3                  Args{min_v = 0, max_v = 120}
4                  dataset
5    return (nSum / rows) -- average calculation
6    where rows = length dataset -- free fact
```

In this code, the noise is calibrated to the sensitivity of the sum, i.e., 120. However, that is not the actual sensitivity of the average: the average changes at most $120/n$ if we change one row in the dataset, which is much less than 120!—observe the free use of n in the sensitivity calculation. Taking advantage of the free availability of n, DP averages can be computed with a special primitive that takes the size of the dataset as an argument:

Listing 2: Average for bounded DP

```
1 query dataset = do
2   nAvg <- dpAvg rows
3                 Laplace{epsilon = 1}
4                 Args{min_v = 0, max_v = 120}
5                 dataset
6   return nAvg -- noisy average
7   where rows = length datasets -- free fact
```

This approach is followed by popular DP libraries like DiffPrivLib [16] and Opacus [41]. We argue that bound DP is not a realistic choice for a programming framework where operations such as filtering or joining data can alter the size of the dataset.

When the dataset size itself is private—such as after filtering some rows or using unbounded DP notions [22]—reasoning about α-β accuracy becomes significantly more challenging. This complexity arises from the interaction between the noises added to the numerator (the differentially private sum) and the denominator (the differentially private dataset size), which together influence the accuracy of the final output. The following pseudocode highlights this challenge by calculating an average working-age of the population in a given the dataset (`dataset'`). In this example, the Laplace mechanism is used to privatize both, the sum of the ages (with *epsilon* = 1), and the size of filtered dataset (with $\epsilon = 0.1$) containing only the working-age population.

Listing 3: Average for transformed datasets

```
query dataset = do
  let dataset' = filter working_age dataset
  nSum <- dpSum Laplace{epsilon = 1}
                Args{min_v = 0, max_v = 120}
                dataset'
  nRows <- dpCount Laplace{epsilon = 0.1}
                   Args{}
                   dataset'
  return (nSum / nRows) -- iCDF?
  where working_age row = (row ^. age >= 18)
                       && (row ^. age <=65)
```

The code describes the most common way to calculate the average, i.e., as post-processing operation between two DP queries: a noisy sum (`nSum`) and a noisy count (`nRows`). This approach is adopted by well-known DP libraries like OpenDP [13], SmartNoise SQL [36], Qrlew [34], and PipelineDP [37].

Unlike simpler cases where iCDFs provide analytical error bounds, deriving accuracy bound for the average in Listing 3 requires analytically determining the distribution of the quotient of two Laplace-distributed random variables (i.e., `nSum / nRows`). This process is non-trivial and introduces significant mathematical complexity. In turn, this added complexity makes it challenging to use iCDFs to deduce accuracy bounds for the final result, particularly in scenarios involving data transformations such as filtering or joining.

The difficulty of reasoning about α-β accuracy for the code in Listing 3 is further underscored by the lack of support for such computations in existing tools. As an approximation, OpenDP [13] proposes *eliminating* one random variable to compute the accuracy of averages. In this approach, the data analyst provides an estimated dataset size, e, which the system uses to sample or impute data. If the estimate is smaller than the actual dataset size, the system samples the specified number of records. If the estimate is larger, the system imputes additional records using a provided default value. The noisy sum is then computed over the sampled or imputed dataset, and the noisy average is calculated as the noisy sum over the estimated record count. Since the record count is treated as constant and the noise distribution of a random variable divided by a constant is known, error estimation for the average calculation becomes straightforward—recall Listing 1. While a resourceful approach, it introduces its own challenges. The accuracy of this method relies on the quality of the analyst's guess, which may be unreliable due to their limited knowledge of the dataset. Additionally, this approach does not account for errors arising from incorrect guesses, further complicating the accuracy guarantees for averages.

The pursuit of accuracy estimations under noisy quantities extends beyond Differential Privacy. In physics, understanding measurement errors and how they propagate through operations is known as *uncertainty propagation* (e.g., [40]). This theory provides equations to handle uncertainties in measurements and their operations, particularly in determining error in the quotients of two measurements with uncertainties, assuming instrument errors *are both independent and small*. In this work, we draw a novel parallel between noisy sums and counts in DP and physical measurements with uncertainties, enabling the use of uncertainty propagation principles to estimate the accuracy of a noisy average or quotients of DP-sums. However, in the context of DP, the noise introduced to ensure privacy may not always align with the assumptions of small uncertainties. The magnitude of DP noise can vary significantly, posing challenges to directly applying classical uncertainty propagation techniques.

The novelty of this paper is the adaptation of uncertainty propagation to α-β accuracy in the context of *unbounded* Differential Privacy. We outline sufficient conditions under which fractional

uncertainties can be employed to analytically estimate the errors in calculating averages and quotients of DP-sums. Our approach is applicable to both the Laplace and Gaussian mechanisms, offering—to the best of our knowledge—*the first analytical accuracy predictions for quotients under these mechanisms.*

Our method identifies *the conditions under which uncertainty propagation can be applied to provide an analytical error* and uses *a noisy count for the accuracy calculations.* Intuitively, a noisy count can be used to reason probabilistically about the bounds of the DP-sums, thus offering a structured way to estimate the sensitivity of the average or the quotients of two DP-sums. This, in turn, facilitates α-β accuracy estimations. Our mathematical equations reveal two critical invariants: the more records there are in the dataset, the more likely it is to provide analytical error bounds; and the greater the variability (range) of the elements involved in the sum, the larger the dataset needs to be to ensure analytical error bounds can be derived.

We also present *accuracy tuners* designed to aid data analysts in understanding and managing the conditions necessary for providing accurate estimation; conditions related to the noisy count, its error bound, the domains of the elements contributing to the DP-sums, along with their privacy settings and error bounds. Since we do not expect data analysts to have all these conditions in mind, our tuners provide an intuitive way to evaluate the trade-offs among the necessary parameters. Specifically, the tuners are useful in the following scenarios:

- *Unknown dataset size*: Given the privacy parameters for counting the numbers of rows, the tuners quickly determine the minimum number of records and the minimum privacy budget required for the DP-sum(s) in order to estimate the accuracy of the average (or quotient of DP-sums).
- *Known dataset size*: For a given noisy count of the number of records, the tuners identify the minimum privacy budget for the DP-sum(s) needed to provide accuracy of the average (or quotient of DP-sums).
- *Budget optimization*: given a privacy budget for the average (or quotient of DP-sums), the tuner suggests how to distribute the budget between the noisy count and the DP-sum(s) to ensure that accuracy can be analytically predicted.

In addition to our theoretical contributions, we include an evaluation section that explores concrete scenarios and parameter values under which our hypotheses about using fractional uncertainties for computing α-β error hold. This empirical analysis provides practical insights into the conditions required for the successful application of our methods. We examine various combinations of dataset sizes, ranges of DP-sums, and privacy parameters to showcase when fractional uncertainties yield accuracy estimations. Our evaluation validates the sufficient conditions for applying fractional uncertainties, illustrating the impact various elements have on their practical applicability. These results not only reinforce the soundness of our approach but also provide actionable guidance for data analysts aiming to balance privacy guarantees and accuracy in their computations.

In summary, the contributions of this work are as follows:

- *Theoretical Framework for Analytical Accuracy*: we provide a novel approach to compute α-β accuracy for averages and quotients of DP-sums. Our novel method identifies conditions under which uncertainty propagation can be applied, enabling accurate error estimation despite the challenges of noise interactions.
- *Accuracy Tuners*: we introduce accuracy tuners to simplify the practical application of our methods.
- *Empirical Validation*: through an evaluation section, we validate our theoretical hypotheses and demonstrate when and how fractional uncertainties can be leveraged to compute analytical errors with concrete parameters values, providing actionable guidance for balancing privacy and accuracy in real-world scenarios.

2 Preliminaries

Differential Privacy (DP) is a quantitative notion of privacy that bounds how much a single individual's private data can affect the result of a data analysis. Formally, differential privacy is a property of a randomized query $\tilde{Q}(\cdot)$ representing the data analysis, as follows.

Definition 2.1 (Differential Privacy [11]). A randomized query $\tilde{Q}(\cdot) : \text{db} \to \mathbb{R}$ satisfies (ε, δ)-differential privacy if and only if for all pairs of neighboring datasets D_1 and D_2 in db differing in at most one element, and for all measurable sets S in the range of $\tilde{Q}$ (i.e., $S \subset \mathbb{R}$), it holds that

$$\Pr\big[\tilde{Q}(D_1) \in S\big] \leq e^{\varepsilon} \Pr\big[\tilde{Q}(D_2) \in S\big] + \delta.$$

In the definition above, the parameters (ε, δ) determine a bound on the distance between the distributions induced by $\tilde{Q}(\cdot)$ when adding or removing an individual from the dataset. When the parameter $\delta = 0$, the definition above is known as pure-DP, while when $\delta > 0$ is called approximated-DP.

To protect all the different ways in which an individual's data can affect the result of a query, the noise needs to be calibrated to the maximal change that the result of the query can have when changing an individual's data. This is formalized through the notion of *sensitivity.*

Definition 2.2 (Sensitivity [11]). The (global) sensitivity of a deterministic query is a measure of how much the result of the query can change when adding or removing an individual from the dataset. Concretely the sensitivity of a query $Q(\cdot) : \text{db} \to \mathbb{R}$ is defined as the quantity:

$$\Delta_Q = \max\{|Q(D_1) - Q(D_2)|\}$$

for D_1, D_2 differing in at most one row.

A well-known method for implementing pure DP queries is the Laplace mechanism, which relies on noise drawn from the Laplace distribution.

Theorem 2.3 (Laplace Mechanism [11]). *Let $Q(\cdot) : db \to \mathbb{R}$ be a deterministic query with sensitivity Δ_Q. Let $\tilde{Q}(\cdot) : db \to \mathbb{R}$ be a randomized query defined as*

$$\tilde{Q}(D) = Q(D) + Lap\left(\frac{\Delta_Q}{\varepsilon}\right),$$

where $Lap(\Delta_Q/\varepsilon)$ denotes the Laplace distribution with mean $\mu = 0$ and scale $b = \frac{\Delta_Q}{\varepsilon}$. Then, $\tilde{Q}(\cdot)$ is $(\varepsilon, 0)$-differentially private, or simply ε-differentially private.

A standard approach to achieve approximate-DP is based on the addition of noise sampled from the Gaussian distribution, this method is known as the Gaussian mechanism.

Theorem 2.4 (Gaussian Mechanism [11]). *Let $Q(\cdot) : db \to \mathbb{R}$ be a deterministic query with sensitivity Δ_Q. Let $\tilde{Q}(\cdot) : db \to \mathbb{R}$ be a randomized query defined as*

$$\tilde{Q}(D) = Q(D) + \mathcal{N}\left(\sqrt{2 * \log\left(\frac{1.25}{\delta}\right)} * \frac{\Delta_Q}{\varepsilon}\right),$$

where $\epsilon, \delta \in (0, 1)$, and $\mathcal{N}\left(\sqrt{2 * \log(1.25/\delta)} * (\Delta_Q/\varepsilon)\right)$ denotes the Gaussian distribution with scale mean $\mu = 0$ and standard deviation $\sigma = \sqrt{2 * \log(1.25/\delta)} * (\Delta_Q/\varepsilon)$. Then, $\tilde{Q}(\cdot)$ is (ε, δ)-differentially private.

In general, the notion of α-β accuracy using *confidence intervals* can be defined as follows.

Definition 2.5 (Accuracy [11]). Given an (ε, δ)-differentially private query $\tilde{Q}(\cdot)$, a target deterministic query $Q(\cdot)$, a distance function $\mathrm{d}(\cdot)$, a bound α, and the probability β; $\tilde{Q}(\cdot)$ is $(\mathrm{d}(\cdot), \alpha, \beta)$-accurate with respect to $Q(\cdot)$ if and only if for any dataset D, it holds that

$$\Pr\left[\mathrm{d}(\tilde{Q}(D), Q(D)) > \alpha\right] \leq \beta$$

This definition allows one to express data-independent error statements such as: with probability at least $1 - \beta$ the result of the query $\tilde{Q}(\cdot)$ diverges from the result of $Q(\cdot)$, in terms of the distance $\mathrm{d}(\cdot)$, for at most α. Then, we will refer to α as the error, β as the confidence probability, and $[-\alpha, \alpha]$ as the confidence interval. For the rest of the document, the considered distance function is that on real numbers: $\mathrm{d}(x, y) = |x - y|$. There are known results about the accuracy of queries using the Laplace and Gaussian Mechanisms.

