Debiasing Functions of Private Statistics in Postprocessing

Flavio Calmon
Harvard University, Cambridge, MA, USA

Elbert Du
Harvard University, Cambridge, MA, USA

Cynthia Dwork
Harvard University, Cambridge, MA, USA

Brian Finley
U.S. Census Bureau, Suitland, MD, USA

Grigory Franguridi
Center for Economic and Social Research, University of Southern California, Los Angeles, CA, USA

Abstract

Given a differentially private unbiased estimate $\tilde{q} = q(D) + \nu$ of a statistic $q(D)$, we wish to obtain unbiased estimates of *functions* of $q(D)$, such as $1/q(D)$, solely through post-processing of $\tilde{q}$, with no further access to the confidential dataset D. To this end, we adapt the deconvolution method used for unbiased estimation in the statistical literature, deriving unbiased estimators for a broad family of twice-differentiable functions – those that are tempered distributions – when the privacy-preserving noise ν is drawn from the Laplace distribution (Dwork *et al.*, 2006). We further extend this technique to functions other than tempered distributions, deriving approximately optimal estimators that are unbiased for values in a user-specified interval (possibly extending to $\pm\infty$).

We use these results to derive an unbiased estimator for private means when the size n of the dataset is not publicly known. In a numerical application, we find that a mechanism that uses our estimator to return an unbiased sample size and mean outperforms a mechanism that instead uses the previously known unbiased privacy mechanism for such means (Kamath *et al.*, 2023). We also apply our estimators to develop unbiased transformation mechanisms for per-record differential privacy, a privacy concept in which the privacy guarantee is a public function of a record's value (Seeman *et al.*, 2024). Our mechanisms provide stronger privacy guarantees than those in prior work (Finley *et al.*, 2024) by using Laplace, rather than Gaussian, noise.

Finally, using a different approach, we go beyond Laplace noise by deriving unbiased estimators for polynomials under the weak condition that the noise distribution has sufficiently many moments.

2012 ACM Subject Classification Security and privacy → Data anonymization and sanitization; Mathematics of computing → Probability and statistics; Theory of computation → Theory of database privacy and security

Keywords and phrases Differential privacy, deconvolution, unbiasedness

Digital Object Identifier 10.4230/LIPIcs.FORC.2025.17

Related Version *Full Version*: `https://arxiv.org/pdf/2502.13314`

Supplementary Material *Software*: `https://github.com/franguridi/debiased-dp`
archived at `swh:1:dir:247451f8691deada77287580f8c5b1c328957a1e`

Funding *Flavio Calmon*: This work was supported in part by Simons Foundation Grant 733782 and Cooperative Agreement CB20ADR0160001 with the United States Census Bureau. This material is also based upon work supported by the National Science Foundation under Grant No. CIF-2312667 and CIF-2231707.

1 Introduction

Differential privacy (DP) has become widely accepted as a "gold standard" of privacy protection in statistical analysis. In particular, it has been adopted by many companies such as Google, Meta, and Apple to protect customer data and by the U.S. Census Bureau to protect respondent data in the 2020 Census [1]. DP mechanisms work by introducing randomness into the computation of all statistics published from a protected database. This added noise guarantees that no attacker can confidently determine from the published statistics whether a particular record is included in the dataset, thereby preserving its privacy.

Among DP mechanisms, *additive mechanisms* are canonical and widely used. These simply add data-independent, zero-mean random noise to the statistics. For example, the Laplace mechanism adds Laplace-distributed noise and is one of the first and most fundamental DP mechanisms [6]. The scale of the added noise needs to be proportional to the statistics' global sensitivity – the greatest amount by which the statistic could change upon the addition or deletion of a single record. Intuitively, this ensures that there is enough noise to mask the presence or absence of any particular record.

Often, however, the statistics to which noise is added differ from the final statistics of interest. In these cases, the statistics of interest must be estimated using the available noisy statistics. Suppose that the noisy statistic $\tilde{q}$ is formed by applying an additive mechanism to the univariate statistic q, but that we want to learn $f(q)$, not q. Even though $\tilde{q}$ is unbiased for q, the plug-in estimator $f(\tilde{q})$ is not generally unbiased for $f(q)$. When unbiasedness is desired, other estimators must be used.

To address this problem, we first derive unbiased estimators for the Laplace mechanism, for a general class of twice-differentiable functions f – those which are also *tempered distributions* (Section 3). [11] develops recursive algorithms that are unbiased estimators for polynomials in Laplace variables. Our paper provides estimators for a large class of non-polynomial functions and gives a simple, closed-form estimator for polynomials. We also provide methods to adapt functions that are not tempered distributions in a way that permits unbiased estimation over a subset of q's domain (Section 4). This extension lets us provide unbiased estimators for the case when $f(q) = 1/q$, which, in turn, lets us provide unbiased estimators of ratio statistics. Such cases arise frequently in practice, as discussed below. Finally, we derive unbiased estimators for a very general class of additive mechanisms when f is a polynomial (see Section 7).