Definition 2.6 (Accuracy for the Laplace Mechanism [11]). Given an ε-differentially private query $\tilde{Q}(\cdot) : \mathrm{db} \to \mathbb{R}$ implemented with the Laplace Mechanism, it holds that:

$$\Pr\left[\left|\tilde{Q}(D) - Q(D)\right| > \log\left(\frac{1}{\beta}\right) * \frac{\Delta_Q}{\varepsilon}\right] \leq \beta$$

Definition 2.7 (Accuracy for the Gaussian Mechanism [11]). Given a (ε, δ)-differentially private query $\tilde{Q}(\cdot) : \mathrm{db} \to \mathbb{R}$ implemented with the Gaussian Mechanism where $\epsilon, \delta \in (0, 1)$, it holds that:

$$\Pr\left[\left|\tilde{Q}(D) - Q(D)\right| > \sigma * \sqrt{2 * \log\left(\frac{2}{\beta}\right)}\right] \leq \beta$$

whit standard deviation $\sigma = \sqrt{2 * \log(1.25/\delta)} * (\Delta_Q/\varepsilon)$.

These definitions use the inverse cumulative distribution function (iCDF) of the noise distribution to provide the corresponding error bounds. Concretely, from the definitions above we have that the iCDF of the Laplace distribution is given by $\mathrm{icdf}(\Delta_Q, \epsilon, 0, \beta) = \log(1/\beta) * \Delta_Q/\varepsilon$, and the iCDF of the Gaussian distribution is given by $\mathrm{icdf}(\Delta_Q, \epsilon, \delta, \beta) = \sigma * \sqrt{2 * \log(2/\beta)}$. We note that the iCDF of Laplace is exact while the one for Gaussian noise is an approximation. There is work on obtaining tighter bounds and relaxing the restriction of $\epsilon < 1$ [2] called analytical Gaussian Mechanism. In this work, however, we focus on the Gaussian Mechanism as described in [11] and leave extending our approach to the analytical Gaussian as future work.

In the field of physics, reasoning about the uncertainties in measurements and how they propagate through operations is a fundamental task [40]. Uncertainty propagation assuming small and independent errors on the measurements being combined. In particular, we have the following uncertainty propagation formula for the quotient of two independent measurements.

Definition 2.8 (Error propagation for the quotien of two measurements [40]). Given two measurements $\tilde{x}$ and $\tilde{y}$ with uncertainties δ_x and δ_y, the error propagation for the quotient $\tilde{r} = \tilde{x}/\tilde{y}$ is given by:

$$\delta_r = |\tilde{r}| * \left(\frac{\delta_x}{|\tilde{x}|} + \frac{\delta_y}{|\tilde{y}|}\right)$$

with $\delta_x/|\tilde{x}|$ and $\delta_y/|\tilde{y}|$ being *small*.

In the definition above, $\delta_x/|\tilde{x}|$ and $\delta_y/|\tilde{y}|$ are known as fractional uncertainties.

3 Fractional uncertainties for DP quotients

In this section, we will demonstrate the originality of our approach for calculating the α-β accuracy for noisy averages and quotients involving noisy sums. In a nutshell, our methodology involves initially determining a noisy count of the records used in the computation of the noisy sums. Unlike the bounded case, where the original size of the dataset is freely available, our approach approximates the accuracy of the averages or quotients between sums by utilizing a noisy count and static information about the range of possible values for each element in the sums. For clarity, we first focus on the accuracy of averages before extending the discussion to quotients.

3.1 Accuracy of averages

We define the differentially private average as the ratio of two differentially private queries: a count and a sum. Concretely, we define the DP average and its privacy guarantees as follows:

Definition 3.1 (Differetially Private Average). Given a $(\varepsilon_c, \delta_c)$-differentially private count $\tilde{c}$ and an $(\varepsilon_s, \delta_s)$-differentially private sum $\tilde{s}$. Then $\tilde{\mathrm{avg}} = \frac{\tilde{s}}{\tilde{c}}$ is an $(\varepsilon_c + \varepsilon_s, \delta_c + \delta_s)$-differentially private average of the dataset.

We aim to compute the error bound, α_{avg}, for the noisy average, $\tilde{\mathrm{avg}}$. Specifically, we seek to determine α_{avg} such that for a given probability β, the following holds:

$$\Pr\left[|\tilde{\mathrm{avg}} - \mathrm{avg}| > \alpha_{\mathrm{avg}}\right] \leq \beta$$

However, analytically determining the distribution of $\tilde{\mathrm{avg}}$ is challenging, as it arises from the ratio of two random variables. In light of this difficulty, we propose an alternative approach: using uncertainty propagation for the quotient of two measurements to

analytically estimate α_{avg}. Recalling Definition 2.8, we would like to obtain an α-β accuracy guarantee as follows:

$$\Pr\left[\left|\frac{\tilde{s}}{\tilde{c}} - \frac{s}{c}\right| > \left|\frac{\tilde{s}}{\tilde{c}}\right| * \left(\frac{\alpha_c}{|\tilde{c}|} + \frac{\alpha_s}{|\tilde{s}|}\right)\right] \leq \beta \tag{1}$$

where s and c are the true sum and count of the dataset, respectively. Furthermore, $\tilde{c}$ is a $(\varepsilon_c, \delta_c)$-differentially private count with error bound computed as $\alpha_c = \text{icdf}(1, \varepsilon_c, \delta_c, \beta_c)$, and $\tilde{s}$ is a $(\varepsilon_s, \delta_s)$-differentially private sum with error bound computed as $\alpha_s = \text{icdf}(\Delta_s, \varepsilon_s, \delta_s, \beta_s)$. The challenge about equation (1) is that we need to reason about the distribution of the random variable $\tilde{s}/\tilde{c}$, which is something far from trivial. To simplify the problem, we could simply sample from both random variables, i.e., $\tilde{v}_{\tilde{c}} \leftarrow \tilde{c}$ and $\tilde{v}_{\tilde{s}} \leftarrow \tilde{s}$, and then compute the error of the average as indicated by Definition 2.8, i.e., $\alpha_{\text{avg}} = \left|\frac{\tilde{v}_{\tilde{s}}}{\tilde{v}_{\tilde{c}}}\right| * \left(\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} + \frac{\alpha_s}{|\tilde{v}_{\tilde{s}}|}\right)$. For that, however, the fractional uncertainties $\alpha_c/|\tilde{v}_{\tilde{c}}|$ and $\alpha_s/|\tilde{v}_{\tilde{s}}|$ need to be *small*, but what does small mean? Furthermore, how does the requirement of being *small* affect any of the parameters of the mechanism for the noisy count and sum?

To answer the first question, we will introduce $0 < \gamma < 1$ as a *small* constant, and require that, when we sample from the mechanisms, i.e., $\tilde{v}_{\tilde{c}} \leftarrow \tilde{c}$ and $\tilde{v}_{\tilde{s}} \leftarrow \tilde{s}$, then it holds (with certain probability) that γ is an upper bound to the ratios $\alpha_c/|\tilde{v}_c|$ and $\alpha_s/|\tilde{v}_s|$, representing the *fractional uncertainty* of the count and sum, respectively. By doing that, the product of relative uncertainties, i.e., $\alpha_c/|\tilde{v}_{\tilde{c}}| \cdot \alpha_s/|\tilde{v}_{\tilde{c}}|$ can be neglected, thus being able to apply the formula in Definition 2.8. More technically, any values $0 < \gamma < 1$ that allow to approximate the binomial theorem $1/1-z = \sum_{i=0}^{\infty} z^i$ by $1/1-z \approx 1 + z$ works [40]—in our case, $z = \alpha_c/|\tilde{v}_{\tilde{c}}|$.

The subsequent propositions address the second question for both the Laplace and Gaussian mechanism. In simpler term, the ability to apply the uncertainty propagation formula α_{avg} is affected by (i) the size of the dataset, (ii) the privacy budget, and (iii) the sensitivity of the sum —detailed proofs can be found in the extended version of this paper. The following proposition states that the more rows a dataset has, the more likely to make the fractional uncertainty $\alpha_c/|\tilde{c}|$ small, that is, every time that we sample $\tilde{v}_{\tilde{c}} \leftarrow \tilde{c}$, it is likely to hold that $\alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma$.

Proposition 3.2. *Given $0 < \gamma < 1$, and a $(\varepsilon_c, \delta_c)$-differentially private count $\tilde{c}$ with error bound computed as $\alpha_c = icdf(1, \varepsilon_c, \delta_c, \beta_c)$, then it holds with probability $1 - \beta_c$ that*

$$c \geq \frac{\alpha_c * (1 + \gamma)}{\gamma} \Rightarrow \alpha_c/|\tilde{c}| \leq \gamma$$

where c is the number of records in the dataset.

For the fractional uncertainty $\alpha_s/|\tilde{s}|$ to be small, the following proposition indicates that most of the budget for the average must be used on the sum rather than the count when a is much smaller than b—see hypothesis $b/\varepsilon_s \leq a/\varepsilon_c$. As the proposition shows, our results hold for sums computed from positive values. We also see that the amount of records in the dataset (c) times the minimum value of the sum (a) is bigger than the error of it (α_s).

Proposition 3.3. *Given $0 < \gamma < 1$, $\alpha_c = icdf(1, \varepsilon_c, \delta, \beta)$, and the privacy parameters (ε_s, δ) for a differentially private sum to be performed with error bound computed as $\alpha_s = icdf(\Delta_s, \varepsilon_s, \delta, \beta)$, where $\Delta_s = \max\{a, b\}$, $a > 0$, a and b being the lower and upper bounds of the values to be added, respectively, and c as the number of records in the dataset, then it holds with probability $(1 - \beta)^2$ that*

$$(c * a > \alpha_s) \wedge \left(\frac{\alpha_c}{c - \alpha_c} \leq \gamma\right) \wedge \left(\frac{b}{\varepsilon_s} \leq \frac{a}{\varepsilon_c}\right) \Rightarrow \alpha_s/|\tilde{s}| \leq \gamma \tag{2}$$

Observe that the propositions above require access to the true count c, which is not available in practice—recall the motivation from Section 1. Instead, we would like to use the value of a differentially-private count $\tilde{v}_{\tilde{c}}$, where $\tilde{v}_{\tilde{c}} \leftarrow \tilde{c}$, to provide a lower and upper bound on the result of the sum (with certain probability). In that manner, a data analyst can first spend some budget into performing a DP-count to obtain the size of the database for then obtaining the accuracy of the average before sampling from the DP-sum. In other words, we will be able to compute the average's error bound (and its preconditions) using only a concrete DP-count value.

The constraint $c * a > \alpha_s$ is required to avoid division by zero in our mathematical development, which then implies that $a > 0$. Furthermore, the proposition utilizes the fact that $a > 0$ to provide upper and lower bounds for $|\tilde{s}|$ in terms of the real count c and the limits a and b, which then enables to bound $\alpha_s/|\tilde{s}|$.

Proposition 3.4. *Given a $(\varepsilon_c, \delta_c)$-differentially private count $\tilde{c}$ with error bound computed as $\alpha_c = icdf(1, \varepsilon_c, \delta, \beta_c)$, $\tilde{v}_{\tilde{c}} \leftarrow \tilde{c}$, the privacy parameters (ε_s, δ) for a differentially private sum with error bound computed as $\alpha_s = icdf(\Delta_s, \varepsilon_s, \delta, \beta_s)$, where $\Delta_s = \max\{a, b\}$, $a > 0$, a and b being the lower and upper bounds of the values to be added, respectively, then it holds with probability $(1 - \beta_c) \cdot (1 - \beta_s)$:*

$$\frac{\alpha_s}{|\tilde{s}|} \leq \frac{\alpha_s}{a * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_s} \tag{3}$$

This proposition provides an upper bound to the fractional uncertainty of the differentially-private sum about to be performed based on a given noisy count.

The following proposition indicates under which conditions, and based on the available noisy count, when the fractional uncertainty of the differentially-private sum is likely to be small.

Proposition 3.5. *Given $0 < \gamma < 1$, a $(\varepsilon_c, \delta_c)$-differentially private count $\tilde{c}$ with error bound computed as $\alpha_c = icdf(1, \varepsilon_c, \delta, \beta)$, $\tilde{v}_{\tilde{c}} \leftarrow \tilde{c}$, such that $\alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma$, the privacy parameters (ε_s, δ) for a differentially private sum with error bound computed as $\alpha_s = icdf(\Delta_s, \varepsilon_s, \delta, \beta)$, where $\Delta_s = \max\{a, b\}$, $a > 0$, a and b being the lower and upper bounds of the values to be added, respectively, then it holds with probability $(1 - \beta)^2$:*

$$\frac{b}{\varepsilon_s} \leq \frac{a}{\varepsilon_c} * \left(\frac{1 - \gamma}{1 + \gamma}\right) \Rightarrow \frac{\alpha_s}{|\tilde{s}|} \leq \gamma$$

Under the assumptions of Proposition 3.5, and that $\alpha_c/|\tilde{v}_{\tilde{c}}|$ is also small, we can apply the uncertainty propagation equations in Definition 2.8 to estimate α_{avg}. Unfortunately, calculating α_{avg} requires to sample from the DP-sum. Can we obtain an error estimate for the average without sampling from the DP-sum?

With that in mind, we propose a new equation for the error bound of the average that does not depend on the result of the DP-sum but rather on an approximation in terms of the noisy DP-count.

To attain our goal, we need to determine an upper bound for the error for the mean as dictated by Definition 2.8 when sampling

from the DP-sum, i.e., when obtaining $\tilde{v}_{\tilde{s}} \leftarrow \tilde{s}$, then $\alpha_s/|\tilde{v}_{\tilde{s}}| \leq \gamma$ and $\alpha_{\text{avg}} = |\tilde{v}_{\tilde{s}}/\tilde{v}_{\tilde{c}}| * (\alpha_c/|\tilde{v}_{\tilde{c}}| + \alpha_s/|\tilde{v}_{\tilde{s}}|)$.

In what follows, we keep using $\tilde{s}$ since we want our reasoning to hold (with certain probability) for any given sampling of the DP-sum. Informally, we proceed as follows where $a > 0$ and $\beta_c = \beta_s$:

$$
\begin{aligned}
&\left|\frac{\tilde{s}}{\tilde{v}_{\tilde{c}}}\right| * \left(\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} + \frac{\alpha_s}{|\tilde{s}|}\right) \\
&\leq \left|\frac{\tilde{s}}{\tilde{v}_{\tilde{c}}}\right| * \left(\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} + \frac{\alpha_s}{a * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_s}\right) \text{; by Proposition 3.4} \\
&\leq \frac{|\tilde{s}|}{|\tilde{v}_{\tilde{c}}|} * \left(\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} + \frac{\alpha_s}{a * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_s}\right) \text{; by properties of } |\cdot| \\
&\leq \frac{b * c + \alpha_s}{|\tilde{v}_{\tilde{c}}|} * \left(\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} + \frac{\alpha_s}{a * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_s}\right) \\
&\quad \text{; by upper bound of } |\tilde{s}| \text{ with prob. } (1 - \beta_s) \\
&\leq \frac{b * (|\tilde{v}_{\tilde{c}}| + \alpha_c) + \alpha_s}{|\tilde{v}_{\tilde{c}}|} * \left(\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} + \frac{\alpha_s}{a * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_s}\right) = \alpha^*_{\text{avg}} \\
&\quad \text{; by upper bound of } c \text{ with prob. } (1 - \beta_c)
\end{aligned}
$$

where we obtain the error for the average α^*_{avg} which does not depend on the result of the (to be performed) noisy sum but rather the (already performed) noisy count.

Proposition 3.6. *Given $0 < \gamma < 1$, the privacy parameters (ε_c, δ) for a differentially private count with error bound computed as $\alpha_c = icdf(1, \varepsilon_c, \delta, \beta)$, $\tilde{v}_{\tilde{c}} \leftarrow \tilde{c}$, the privacy parameters (ε_s, δ) for a differentially private sum with error bound computed as $\alpha_s = icdf(\Delta_s, \varepsilon_s, \delta, \beta)$, where $\Delta_s = \max\{a, b\}$, $a > 0$, a and b being the lower and upper bounds of the values being added, respectively, and*

- $\alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma$
- $\frac{b}{\varepsilon_s} \leq \frac{a}{\varepsilon_c} * \left(\frac{1-\gamma}{1+\gamma}\right)$

then it holds

$$\Pr\left[\left|\frac{\tilde{s}}{\tilde{v}_{\tilde{c}}} - \frac{s}{c}\right| > \alpha^*_{avg}\right] \leq 2\beta$$

where

$$\alpha^*_{avg} = \frac{b * (|\tilde{v}_{\tilde{c}}| + \alpha_c) + \alpha_s}{|\tilde{v}_{\tilde{c}}|} * \left(\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} + \frac{\alpha_s}{a * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_s}\right)$$

From this result, we can derive several observations. First, that both reducing α_c and α_s or a big noisy count minimizes the overall error bound α^*_{avg}. Second, smaller ranges a and b also improve error estimation, highlighting the importance of tight data bounds during analysis. Third, allocating privacy budgets effectively between the noisy count and the sum is crucial. Lastly, when real counts c are significantly larger than α_c/γ, it is likely that $|\tilde{v}_{\tilde{c}}| \geq \alpha_c/\gamma$, which implies $\alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma$—which allows us to provide accuracy bounds most of the time when the constraints around the epsilons are satisfied. Extending the proposition above to consider non-positive elements would require case-specific analyses to derive the corresponding preconditions, which might not end up in a compact, elegant formulation that is independent of the DP mechanism—thus we leave it as future work.

In our empirical evaluation, we took $\gamma = 0.1$, corresponding to a 10% threshold on fractional uncertainties, as a reasonable and intuitive baseline for determining when uncertainty propagation yields error bounds. However, this choice is not fundamental, but it can affect the applicability of our approach. We further explore the implications of varying γ in Section 5.

3.2 Accuracy for quotients of DP-sums

Using analogous steps for estimating accuracy for averages, in this section we show that we can approximate the accuracy of quotients computed from two DP-sums. Concretely, let $\tilde{s}_1$ and $\tilde{s}_2$ be the DP-sums of two different queries over the same dataset or two datasets with the same number of records c. Let $\tilde{s}_1$ be computed over values in the range $[a_1, b_1]$ and $\tilde{s}_2$ over values in the range $[a_2, b_2]$, where $a_i > 0$. Then, the noisy ratio of these two sums defined as $\tilde{s}_1/\tilde{s}_2$ has an error bound $\alpha_{\tilde{s}_1/\tilde{s}_2}$ that can be formulated in function of a noisy count.

Given a comparison to the average scenario, the following proposition for quotients of DP-sums is simply derived by imposing further constraints to the parameters of $\tilde{s}_2$ in the same way as we did for the privacy parameters of $\tilde{s}_1$.

Proposition 3.7. *Given $0 < \gamma < 1$, the privacy parameters $(\varepsilon_c, \delta_c)$ for a differentially private count with error bound computed as $\alpha_c = icdf(1, \varepsilon_c, \delta, \beta)$, $\tilde{v}_{\tilde{c}} \leftarrow \tilde{c}$, $i \in \{1, 2\}$, the privacy parameters $(\varepsilon_{s_i}, \delta)$ for a differentially private sum with error bound computed as $\alpha_{s_i} = icdf(\Delta_{s_i}, \varepsilon_{s_i}, \delta, \beta)$, where $\Delta_{s_i} = \max\{a_i, b_i\}$, $a_i > 0$, b_i being the lower and upper bounds of the values in the dataset, respectively, and*

- $\alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma$
- $\frac{b_i}{\varepsilon_{s_i}} \leq \frac{a_i}{\varepsilon_c} * \left(\frac{1-\gamma}{1+\gamma}\right)$,

then it holds

$$\Pr\left[\left|\frac{\tilde{s_1}}{\tilde{s_2}} - \frac{s_1}{s_2}\right| > \alpha^*_{\tilde{s_1}/\tilde{s_2}}\right] \leq 3\beta + \beta^3$$

where

$$\alpha^*_{\tilde{s_1}/\tilde{s_2}} = \frac{b_1 * (|\tilde{v}_{\tilde{c}}| + \alpha_c) + \alpha_s}{a_2 * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_{s_2}} * \left(\frac{\alpha_{s_1}}{a_1 * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_{s_1}} + \frac{\alpha_{s_2}}{a_2 * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_{s_2}}\right)$$

As with the prediction of accuracy of averages, larger dataset sizes improves error approximation, as do tight sensitivity bounds for the sums (*i.e.*, a_i, b_i). However, there is a notable distinction here. The term $b_1 * (|\tilde{v}_{\tilde{c}}| + \alpha_c) + \alpha_s / a_2 * (|\tilde{v}_{\tilde{c}}| - \alpha_c) - \alpha_{s_2}$ in $\alpha^*_{\tilde{s_1}/\tilde{s_2}}$ highlights the critical role of the ratio b_1/a_2. A large b_1 relative to a_2 can inflate the error estimation. To address this, it is beneficial to balance the scales of the two sums, e.g., by rescaling. It might also be beneficial to place the sum with the smaller b_i as the numerator when computing the quotient.

4 Tuners

The findings in Section 3 outline several prerequisites for estimating the precision of quotients, which may pose a challenge for data analysts. To alleviate this issue, we provide here a series of *tuners* that facilitate the validation of average error estimation under these conditions. The tuners are designed to work in three different modes of exploration, each of which is useful for different scenarios. For simplicity, we show the tuners for averages that can be easily extended to work on quotients of DP sums. We assume

that the lower and upper limits for the range of values of the sum lies in the positive interval $[a, b]$.

4.1 Mode I: Unknown dataset dimension

This mode is designed to help users determine the minimal number of records required to likely satisfy the precondition $\alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma$ in Proposition 3.6 and 3.7 for a desired level of privacy for the count (ε_c, δ_c). In this mode, the tuner takes as input the desired label of privacy for the count (ε_c, δ_c) and its confidence parameter β_c, then, the tuner provides the minimal number of records $c_{\min}$. Additionally, the tuner suggests the minimal privacy parameter for the sum $\varepsilon_{s_{\min}}$ such that it satisfies the precondition(s) $b/\varepsilon_{s_{\min}} \leq a/\varepsilon_c * (1-\gamma/1+\gamma)$.

Algorithm 1: Tuner Mode I

Function `modeI`$(\gamma, \varepsilon_c, \delta_c, \beta_c, a, b)$**:**

$c_{\min} \leftarrow \left\lceil \frac{\alpha_c(1+\gamma)}{\gamma} \right\rceil$ ▷ *by Prop. 3.2*;

$\varepsilon_{s_{\min}} \leftarrow \frac{b\varepsilon_c}{a} \cdot \left(\frac{1+\gamma}{1-\gamma}\right)$ ▷ *by hypothesis in Prop. 3.5*;

$\delta_s \leftarrow \delta_c$;

$\beta_s \leftarrow \beta_c$;

return $(c_{\min}, (\varepsilon_{s_{\min}}, \delta_s), \beta_s)$;

Algorithm 1 describes the tuner running in mode I. We have already seen in Proposition 3.3 that the minimal number of records $c_{\min}$ can be determined based on the desired privacy level for the count ε_c and the count's parameter β_c in such a way that the condition $\alpha_c/|\tilde{c}| \leq \gamma$ is likely to hold. To determine the minimal privacy parameter for the sum $\varepsilon_{s_{\min}}$, we elaborate on the condition $b/\varepsilon_{s_{\min}} \leq a/\varepsilon_c * (1-\gamma/1+\gamma)$, which applies for both the Laplace and Gaussian mechanisms.