There are several reasons why noise may not be added directly to the statistic of interest and bias in the plug-in statistic must be considered. A leading case occurs when $f(q)$ has a much higher global sensitivity than q. For example, when the domain of q includes 0 or values arbitrarily close to 0, the global sensitivity of $f(q) = 1/q$ is typically infinite and no amount of noise provides a finite DP guarantee. This same problem affects the many statistics which can be expressed as ratios of low-sensitivity statistics. For example, the mean is the ratio of a sum and a sample size. Likewise, in a simple linear regression of the regressand y on the regressor x, the ordinary least squares (OLS) estimator of the slope coefficient is the ratio of the empirical (co)variances $\mathrm{Cov}[x, y]$ and $\mathrm{Var}[x]$. It is common, then, to add noise to the low-sensitivity statistics that form these ratios and use the plug-in estimator for the ratio statistic of interest. See, for example, [2] for such a treatment of the OLS estimator.

Noise may also be added to statistics that are not of direct interest because data curators, such as government agencies, may publish noisy microdata or a noisy predetermined set of aggregates for general-purpose use. For example, a researcher trying to learn the proportion of the population with doctoral degrees may only have access to published noisy totals of the general population and the population of degree holders. The plug-in estimator of the mean is, as above, the ratio of these noisy totals.

This situation may also arise because, under DP, there is a limited "privacy budget" which is drawn upon every time we use the raw data to release another (noisy) statistic. Splitting the budget among more statistics requires that more noise be added to each of them. This makes it beneficial to "re-use" statistics whenever possible. For example, in Section 5, we develop a DP mechanism that uses our results to provide private unbiased estimates of a mean and sample size. This mechanism obtains a noisy sample size via the Laplace mechanism and then re-uses it to estimate the denominator of the mean statistic. We find that this approach outperforms an alternative mechanism that uses an existing method from [12] to add noise directly to the mean query, without re-using the noisy sample size.

Again, these scenarios all have in common that the plug-in estimator $f(\tilde{q})$ will typically be biased.[1] To see why unbiasedness is desirable, recall that the bias and variance of a sum of n uncorrelated estimates are respectively the sums of the estimates' biases and variances. Accordingly, the sum's bias increases at the rate $O(n)$ while its standard deviation grows at the rate $O(\sqrt{n})$. The sum's overall RMSE therefore grows at the $O(\sqrt{n})$ rate if the estimates are unbiased, but at the faster $O(n)$ rate otherwise.

For example, consider the following simple example: suppose the true value of some quantity of interest is 1, but each time we try to learn the value of this quantity, we get a fresh draw from the distribution $\mathcal{N}(1, 100)$. A mechanism that ignores the data and returns 0 has bias 1 and variance 0, resulting in an overall RMSE of 1. On the other hand, reporting the value of any single draw would have bias 0 and variance 100. On average, this will be off by around 10. Thus, on individual draws, the first mechanism is more accurate. However, if we take the mean of $10,000$ such draws, the mechanism that always returned 0 still gives a mean of 0, which is still off by 1. On the other hand, the mean of the $10,000$ draws now has variance $\frac{100}{10,000} = \frac{1}{100}$, resulting in an RMSE of $\frac{1}{10}$. That is, we would now expect this estimate to be off by around 0.1, which is a significant improvement over the biased estimator.

This makes unbiasedness very important in meta-analyses, which aggregate multiple estimates. It is also important when adding noise to a large number of quantities whose sums are of independent interest. This situation commonly arises in the local model of DP, where an extra layer of privacy is obtained by adding noise to every record even before entrusting it to the data curator. Sums or means using these noisy records could then be subject to severe error if the record-level estimates being summed are biased. For example, with network data, the count of k-stars (i.e., sets of k edges sharing a node) is a sum of polynomials in each node's degree. In experiments with network data protected by local DP, [11] find that a mechanism that simply sums unbiased estimates of these polynomials outperforms the L2 error of prior work by factors as high as 5 orders of magnitude.

[1] In fact, in the case with the Laplace mechanism and $f(q) = 1/q$, the expectation and all higher moments of the plug-in estimator $f(\tilde{q})$ fail to even exist, implying that the estimator has extremely fat tails and is very prone to returning extreme outliers. This affects the ratio statistics discussed above, as well. The unbiased estimator we develop for this case in Section 4.1 possesses finite moments of all orders.

Likewise, unbiasedness is key when noise is added to disaggregate sums with the expectation that they can be aggregated further to obtain sums for larger groups. For example, [13] and [8] develop mechanisms for use with per-record DP – a variant of DP whose privacy guarantees differ between records, and which is being considered by the Census Bureau for use with its County Business Patterns (CBP) data product [3]. [8] develop *transformation mechanisms* for this purpose, which improve privacy guarantees by adding noise to concave functions of q rather than to q itself. Estimates of q must then be obtained from these noisy transformed values. The CBP data includes sums of employment and payroll, grouped by finely divided geographies and industry codes. If these transformation mechanisms were used for the CBP and the estimator of q for these sums were biased, further aggregates of these estimates to obtain, say, state-level sums would be subject to severe biases.