4.1.1 An example. To illustrate the tuner's functionality in this mode, let's consider several values for the count's privacy parameter $\varepsilon_c \in [0.001, 0.1)$ and its confidence parameter $\beta_c \in 0.05, 0.1$, $a = 1$, $b = 10$, and $\gamma = 0.1$. We assume we will use the Laplace mechanism so $\delta_c = 0$, and we omit it in what follows. For each combination ε_c and β_c, we use the tuner to provide the minimal number of records $c_{\min}$ and the minimal privacy parameter for the sum $\varepsilon_{s_{\min}}$. Table 1 shows some of the results obtained. As we can see, the minimal number of records $c_{\min}$ decreases as the privacy level for the count ε_c and the confidence parameter β_c increases. On the other hand, the minimal privacy parameter for the sum $\varepsilon_{s_{\min}}$ increases proportionally to the privacy level for the count ε_c.

ε_c	β_c	$c_{\min}$	$\varepsilon_{s_{\min}}$
0.001	0.05	33.0K	0.01
	0.1	25.3K	0.01
0.05	0.05	661	0.61
	0.1	508	0.61
0.09	0.05	368	1.1
	0.1	283	1.1

Table 1: Example for tuner Mode I.

4.2 Mode II: Known dataset dimension

This mode is designed for when the user has already obtained information about of the number of records in the dataset by performing a DP-count and is interested in determining the privacy parameters for the sum such that the preconditions for the average's error estimation are likely valid. As such, the tuner takes as input the result of the DP-count $\tilde{v}_{\tilde{c}}$ together with the used privacy budget $(\varepsilon_c, \delta_c)$, and produces $\varepsilon_{s_{\min}}$ such that $b/\varepsilon_{s_{\min}} \leq a/\varepsilon_c * (1-\gamma/1+\gamma)$ is likely satisfied.

Algorithm 2: Tuner Mode II

Function `modeII`$(\gamma, \varepsilon_c, \delta_c, \beta_c, a, b, \tilde{v}_{\tilde{c}})$**:**

if $\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} > \gamma$ **then**

return Error: Condition not met

else

$(c_{\min}, (\varepsilon_{s_{\min}}, \delta_s), \beta_s) \leftarrow$ `modeI`$(\gamma, \varepsilon_c, \delta_c, \beta_c, a, b)$

return $((\varepsilon_{s_{\min}}, \delta_s), \beta_s)$

Algorithm 2 describes the tuner running in mode II. This mode primarily ensures that the fractional uncertainty of the noisy count is minimal, triggering a transition to the tuner in mode I while discarding the reported minimum number of records.

4.2.1 An example. Similar as we did before, to illustrate the tuner's functionality in this mode, let us consider several values for the count's privacy parameter $\varepsilon_c \in \{0.001, 0.05\}$ and its confidence parameter $\beta_c = 0.05$, $a = 1$, $b = 100$, and $\gamma = 0.1$. We consider datasets with real counts $c \in \{10.000, 100.000, 1.000.000\}$ records. We assume we will use the Laplace mechanism so $\delta_c = 0$, and we omit it in what follows. Table 2 shows the results obtained.

ε_c	c	$\tilde{v}_{\tilde{c}}$	$\varepsilon_{s_{\min}}$
0.001	10K	8899	unsat cond
	100K	99937	0.12
0.05	10K	9971	6.11
	100K	99979	6.11
0.09	10K	9999	11
	100K	100004	11

Table 2: Example for tuner Mode II.

The table indicates that, under the given privacy parameters, the dataset with a real count of $10K$ records are not enough to be able to provide error estimations when the epsilon is too small, e.g., $\varepsilon_c = 0.001$, $c = 10K$, and $\tilde{v}_{\tilde{c}} = 8899$. However, average accuracy estimation are likely possible with datasets with real counts are higher (e.g., $100K$ and $1M$) or the privacy parameter epsilon for the noisy count is bigger.

The table also shows that the more budget we spend on the noisy count, the higher the budget that we need for the noisy sum—some ε_s are higher than those recommend by good practices, but they have been selected for making this point clear. This behavior comes from the requirement $\frac{b}{\varepsilon_s} \leq \frac{a}{\varepsilon_c} * \left(\frac{1-\gamma}{1+\gamma}\right)$. Observe that the higher the ε_c, the smaller $\frac{a}{\varepsilon_c} * \left(\frac{1-\gamma}{1+\gamma}\right)$, hence ε_s needs to increase to make $\frac{b}{\varepsilon_s}$ smaller and satisfy the inequality. Alternatively, the dataset can be manipulated so that the lower bound of the sum a gets increased so that the inequality holds with a possible small budget requirements for the sum.

4.3 Mode III: Budget optimization

This mode is designed to help users explore different privacy allocation of the budget for both the DP-count and DP-sum so that the preconditions to report the accuracy of the average are likely to be satisfied. This mode deviates from its predecessors by requiring users *to estimate or forecast the total number of records within the dataset*—this requirement aligns with the methodology suggested by OpenDP when operating under an unbounded Differential Privacy model.

In this mode, the tuner takes a range of maximal values for average's privacy level $\mathcal{E}_{\text{avg}} = \{(\varepsilon_{\text{avg}_1}, \delta_{\text{avg}_1}), (\varepsilon_{\text{avg}_2}, \delta_{\text{avg}_2}), \ldots, (\varepsilon_{\text{avg}_n}, \delta_{\text{avg}_n})\}$ as well as a range of estimates for the number of records $\hat{C} = \{\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_m\}$. The tuner then provides the minimal privacy parameters for the count $\varepsilon_{c_{\text{min}}}$ and the sum $\varepsilon_{s_{\text{min}}}$ for each combination $((\varepsilon_{\text{avg}_i}, \delta_{\text{avg}_i}), \hat{c}_j)$ with $1 \le i \le n; 1 \le j \le m$ such that the preconditions for the average's error estimation are likely to be satisfied provided that the budget $\varepsilon_{\text{avg}_i}$ is enough. Algorithm 3

Algorithm 3: Tuner Mode III (Laplace mechanism)

Function `oneStep`$(\gamma,\ \hat{c},\ \varepsilon_{avg_{max}},\ \beta,\ a,\ b)$**:**

$\varepsilon_{c_{\text{min}}} \leftarrow \dfrac{\log\left(\frac{1}{\beta/2}\right)(1+\gamma)}{\hat{c}\cdot\gamma}$ ▷ *by $\alpha_c/(c-\alpha_c) \le \gamma$ in Prop. 3.3*

$(c_{\text{min}}, (\varepsilon_{s_{\text{min}}}, \delta_s), \beta_s) \leftarrow$ `modeI`$(\gamma,\ \varepsilon_{c_{min}},\ 0,\ \beta/2,\ a,\ b)$

$\varepsilon_{\text{avg}_{\text{min}}} \leftarrow \varepsilon_{c_{\text{min}}} + \varepsilon_{s_{\text{min}}}$

if $\varepsilon_{avg_{min}} > \varepsilon_{avg_{max}}$ **then**

  return $\{\}$ ▷ *not enough budget*

return $\{(\varepsilon_{c_{\text{min}}},\ \varepsilon_{s_{\text{min}}},\ \varepsilon_{\text{avg}_{\text{min}}})\}$

Function `modeIII`$(\gamma,\ \hat{C},\ \mathcal{E}_{avg},\ \beta,\ a,\ b)$**:**

$R \leftarrow \emptyset$

foreach $(\hat{c}, \varepsilon_{avg}) \in \hat{C} \times \mathcal{E}_{avg}$ **do**

  $R \leftarrow R \cup$ `oneStep`$(\gamma,\ \hat{c},\ \varepsilon_{avg},\ \beta,\ a,\ b)$

return R

defines the tuner for the Laplace mechanism. Different from Algorithm 1 and 2, the code for the tuner is mechanism-specific when it comes to calculating $\varepsilon_{c_{\text{min}}}$. The algorithm performs the Cartesian product of the privacy budgets proposed for the average $\mathcal{E}$ and the predicted sizes of the dataset $\hat{C}$. For each element in this product, the tuner computes the minimum epsilon for the counter and then uses that to call the tuner in Mode II to obtain the minimum epsilon for the sum.

For reasons of space, we do not present this mode for the Gaussian mechanism. It is very similar to Algorithm 3 excepts that it needs to account for the δ. In short, given a desired δ_{avg}, it splits it in two and uses $\delta_{\text{avg}}/2$ when calling Algorithm 1 rather than 0.

While our methodology computes each $(\varepsilon_{\text{avg}}, \delta_{\text{avg}})$ configuration independently, we allow the tuner to take a list $\mathcal{E}_{\text{avg}}$ of such tuples to facilitate *tabular exploration* of how different privacy budgets interact with varying record count estimates. This design choice supports practical scenarios where analysts wish to compare trade-offs across multiple budget configurations at once—something particularly useful during the planning or tuning phase of a DP analysis. For users interested in a single configuration, our formulation still accommodates this by passing a singleton set—thus without adding complexity for simpler use cases.

4.3.1 An example. To illustrate the aid that the tuner can provide, we present an example where we consider the Laplace mechanism, $\beta_{\text{th}} = 0.05$, $a = 1$, $b = 10$, $\gamma = 0.1$, $\varepsilon_{\text{avg}_{\text{max}}} \in \{0.041, 0.081\}$, and $\hat{c} \in \{10K, 100K, 1M\}$. For each combination of $\varepsilon_{\text{avg}_{\text{max}}}$ and $\hat{c}$, the tuner suggests values for the count and sum privacy parameters, $\varepsilon_{c_{\text{min}}}$ and $\varepsilon_{s_{\text{min}}}$, following the procedure described in Algorithm 3:

$\varepsilon_{\text{avg}_{\text{max}}}$	$\hat{c}$	$\varepsilon_{c_{\text{min}}}$	$\varepsilon_{s_{\text{min}}}$	$\varepsilon_{\text{avg}_{\text{min}}}$	Satisfied conditions
0.041	10K	4.06e-03	4.96e-02	5.37e-02	False
	100K	4.10e-04	4.96e-03	5.37e-03	True
	1M	4.10e-05	4.97e-04	5.38e-04	True
0.081	10K	4.10e-03	4.96e-02	5.37e-02	True
	100K	4.10e-04	4.96e-03	5.37e-03	True
	1M	4.10e-04	4.97e-03	5.38e-03	True

Table 3: Example for tuner Mode III.

Each row demonstrates how, for a given combination of $\varepsilon_{\text{avg}_{\text{max}}}$ and $\hat{c}$, the tuner suggests values for $\varepsilon_{c_{\text{min}}}$ and $\varepsilon_{s_{\text{min}}}$, and whether the conditions for the privacy protection are satisfied. In particular, we can see that for $\varepsilon_{\text{avg}_{\text{max}}} = 0.041$ and $\hat{c} = 10K$ the conditions are not satisfied since the minimal privacy parameter for the average exceeds the maximum allowed value—see the extended version of this work for a more detailed exploration of parameters when using this tuner.

Overall, the tuner proves to be a valuable tool for practitioners, offering a clear pathway for setting privacy parameters in the error estimation of averages. However, it is crucial to remember that the tuner is not a one-size-fits-all solution, and users must take responsibility for ensuring that any estimations or assumptions made are aligned with the actual characteristics of the dataset.

5 Evaluation

To assess the utility of the proposed error bound for the quotient of random variables, we conducted a series of experiments. These experiments focused on validating the assumptions and theoretical predictions under varying conditions, with particular emphasis on the correctness and applicability of the derived bounds.

5.1 Fractional uncertainties in a DP setting

Firstly, we are interested in verifying the soundness of the error estimation by fractional uncertainties for differentially private averages as proposed in equation (1). That is, if we have a noisy count $\tilde{v}_{\tilde{c}} \leftarrow \hat{c}$ and a noisy sum $\tilde{v}_{\tilde{s}} \leftarrow \hat{s}$ with *small* fractional uncertainties, does it hold that the error of the average can be approximated by $|\tilde{v}_{\tilde{s}}/\tilde{v}_{\tilde{c}}| * (\alpha_c/|\tilde{v}_{\tilde{c}}| + \alpha_s/|\tilde{v}_{\tilde{s}}|)$ with certain confidence? To answer that question involves comparing the theoretical error bounds—as provided by fractional uncertainties—with empirical observations. The goal is to determine whether the theoretical error bounds provide an upper limit on the actual error observed in practice when the preconditions that ensure *small* fractional uncertainties are satisfied—recall Proposition 3.5.

In our evaluation, we compare the theoretical error bound with the empirical error computed from the noisy average. Specifically,

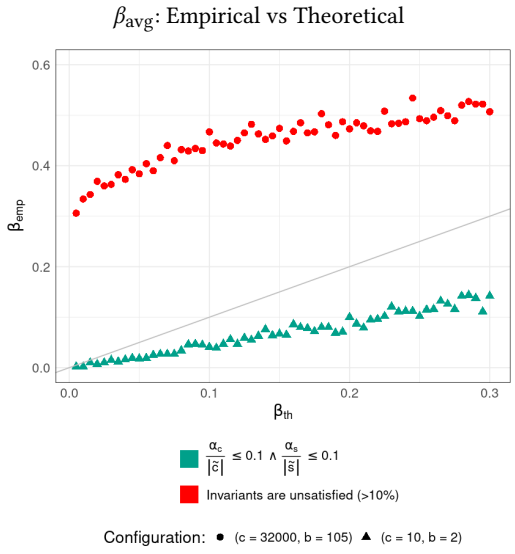


Figure 1: Soundness of error estimation for averages with $\varepsilon_c = 0.02$ and $\varepsilon_s = 0.8$

for a given dataset with c records, and privacy parameters ε_c and ε_s (for the count and sum under pure-DP, respectively), we take the confidence parameter β, called β_{th}, to vary between the range $[0.005, 0.3]$. For each value of β_{th}, we compute several DP-counts and DP-sums, together with their corresponding error bounds α_c and α_s with $\beta_c = \beta_s = \beta_{\text{th}}/2$, and then calculate the empirical and theoretical values for α_{avg}. Concretely, the empirical error is defined as the absolute value of the difference between the noisy average and the true average, while the theoretical error is determined by the uncertainty propagation formula from Definition 3.1.

$$\alpha_{\text{avg-emp}} = \left| \frac{\tilde{v}_{\tilde{s}}}{\tilde{v}_{\tilde{c}}} - \frac{s}{c} \right|$$

$$\alpha_{\text{avg-th}} = \left| \frac{\tilde{v}_{\tilde{s}}}{\tilde{v}_{\tilde{c}}} \right| * \left(\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} + \frac{\alpha_s}{|\tilde{v}_{\tilde{s}}|} \right)$$

Moreover, for each pair of sampled $(\tilde{v}_{\tilde{c}}, \tilde{v}_{\tilde{s}})$, our evaluation also checks whether the conditions from Proposition 3.5 are satisfied and provide validity percentages of all the samples defined as:

$$\text{valid}_c = \frac{\text{\# of noisy count samples where } \alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma}{\text{\# of noisy count samples}}$$

$$\text{valid}_s = \frac{\text{\# of noisy sum samples where } \alpha_s/|\tilde{v}_{\tilde{s}}| \leq \gamma}{\text{\# of noisy sum samples}}$$

We can then calculate an empirical value for β, called β_{emp}, by checking the proportion of empirical errors $\alpha_{\text{avg-emp}}$ that are above the theoretical bound computed using the non-parametric formula $\alpha_{\text{avg-th}}$. A correct estimation of the error bound by fractional uncertainties should yield $\beta_{\text{emp}} \leq \beta_{\text{th}}$, indicating that the error estimation is indeed an upper bound of the real error.

Figure 1 shows the result of our evaluation. Concretely, the experiments are conducted with $a = 1$, $\gamma = 0.1$, $\varepsilon_c = 0.02$, and $\varepsilon_s = 0.8$ while choosing different configurations for the count c and data's range upper limit b. Points marked with ● represent the cases where the configuration is set as $c = 32000, b = 105$, while those marked with ▲ represent the cases where $c = 10, b = 2$. The points filled with green represent the values of β_{emp} obtained from 1000 samples, where the conditions necessary for achieving small fractional uncertainties are typically met. Specifically, this occurs when at least 90% of the samples fulfill the required preconditions. In contrast, the points filled with red indicate that at least a 10% of the samples do not satisfy the preconditions. The diagonal line represents the boundary between the theoretical and empirical βs. A more detailed depiction of some of the results obtained in this evaluation can be found in Table 4. Importantly, it includes a more granular overview of the validity percentages for each precondition serving as indicators of the overall reliability of the error estimation.

c	a	b	β_{th}	β_{emp}	valid_c	valid_s
3200	1	105	0.005	0.004	100%	100%
			0.05	0.019	100%	100%
			0.3	0.131	100%	100%
10	1	2	0.005	0.331	0%	0%
			0.05	0.404	0%	0%
			0.3	0.526	0%	0%

Table 4: Results for the experiments in Fig. 1

As evident in these results, the configuration with a larger data size (32000) always yields a correct error estimation as the values of β_{emp} remain lower than those of β_{th}, i.e., below the diagonal line. Interestingly, the validity percentages for both conditions are 100%, indicating that the preconditions are always satisfied. On the other hand, the configuration with a smaller data size (10) does not satisfy the preconditions—even though the data variability is significantly lower—as indicated by the assigned color and the validity percentages are 0% on each condition. Furthermore, the error estimation is incorrect since we see that the empirical error always surpasses the empirical bound, i.e., it is above the diagonal line.

This experiment demonstrates that indeed the proposed error estimation using propagation of uncertainties is an upper bound for the empirical error when the preconditions are satisfied. Moreover, the experiment also underscores that the preconditions are necessary to ensure the applicability of the error estimation by using fractional uncertainties.

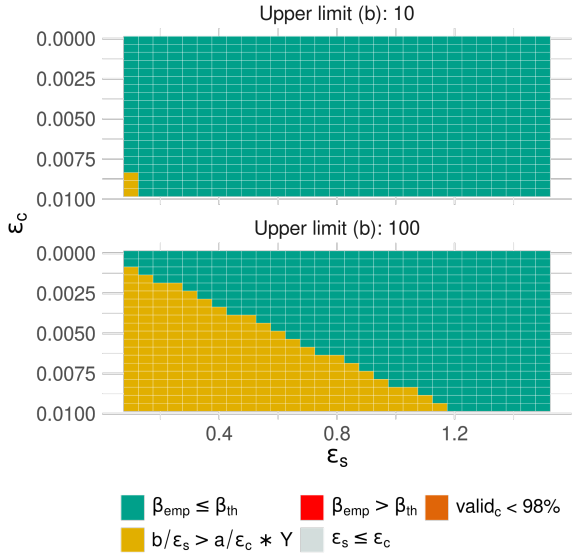
5.2 Error estimations

As show in Section 3, the average's error bound can be approximated by applying fractional uncertainties with a noisy count and theoretical error bounds (i.e., α_c and α_s) for the count and the sum. In what follows, we check that the error estimation described in Proposition 3.6 is a sound approximation of the empirical error.

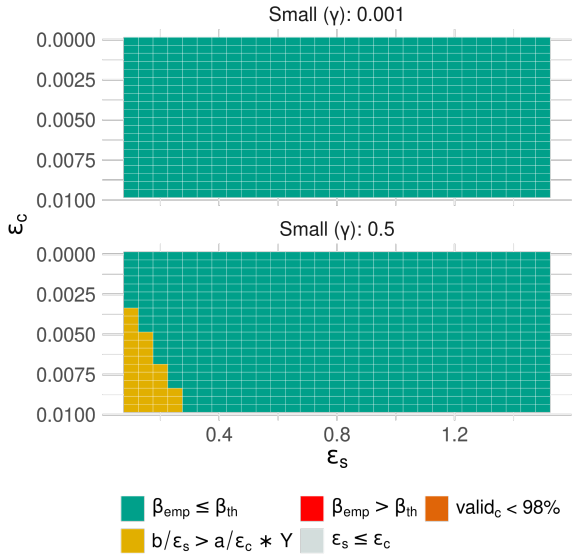
To do so, we take several values for the privacy parameters ε_c and ε_s, importantly, we only consider the cases where $\varepsilon_s > \varepsilon_c$ since we need to satisfy the inequality $b/\varepsilon_s \leq a/\varepsilon_c * (1-\gamma/1+\gamma)$.

For each pair of privacy parameters, we compute the minimal number of records c_{min} required to satisfy that the fractional uncertainty of the noisy count is *small*—recall Proposition 3.2. We then generate a synthetic dataset with c_{min} records with values sampled uniformly within the range $[a, b]$. We compute the real

$$\frac{\alpha_c}{|\tilde{v}_{\tilde{c}}|} \leq \gamma \wedge \frac{b}{\varepsilon_s} \leq \frac{a}{\varepsilon_c} * \left(\frac{1-\gamma}{1+\gamma}\right) \Rightarrow \Pr\left[|\tilde{\text{avg}} - \text{avg}| > \alpha^*_{\text{avg}}\right] \leq \beta_{\text{th}}$$



(a) Varying data's upper limit (b)



(b) Varying the "small" threshold (γ)

Figure 2: Soundness of α^*_{avg} with $a = 1$, $\beta_{\text{th}} = 0.05$

and noisy averages under the Laplace Mechanism for this dataset, using the corresponding counts and sums. Then, we compare the theoretical error bound, α^*_{avg}, with the empirical error. As with the previous experiment, when analyzing the proportion of empirical errors surpassing the theoretical error bound, we expect to have $\beta_{\text{emp}} \leq \beta_{\text{th}}$.

In Figure 2a we show the results of the experiments for the average's error estimation α^*_{avg}. Concretely, the first matrix depicts

b	ε_c	ε_s	$c_{\min}$	α_c	α_s	valid$_c$	valid$_s$
10	0.004	0.20	9898	900	184	98%	True
		0.14			26	99%	True
	0.009	0.20	4461	405	184	99%	True
		0.14			26	100%	True
100	0.004	0.20	9898	900	1844	99%	False
		0.14			263	99%	True
	0.009	0.20	4461	405	1844	98%	False
		0.14			263	99%	True

Table 5: Results for the experiments in Fig. 2a

the case where $a = 1$, $b = 10$, $\gamma = 0.1$, and $\beta_{\text{th}} = 0.05$. In this case, we can see that all cases satisfy the first condition $\alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma$ with high probability, as the validity percent is always above 98% (i.e., there are no cells marked in orange). However, this is not the case for the second condition, as there are some cells marked in yellow indicating that $b/\varepsilon_s > a/\varepsilon_c * Y$ where $Y = (1-\gamma/1+\gamma)$ and consequently rendering the error estimation invalid for those $(\varepsilon_s, \varepsilon_c)$ pairings.

These results highlight the importance of correctly distributing the privacy budget between the count and the sum to ensure the applicability of the error estimation. Observe that for those cases where the conditions hold (i.e., the cell is not marked in grey, yellow, or orange), the empirical error remains below the theoretical one as desired, marked in green, suggesting that the conditions are a sufficient constraint to ensure the correctness of the error estimation.

The second matrix in Figure 2a shows the case where the range of data is increased to $b = 100$, the effect of this change is that there are more cases where the second precondition $b/\varepsilon_s \leq a/\varepsilon_c * Y$ is not satisfied, leading to a higher proportion of cells marked in yellow, this is because sum's sensitivity is determined by the upper bound b, i.e., the bigger b, the bigger ε_s should be in order to reduce b/ε_s. However, for those cases where the conditions hold, the error estimation remains correct.

Table 5 highlights key results from these experiments, illustrating the interplay between the privacy parameters ε_c and ε_s and the dataset range b in meeting the conditions required for accurate error estimation of the average. For $b = 10$, the conditions are satisfied for most parameter combinations, with validity percentages consistently exceeding 98%, and always satisfying the inequality $b/\varepsilon_s \leq a/\varepsilon_c * Y$ for both $\varepsilon_c = 0.004$ and $\varepsilon_c = 0.009$. However, when the dataset range increases to $b = 100$, the validity of the second condition declines, as indicated by the higher number of invalid cases (marked as False). This decline occurs because a larger range increases the sum's sensitivity, requiring a proportionally higher privacy budget ε_s to satisfy the condition. These findings underscore the critical importance of distributing the privacy budget effectively between the count and the sum, particularly for datasets with wider ranges, to ensure the applicability of the error estimation.

Lastly, we aim to investigate the effect of the parameter γ on the soundness of the error estimation. To this end, we performed similar experiments as before but where the value of γ was varied among 0.001, 0.5, while keeping the other parameters fixed as follows: $a = 1$, $b = 10$, and $\beta_{\text{th}} = 0.05$. The results of this evaluation

c	mean	standard deviation	β_{emp}	valid$_c$	valid$_s$
1K	8.38e-02	2.04e-04	1e-04	100%	100%
10K	7.93e-03	1.75e-06	2e-04	100%	100%
100K	7.86e-04	1.76e-08	2e-04	100%	100%
1M	7.86e-05	1.73e-10	5e-04	100%	100%
10M	7.86e-06	1.76e-12	1e-04	100%	100%

Table 6: Relative uncertainty ($\alpha^*_{\text{avg}}/|a\tilde{v}g|$) analysis

are presented in Figure 2b. Our findings show that the soundness of the error estimation remains unaffected by the value of γ. However, we observe that different values of γ lead to varying applicability restrictions. Specifically, increasing γ does not necessarily result in a higher number of valid combinations of $(\varepsilon_c, \varepsilon_s)$ values. This can be explained by the fact that, while increasing γ makes the first condition easier to satisfy, the second condition becomes more restrictive. As γ approaches zero, the expression $^{1-\gamma}/_{1+\gamma}$ tends to zero, which imposes stricter constraints on the values of ε_s.

5.3 Estimation tightness

The proposed error estimation for the average has been shown to be sound, and the conditions necessary for its applicability are sufficient. The remaining question is whether the error estimation is sufficiently *tight*, meaning whether the provided bounds are small enough to be practical in real-world scenarios. To assess the tightness of the error estimation, we compute the average's relative uncertainty $\alpha^*_{\text{avg}}/|a\tilde{v}g|$. A lower ratio indicates a tighter bound.

To evaluate the relative uncertainty, we revisit the example from Section 1 about calculating the average over the working-age population (see Listing 3). Our goal is to determine whether the error estimation in this context is practical. For this evaluation, we vary the dataset size while keeping the parameters fixed as follows: $\varepsilon_c = 0.1$, $\varepsilon_s = 1$, $a = 18$, $b = 65$, $\gamma = 0.1$, and $\beta_{\text{th}} = 0.05$.

For each dataset size, we conduct 10,000 iterations, calculating the relative uncertainty for each sample, and then report the mean and standard deviation across all samples. Additionally, we provide the proportion of samples that satisfy the preconditions. Table 6 summarizes the findings were several trends are evident. The relative uncertainty decreases steadily with increasing dataset size, highlighting the scalability of the method. Importantly, the relative uncertainty remains below 1 across all dataset sizes, demonstrating practical tightness even for small datasets. Additionally, the standard deviation of the uncertainty diminishes sharply for larger datasets, reflecting increased reliability. The empirical confidence (β_{emp}) remain lower than that of the theoretical one as expected, indicating that the error estimation is consistently an upper bound on the empirical error. These findings underline the robustness and practicality of the proposed estimation technique across various dataset scales and real world scenarios.

Finally, considering a concrete instance of evaluating the error in the average operation from Listing 3, let's assume we have a dataset of 1k (10k) records (after applying the age range filter) and a corresponding noisy count of $\tilde{v}_{\tilde{c}} = 998$ ($\tilde{v}_{\tilde{c}} = 10006$). In this case, we can say with 95% confidence that the error of the average working-age population is $\alpha^*_{avg} = 3.44$ ($\alpha^*_{avg} = 0.32$). Therefore, the true average of the working-age population lies within the range $[a\tilde{v}g - 3.44, a\tilde{v}g + 3.44]$ ($[a\tilde{v}g - 0.32, a\tilde{v}g + 0.32]$) with 95% confidence.

ε_c	c	s/c	$\tilde{v}_{\tilde{s}}/\tilde{v}_{\tilde{c}}$	proposed	basic
0.0001	100	52.04	-2.03	-	$(-\infty, \infty)$
	1K	50.51	1.75	-	$(-\infty, \infty)$
	10K	49.98	76.6	-	$(-\infty, \infty)$
	100K	50.49	46.84	-	[34.75, 71.53]
	1M	50.51	50.4	[44.56, 56.24]	[48.59, 52.35]
	10M	50.5	50.51	[49.95, 51.07]	[50.32, 50.7]
	100M	50.5	50.52	[50.46, 50.57]	[50.5, 50.54]
	1G	50.5	50.5	[50.49, 50.51]	[50.5, 50.5]
0.00001	100	51.77	-0.02	-	$(-\infty, \infty)$
	1K	49.97	-0.61	-	$(-\infty, \infty)$
	10K	50.85	3.8	-	$(-\infty, \infty)$
	100K	50.6	8.83	-	[5.36, 24.81]
	1M	50.49	45.59	-	[34.19, 68.39]
	10M	50.49	49.68	[45.72, 53.64]	[47.94, 51.55]
	100M	50.5	50.6	[50.21, 50.99]	[50.42, 50.79]
	1G	50.5	50.5	[50.46, 50.54]	[50.48, 50.52]

Table 7: Comparison against the basic approach with $a = 1$, $b = 100$, $\gamma = 0.1$, and $\varepsilon_s = 0.02$

5.4 Comparison with a basic approach for CI

To provide a clear picture regarding the advantages of the proposed method, we consider a basic approach for estimating the error of the average. This approach consists of deriving the confidence interval (CI) of the average using those of the noisy count and noisy sum together with the result of both queries. Specifically, given a noisy count $\tilde{v}_{\tilde{c}}$, a noisy sum $\tilde{v}_{\tilde{s}}$, and their corresponding error bounds α_c and α_s, the CI for the average can be approximated by evaluating the extreme values of the count and sum. This involves determining the boundary cases for the ratio s/c as follows:

$$
\begin{aligned}
&\min_c = \tilde{v}_{\tilde{c}} - \alpha_c \qquad \max_c = \tilde{v}_{\tilde{c}} + \alpha_c \\
&\min_s = \tilde{v}_{\tilde{s}} - \alpha_s \qquad \max_s = \tilde{v}_{\tilde{s}} + \alpha_s \\
&\min_{\text{avg}} = \min\left\{\frac{\min_s}{\min_c}, \frac{\max_s}{\min_c}, \frac{\min_s}{\max_c}, \frac{\max_s}{\max_c}\right\} \\
&\max_{\text{avg}} = \max\left\{\frac{\min_s}{\min_c}, \frac{\max_s}{\min_c}, \frac{\min_s}{\max_c}, \frac{\max_s}{\max_c}\right\} \\
&\qquad \Rightarrow \frac{s}{c} \in \left[\min_{\text{avg}}, \max_{\text{avg}}\right] = \text{basic}
\end{aligned} \tag{4}
$$

Importantly, if the interval for the count ($c \in [\min_c, \max_c]$) contains zero, the average s/c becomes unbounded, and we assume $s/c \in (-\infty, \infty)$.

To compare our error bound with this basic approach, we take fixed values for the parameters a, b, ε_s, and γ, and choose different values for the count's privacy parameter ε_c and the size of the dataset c. For each dataset size, we generate a synthetic dataset with values sampled uniformly within the range $[a, b]$. We then compute the noisy count $\tilde{v}_{\tilde{c}}$ and the noisy sum $\tilde{v}_{\tilde{s}}$, along with their corresponding error bounds α_c and α_s. Consequently, we compute the average's error α^*_{avg} using the noisy count and the sum's privacy parameter, then we determine the CI for the average as

proposed $= \left[\tilde{v}_{\text{avg}} - \alpha^*_{\text{avg}}, \tilde{v}_{\text{avg}} + \alpha^*_{\text{avg}}\right]$, where $\tilde{v}_{\text{avg}} = \tilde{v}_s/\tilde{v}_c$—it is important to note that the noisy average is used solely for constructing comparable CIs, and it is not required to provide the error estimation with our method. Lastly, we compute the basic approach's CI for the average as described in Equation 4, and we compare the two intervals.

Table 7 shows the results of this comparison for different dataset sizes c and privacy parameters ε_c with fixed values of $a = 1$, $b = 100$, $\varepsilon_s = 0.02$, and $\gamma = 0.1$. While the basic approach is conceptually simple and requires only elementary calculations, making it appealing for quick estimates, it suffers from several important limitations in practice. As shown in Table 7, the intervals it produces are often either uninformative or misleading, particularly for small dataset sizes or when the privacy budget is heavily constrained. From these results we can highlight three important observations when comparing the two methods. First, it is important to note that the basic approach requires executing both the count and sum queries, thereby fully consuming the privacy budget allocated for the average—even in cases where the resulting CI is uninformative (e.g., when it becomes unbounded). In contrast, our method provides CIs only when its preconditions are met, offering early feedback and enabling practitioners to reallocate the privacy budget more effectively, particularly toward the sum.

Second, the basic approach can yield intervals that are not valid. For instance, when $\varepsilon_c = 0.00001$ and $c = 100$K, it produces an interval of $[5.36, 24.81]$ which does not contain the true average of 50.6, leading to erroneous conclusions. While our method can be conservative—sometimes yielding no interval even when the basic approach provides a seemingly valid one (e.g., for $\varepsilon_c = 0.0001$, $c = 100$K or $\varepsilon_c = 0.00001$, $c = 1$M)—this can be seen as the price to pay for ensuring the correctness of the output.

Finally, as the dataset size increases, the CIs calculated using our approach naturally converge to those obtained with the basic method, providing tight intervals while maintaining privacy guarantees, and notably, without consuming additional privacy budget for the sum. This scalability makes our approach particularly attractive in large-scale data analysis scenarios.

These observations are consistent across different variations of the parameters. However, the applicability of our method is highly dependent on the relationship among the parameters. For instance, it cannot be employed when the privacy requirements for the count and the sum are uniform. To further investigate this dependency, we explore the applicability regions in the next set of experiments.

5.5 Applicability exploration

To assess the applicability of our proposed error estimation method, we systematically explore the parameter space defined by a, b, c, ε_c, ε_s, and γ. The goal is to identify regions where the conditions for the validity of the error estimation are met and as such understanding the practical scenarios in which the method can be effectively applied. Our approach involves iterating through various combinations of the parameters. For a given γ and a, and for each combination of $(b, c, \varepsilon_c, \varepsilon_s)$, we generate a uniformly distributed dataset of size c with values in the range $[a, b]$. Next, we compute the noisy count $\tilde{v}_{\tilde{c}}$ along with its corresponding error bound α_c, as well as the error bound α_s for the noisy sum. Using these values, we calculate the error estimation for the average, α^*_{avg}, and verify whether the conditions for validity are satisfied. Parameter combinations that meet the conditions are marked as part of the applicability region, providing a clear picture of where the error estimation method can be reliably applied.

Figure 3 illustrates the applicability regions for large-scale (i.e., $c = 1000$ and $c = 1000000$) and small-scale (i.e., $c = 100$ and $c = 300$) datasets under different privacy regimes. In these graphs, the x-axis represents the privacy parameter ε_c, while the y-axis represents ε_s. Each shaded region represents combinations of $(\varepsilon_c, \varepsilon_s)$ where the conditions are likely to hold, assuming the dataset values fall within the range $[1, b]$. As such, these regions depict the space where the error estimation is expected to be valid. We also note that the regions are not mutually exclusive: for a given value of b, all points lying above the corresponding line represent valid combinations of parameters where our method can be applied.

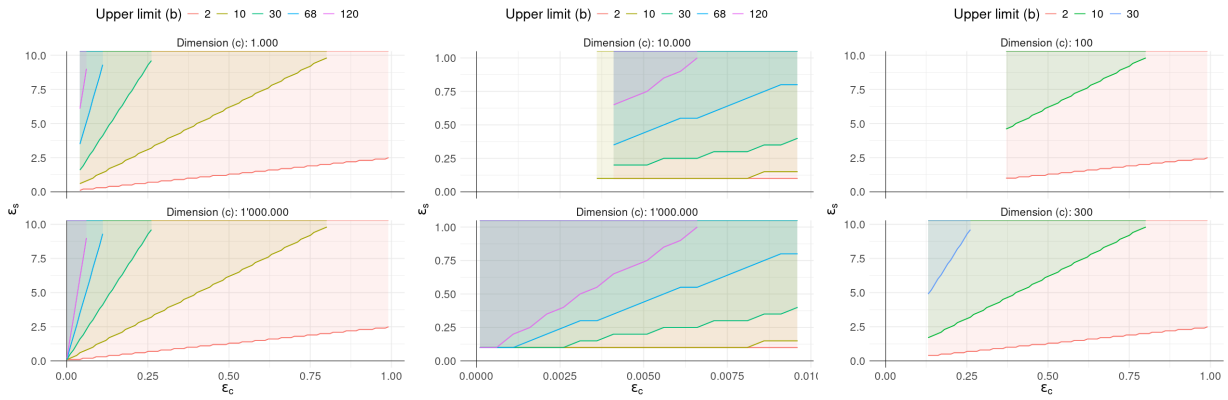
In the first scenario, depicted in Figure 3a, we examine a mixed privacy regime where the count is subject to stricter privacy requirements ($\varepsilon_c \leq 1$) compared to the sum ($\varepsilon_s > 1$). The results reveal that the applicability regions expand significantly for smaller values of b. This indicates that lower variability in the data (i.e., reduced sensitivity of the sum) allows for stronger privacy guarantees for both the sum and the count.

In the second scenario, shown in Figure 3b, we focus on a high privacy regime where both the count and the sum are subject to stricter privacy limits. This scenario provides a more detailed view of the parameter space, highlighting the interplay between different configurations. For instance, by observing the starting points of the regions for various values of b, we can see that as the dataset size decreases, stricter privacy limits are no longer suitable for both the sum and the count. Notably, the dataset size directly influences the privacy budget that can be allocated to the count. Smaller datasets make it more challenging to satisfy the precondition $\alpha_c/|\tilde{v}_{\tilde{c}}| \leq \gamma$, thereby restricting the minimal privacy target that can be enforced on the count. This is evident when comparing the starting points of the regions on the x-axis among different dataset sizes.

In the third scenario, depicted in Figure 3c, we consider a mixed privacy regime with smaller datasets ($c = 100$ and $c = 300$). Compared to previous scenarios, the applicability regions are significantly smaller, highlighting the increased difficulty in satisfying the constraints under limited data availability. This reduction in applicability regions emphasizes how limited dataset sizes affect the feasible parameter space, especially under conditions of high data variability. As dataset size decreases, the method becomes less capable of accommodating wide data ranges under the same privacy budgets.

ε_c	b	c_{min}	$\varepsilon_{s_{\text{min}}}$
0.25	2	164	0.61
	10		3.06
	30		9.17
	68		20.78
	120		36.67
0.75	2	56	1.83
	10		9.17
	30		27.50
	68		62.33
	120		110.00

Table 8: Tuner Mode I with $\gamma = 0.1$, $a = 1$



(a) Mixed privacy regimes, medium/large-scale datasets

(b) High privacy regime, medium/large-scale datasets

(c) Mixed privacy regime, small-scale datasets

Figure 3: Applicability regions with $\gamma = 0.1$, $a = 1$

To further analyze the limitations imposed by the dataset size and the data variability, we explore tuner's suggestions under this restricted scenario. Table 8 contains the tuner's recommendations for various values of b and two choices of ε_c, with fixed parameters $a = 1$ and $\gamma = 0.1$. We observe that the minimum dataset size $c_{\min}$ remains constant across all values of b for a given ε_c. This constancy does not imply a lack of interaction between dataset size and data range; rather, it reflects the tuner's internal logic: $c_{\min}$ is determined by the error bound for the count and the chosen "small" threshold. As such, the influence of b is channeled instead into adjustments of the required minimum privacy budget $\varepsilon_{s_{\min}}$.

The monotonic increase of $\varepsilon_{s_{\min}}$ with b illustrates how expanding the data range demands greater privacy resources. For example, when $\varepsilon_c = 0.25$, increasing b from 10 to 68 results in a seven-fold increase in $\varepsilon_{s_{\min}}$ (from 3.06 to 20.78). This insight explains why, in Figure 3c, no applicability regions exist for certain (c, b) combinations—specifically, when $c = 100$ and $b = 30, 68, 120$—as the dataset is too small to support the corresponding privacy requirements. Moreover, when $b = 68$ and $\varepsilon_c = 0.25$, the required $\varepsilon_{s_{\min}}$ exceeds the maximal privacy budget of 10, rendering the entire configuration inapplicable under any of the considered dataset sizes. This serves as a critical reminder: under high variability and limited data, the parameter space rapidly becomes infeasible, thus requiring a relaxation of the privacy targets.

In the final set of experiments we focus on analyzing the effect of the parameter γ on the applicability regions of the proposed error estimation method. Recall that γ serves as a threshold for the fractional uncertainty of the noisy count, influencing the conditions under which the error estimation is valid. For these experiments, the dataset size is fixed at $c = 1000000$, and a high privacy regime is considered for both the count and the sum. Two values of γ are analyzed: a strict value of $\gamma = 0.001$ and a relaxed value of $\gamma = 0.5$. The results of these experiments are illustrated in Figure 4, which shows the applicability regions for the two selected values of γ.

As expected, when γ is tightened, the applicability regions shrink due to the stricter condition $\alpha_c/|\tilde{o}_{\tilde{c}}| \leq \gamma$, which limits the range of valid combinations for the privacy parameter ε_c. Conversely,

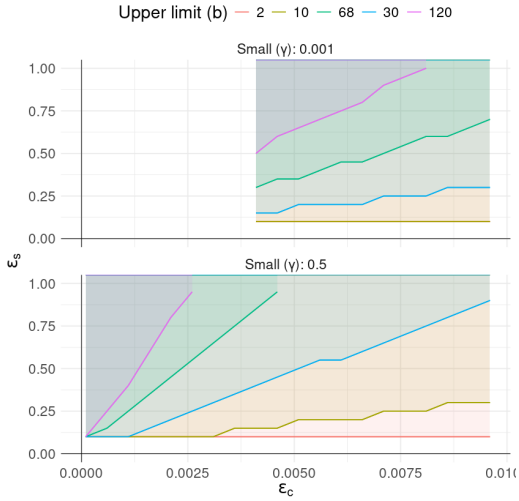


Figure 4: Applicability regions with $c = 1000000$, $a = 1$

relaxing γ expands the applicability regions by making it easier to satisfy this condition. However, this comes at a cost: the second condition, $b/\varepsilon_s \leq a/\varepsilon_c * (1-\gamma/1+\gamma)$, becomes more restrictive as γ increases, reducing the flexibility of valid ε_s values. When compared to the intermediate case of $\gamma = 0.1$ shown in Figure 3b, it becomes evident that increasing γ reduces the space of valid ε_s values as the threshold's slope becomes steeper as the value of γ increases. This trade-off highlights the delicate balance between the two conditions and the importance of carefully selecting γ based on the dataset size, privacy requirements, and data variability.

6 Related work

Statistical mean estimations. In a statistical setting, the goal is to estimate the error of approximating the population mean using finite samples (e.g, [4, 6, 8, 17, 19, 21]). We refer the reader to the

work by Kamath and Ullman for a comprehensive view on private statistical estimation [20]. More recently, PLAN [1] introduces variance-aware noise budgeting, providing tighter error bounds by allocating noise to dimensions with higher variance. Working with mean estimators in a DP setting is far from trivial. In fact, there is an inherent trade-off between bias, accuracy, and privacy of mean estimators: no algorithm can simultaneously have low bias, low error, and low privacy loss for arbitrary distributions (of the underlying data) [18]. Our work focuses on error estimations in the non-distributional, empirical setting. Unlike statistical settings, which often assume distributional properties of the data and focus on sampling-based error estimations, our approach addresses scenarios where the noise stems solely from the privacy mechanism. Statistical mean estimators have a similarity to our approach in the sense that they provide accuracy guarantees under a certain minimum number of samples while our approach does it on the noisy size of datasets.

Empirical setting. This setting exclusively concentrates on the noise generated by the privacy mechanism, rendering it particularly appropriate for systems such as SQL-based query engines. We observe that certain statistical mean confidence interval mechanisms [21] can be adapted to operate within the empirical framework. However, to achieve that, strong assumptions on the distribution of the underlying data are needed (i.e., normality). The closest work to ours is a recent-unpublished manuscript [12]. Focused on improving the accuracy of averages, Fitzsimons et al. show how two DP sums, where one is carefully crafted, can be used to derive a noisy size of the dataset without spending any budget. Authors then show how to compute the variance of both sums and the derived noisy size. However, no variance calculation is provided for the average itself. Different from our work, their technique only focuses on averages with a notion of *mean square error* for Gaussian noise, where accuracy estimations are done *empirically*. Instead, we focus on α-β accuracy with analytical error estimations for averages or quotients of DP sums for both the Laplace and Gaussian mechanisms. The work by Sun et al. shows how to provide private-preserving confidence intervals (i.e., α-β accuracy) for the Exponential mechanism [30] (EM) and the Sparse vector technique [10] (SVT). Their methodology employs additional privacy budget to compute CI bounds, relying on carefully designed utility functions for the EM and thresholds for the SVT. Authors provide CI for averages using CIs for SVT where four DP-queries need to be performed and bound DP is assumed. Interestingly, the work shows that for any DP mechanism, if the CI has a confidence level $\geq 2/3 \approx 0.66$, then the size of the interval must be at least N/n, where N is the maximum non-negative number of the elements of the sum and n is the size of the dataset—which coincides with the numbers we obtained in our evaluation. Different from them, we support unbounded DP and accuracy for quotients of DP-sums. Recently, Lin et al. address the challenge of constructing CIs for population proportions in a DP setting—which requires mathematically sophisticated techniques[25]. Our current formulations do not apply in these settings, as the quantities involved are typically not computed over the *same dataset*: the numerator corresponds to an aggregate over a filtered subset, while the denominator refers to the full population. In such cases, the relationship between numerator and denominator is no longer straightforward, making it difficult to bound the sum in terms of the noisy denominator—a limitation that presents an interesting direction for future work. We remark that our method is not intended as a one-size-fits-all solution but rather as a valuable addition to the practitioner's toolbox for computing DP quotients' accuracy under unbounded DP settings.

Accuracy in DP tools. PSI [14] provides a user interface that allows for the selection of either the desired level of accuracy or the imposed level of privacy. The error estimates provided by PSI are expressed in terms of α-β accuracy. Unfortunately, PSI only supports a restricted set of transformations and primitives, offering α-β accuracy solely at a single noisy measurement, e.g., a noisy count or a noisy sum—an approach also followed by OpenDP [13]. GUPT [32] operates under the sample-and-aggregate framework for differential privacy [35]. GUPT provides analysts with the flexibility to define either the desired accuracy of the output or the required level of privacy. However, this tool only accommodates analyses compatible with the sample-and-aggregate framework and offers only confidence intervals (i.e., α-β) estimates at the individual measurement level. APEx [15] specializes in answering three types of counting queries: WCQ (weighted counting queries), ICQ (iceberg counting queries), and TCQ (top-k counting queries). To address WCQ queries, APEx leverages the matrix mechanism [24] and uses Monte Carlo simulations to empirically derive accuracy bounds in terms of α and β. ICQ queries focus on returning aggregates of bins exceeding a specified threshold, for which APEx introduces novel data-dependent analytical accuracy bounds. For TCQ queries, a generalization of the report-noisy-max mechanism [11] is employed. APEx provides empirical accuracy guarantees for some queries and analytical guarantees for others. DPella [26, 27] emphasizes providing accuracy guarantees for queries alongside their privacy protections. Unlike many other DP libraries, DPella integrates α-β accuracy bounds into its query system and provides support to about the compositional accuracy of complex queries involving multiple DP mechanisms. Our work could complement and extend the mentioned tools with analytical methods for computing α-β accuracy bounds for averages and quotients of DP-sums.

Ratios of distributions. Understanding the distributions of quotients of random variables is a problem known by mathematicians for its complexity and analytical difficulty. Marsaglia provided a closed-formula for the CDF function $F(t)$ and density function $f(t)$ for the ratio $(a+x)/b+y$, where a and b are positive constants and x and y are independent standard normal variables [28]. By running some simulations, he shows that sometimes the resulting distribution is unimodal (i.e., one peak) or bimodal (i.e., two peaks). Forty years later, Marsaglia complemented that work by showing how to transform any ratio of normal variables w/z into the form $(a+x)/b+y$ as well as conditions for a and b to predict if the resulting distribution can be approximated by a normal distribution or is a unimodal or bimodal one [29]. Deriving an iCDF for error estimation under such conditions is inherently non-trivial due to the intricate and often non-symmetric nature of the resulting distribution. Broda and Kan study ratio distributions by performing saddlepoint approximations [7] of the density distribution function [5]. Although saddlepoint approximation offers a mathematical tool, it introduces significant complexity. Unlike these approaches, our method avoids

the need to understand or approximate the shape of the ratio distribution. This provides significant simplicity, and we hope it makes our approach more practical for real-world applications.

7 Conclusions

By connecting α-β accuracy concepts with uncertainty propagation techniques, we derived conditions under which accurate error bounds can be established, highlighting the interplay between dataset size, sensitivity bounds, and the inherent uncertainty introduced by DP mechanisms. The work contributes novel insights into the propagation of uncertainty in DP settings and enables DP tools to provide accuracy guarantees for averages and quotients of DP-sums under specific conditions—an aspect currently missing in many existing frameworks. It would be valuable to explore *specific* uncertainty propagation formulas for Gaussian distributions[1]. A key challenge lies in determining how accurately tangent-plane approximations[2] can capture functions like quotients in the context of DP.

Acknowledgments. We thank the anonymous reviewers for their feedback. This work was supported by VINNOVA Swedish Innovation Agency, Vetenskapsrådet, and the NSF awards CNS 204024.

References

[1] Martin Aumüller, Christian Lebeda, Boel Nelson, and Rasmus Pagh. 2024. PLAN: Variance-Aware Private Mean Estimation. *Proc. on Privacy Enhancing Technologies* 2024 (07 2024), 606–625.

[2] Borja Balle and Yu-Xiang Wang. 2018. Improving the Gaussian Mechanism for Differential Privacy: Analytical Calibration and Optimal Denoising. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*. 403–412.

[3] Skye Berghel, Philip Bohannon, Damien Desfontaines, Charles Estes, Sam Haney, Luke Hartman, Michael Hay, Ashwin Machanavajjhala, Tom Magerlein, Gerome Miklau, et al. 2022. Tumult analytics: a robust, easy-to-use, scalable, and expressive framework for differential privacy. *arXiv:2212.04133* (2022).

[4] Sourav Biswas, Yihe Dong, Gautam Kamath, and Jonathan Ullman. 2020. CoinPress: Practical Private Mean and Covariance Estimation. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc.

[5] Simon A. Broda and Raymond Kan. 2015. On distributions of ratios. *Biometrika* 103, 1 (12 2015), 205–218.

[6] Mark Bun and Thomas Steinke. 2019. Average-Case Averages: Private Algorithms for Smooth Sensitivity and Mean Estimation. In *Advances in Neural Information Processing Systems*, Vol. 32. Curran Associates, Inc.

[7] H. E. Daniels. 1954. Saddlepoint Approximations in Statistics. *The Annals of Mathematical Statistics* 25, 4 (1954).

[8] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. 2013. Local Privacy and Statistical Minimax Rates. In *IEEE Annual Symposium on Foundations of Computer Science*.

[9] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating Noise to Sensitivity in Private Data Analysis. In *Proceedings of the Third Conference on Theory of Cryptography* (New York, NY) *(TCC'06)*. 265–284.

[10] Cynthia Dwork, Moni Naor, Omer Reingold, Guy N. Rothblum, and Salil Vadhan. 2009. On the complexity of differentially private data release: efficient algorithms and hardness results. In *Proc. of Symp. on Theory of Computing (STOC '09)*. ACM.

[11] Cynthia Dwork, Aaron Roth, et al. 2014. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science* 9, 3–4 (2014), 211–407.

[12] Jack Fitzsimons, James Honaker, Michael Shoemate, and Vikrant Singhal. 2024. Private Means and the Curious Incident of the Free Lunch. arXiv:2408.10438 [cs.CR] https://arxiv.org/abs/2408.10438

[13] Marco Gaboardi, Michael Hay, and Salil Vadhan. 2020. A programming framework for OpenDP. *Manuscript, May* (2020).

[1] where uncertainties are calculated as the square root of the sum in quadrature of the fractional uncertainties, e.g., $\frac{\delta_{\tilde{x}/\tilde{y}}}{|\tilde{x}/\tilde{y}|} = \sqrt{(\delta_x/|\tilde{x}|)^2 + (\delta_y/|\tilde{y}|)^2}$

[2] Given f differentiable, and taking two noisy measurements $\tilde{x}$, and $\tilde{y}$, then the equation of the tangent plane is given by $f(x, y) = f(\tilde{x}, \tilde{y}) + f_x(\tilde{x}, \tilde{y})(x - \tilde{x}) + f_y(\tilde{x}, \tilde{y})(y - \tilde{y})$. The challenge is to understand how far the real value $f(x, y)$ is from $f(\tilde{x}, \tilde{y})$.

[14] Marco Gaboardi, James Honaker, Gary King, Jack Murtagh, Kobbi Nissim, Jonathan Ullman, and Salil Vadhan. 2016. Psi (Ψ): a private data sharing interface. *arXiv preprint arXiv:1609.04340* (2016).

[15] Chang Ge, Xi He, Ihab F Ilyas, and Ashwin Machanavajjhala. 2019. Apex: Accuracy-aware differentially private data exploration. In *Proceedings of the 2019 International Conference on Management of Data*. 177–194.

[16] Naoise Holohan, Stefano Braghin, Pól Mac Aonghusa, and Killian Levacher. 2019. Diffprivlib: The IBM Differential Privacy Library. arXiv:1907.02444

[17] Gautam Kamath, Jerry Li, Vikrant Singhal, and Jonathan Ullman. 2019. Privately Learning High-Dimensional Distributions. In *Proceedings of the Thirty-Second Conference on Learning Theory (Proc. of Machine Learning Research, Vol. 99)*, Alina Beygelzimer and Daniel Hsu (Eds.). PMLR, 1853–1902.

[18] Gautam Kamath, Argyris Mouzakis, Matthew Regehr, Vikrant Singhal, Thomas Steinke, and Jonathan Ullman. 2024. A Bias-Accuracy-Privacy Trilemma for Statistical Estimation. *J. Amer. Statist. Assoc.* 0, ja (2024), 1–23.

[19] Gautam Kamath, Vikrant Singhal, and Jonathan Ullman. 2020. Private Mean Estimation of Heavy-Tailed Distributions. In *Proceedings of Thirty Third Conference on Learning Theory (Proceedings of Machine Learning Research, Vol. 125)*, Jacob Abernethy and Shivani Agarwal (Eds.). PMLR.

[20] Gautam Kamath and Jonathan Ullman. 2020. A Primer on Private Statistics. arXiv:2005.00010 [stat.ML] https://arxiv.org/abs/2005.00010

[21] Vishesh Karwa and Salil Vadhan. 2018. Finite Sample Differentially Private Confidence Intervals. In *Theoretical Computer Science Conference (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 94)*, Anna R. Karlin (Ed.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[22] Daniel Kifer and Ashwin Machanavajjhala. 2011. No free lunch in data privacy. In *Proc. of International Conference on Management of Data (SIGMOD '11)*. ACM.

[23] Ios Kotsogiannis, Yuchao Tao, Xi He, Maryam Fanaeepour, Ashwin Machanavajjhala, Michael Hay, and Gerome Miklau. 2019. PrivateSQL: A Differentially Private SQL Query Engine. *Proc. VLDB Endow.* 12, 11 (July 2019), 1371–1384.

[24] Chao Li, Michael Hay, Vibhor Rastogi, Gerome Miklau, and Andrew McGregor. 2010. Optimizing linear counting queries under differential privacy. In *Proc. of ACM SIGMOD Symposium on Principles of Database Systems (PODS '10)*. ACM.

[25] Shurong Lin, Mark Bun, Marco Gaboardi, Eric D Kolaczyk, and Adam Smith. 2024. Differentially private confidence intervals for proportions under stratified random sampling. *Electronic Journal of Statistics* 18, 1 (2024), 1455–1494.

[26] Elisabet Lobo-Vesga, Alejandro Russo, and Marco Gaboardi. 2020. A programming framework for differential privacy with accuracy concentration bounds. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 411–428.

[27] Elisabet Lobo-Vesga, Alejandro Russo, and Marco Gaboardi. 2021. A Programming Language for Data Privacy with Accuracy Estimations. *ACM Trans. Program. Lang. Syst.* 43, 2 (June 2021).

[28] George Marsaglia. 1965. Ratios of Normal Variables and Ratios of Sums of Uniform Variables. *J. Amer. Statist. Assoc.* 60, 309 (1965), 193–204.

[29] George Marsaglia. 2006. Ratios of Normal Variables. *Journal of Statistical Software* 16, 4 (2006), 1–10.

[30] Frank McSherry and Kunal Talwar. 2007. Mechanism design via differential privacy. In *IEEE Symposium on Foundations of Computer Science (FOCS'07)*. IEEE.

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

[32] Prashanth Mohan, Abhradeep Thakurta, Elaine Shi, Dawn Song, and David Culler. 2012. GUPT: privacy preserving data analysis made easy. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*. 349–360.

[33] Arjun Narayan and Andreas Haeberlen. 2012. DJoin: Differentially Private Join Queries over Distributed Databases. In *10th USENIX Symposium on Operating Systems Design and Implementation, OSDI*. USENIX Association.

[34] Victoria de Sainte Agathe Nicolas Grislain, Paul Roussel. 2024. Qrlew: Rewriting SQL into Differentially Private SQL. arXiv:2401.06273

[35] Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. 2007. Smooth sensitivity and sampling in private data analysis. In *Proc. ACM Symposium on Theory of Computing (STOC '07)*. ACM.

[36] OpenDP. 2021. SmartNoise. https://smartnoise.org

[37] OpenMinded. 2022. Differentially private data aggregation. https://pipelinedp.io

[38] Davide Proserpio, Sharon Goldberg, and Frank McSherry. 2014. Calibrating Data to Sensitivity in Private Data Analysis. *PVLDB* 7, 8 (2014).

[39] Dajun Sun, Wei Dong, and Ke Yi. 2023. Confidence Intervals for Private Query Processing. *Proc. VLDB Endow.* 17, 3 (Nov. 2023).

[40] John Robert Taylor and William Thompson. 1982. *An introduction to error analysis: the study of uncertainties in physical measurements*. Vol. 2. Springer.

[41] Ashkan Yousefpour, Igor Shilov, Alexandre Sablayrolles, Davide Testuggine, Karthik Prasad, Mani Malek, John Nguyen, Sayan Ghosh, Akash Bharadwaj, Jessica Zhao, Graham Cormode, and Ilya Mironov. 2022. Opacus: User-Friendly Differential Privacy Library in PyTorch. arXiv:2109.12298

[42] Dan Zhang, Ryan McKenna, Ios Kotsogiannis, Michael Hay, Ashwin Machanavajjhala, and Gerome Miklau. 2018. EKTELO: A Framework for Defining Differentially-Private Computations. In *Proc. International Conference on Management of Data*.