In Section 6, we apply our estimators to create variants of these transformation mechanisms that satisfy a stronger type of per-record DP guarantee than the ones originally proposed in [8].

In this paper, we make the following contributions:

1. We derive closed-form unbiased estimators for a large class of functions – twice-differentiable functions that are tempered distributions – when the Laplace mechanism is used. We also develop estimators that are unbiased for subsets of the statistic's domain for functions that are not in this class.
2. We exposit the deconvolution method from the statistics literature (e.g., [15], page 185) for deriving unbiased estimators. This could be used to derive estimators for further functions and further mechanisms, and we believe its use in DP is novel and of independent interest.
3. We apply our unbiased estimators to create novel unbiased privacy mechanisms for per-record DP, a new variant of DP being considered for use by the Census Bureau [3].
4. We derive closed-form unbiased estimators for polynomial functions of statistics privatized using any of a large class of additive mechanisms.

2 Differential Privacy, Unbiasedness, and Deconvolution

The following is the definition of differential privacy, introduced in [6]:

▶ **Definition 1.** *Datasets D and D' are* neighboring databases *if they differ by the inclusion of at most* 1 *element.*

▶ **Definition 2.** *A mechanism $\mathcal{M}$ is (ϵ, δ)-differentially private ((ϵ, δ)-DP) if, for any pair of neighboring datasets D, D' and any measurable set of possible outcomes S, we have*

$$\Pr[\mathcal{M}(D) \in S] < e^{\epsilon} \cdot \Pr[\mathcal{M}(D') \in S] + \delta.$$

Most of our work uses Fourier transforms [9]. The following definitions and theorems are adapted from the textbook treatment in [14].

▶ **Definition 3.** *The Fourier transform of an absolutely integrable function f is*

$$F[f(x)](y) = \int_{-\infty}^{\infty} e^{-2\pi i y x} f(x) dx.$$

We often denote the Fourier transform of f by $\hat{f}(y)$. The Fourier transform also has an inverse:

$$F^{-1}[\hat{f}(y)](x) = \int_{-\infty}^{\infty} e^{2\pi i y x} \hat{f}(y) dy.$$

There are many important functions we wish to compute unbiased estimates of which are not absolutely integrable. In particular, polynomials and the function $f(q) = 1/q$ are not absolutely integrable, so we must define the Fourier transform over a more general family of functions, *tempered distributions*.[2] Importantly, in this use, the term "distribution" does not refer to a probability distribution. Rather, it refers to a class of objects also known as "generalized functions." For the purposes of this work, we can largely restrict ourselves to working with tempered distributions which are also functions, though there exist tempered distributions which are not functions, such as the Dirac delta "function". Below, we specialize the relevant theory to the case of tempered distributions that are also functions, but the interested reader should see Appendix A of this paper's full version (linked to on the title page) for the more general case.

For our purposes, then, tempered distributions can be thought of as functions that may not be absolutely integrable, but which grow no faster than a polynomial. Formally, this is expressed by the condition that the product of a tempered distribution and any function in the Schwartz space (defined below) is integrable.

▶ **Definition 4.** *The Schwartz space $S(\mathbb{R})$ is defined as follows:*

$$S(\mathbb{R}) = \{s : \mathbb{R} \to \mathbb{C} \mid s \in C^\infty,\ \sup_{x\in\mathbb{R}} |x^m s^{(n)}(x)| < \infty \quad \forall m, n \in \mathbb{N}\},$$

where $\mathbb{N}$ denotes the set of non-negative integers and $s^{(n)}$ denotes the n^{th} derivative of s. That is, functions in $S(\mathbb{R})$ are infinitely differentiable everywhere and they – along with all of their derivatives – go to 0 at a super-polynomial rate.

Note that all the functions $s \in S(\mathbb{R})$ are absolutely integrable, so their Fourier transforms $\hat{s}$ exist. With the Schwartz space so defined, we introduce tempered distributions below.

▶ **Definition 5.** *A function f is a tempered distribution if and only if, for all $s \in S(\mathbb{R})$,*

$$\int_{-\infty}^{\infty} f(x)s(x)dx \in \mathbb{C}.$$

Definition 3 introduces the Fourier transform only for absolutely integrable functions. The following definition extends it to all tempered distributions.[3]

▶ **Definition 6.** *When it exists, $\hat{f}$ is the function such that for all $s \in S(\mathbb{R})$,*

$$\int_{-\infty}^{\infty} \hat{f}(x)s(x)dx = \int_{-\infty}^{\infty} f(x)\hat{s}(x)dx.$$

Technically, the Fourier transform of a tempered distribution always exists and is a tempered distribution, but may not also be a function, even when the distribution being Fourier-transformed is a function. See Appendix A of this paper's full version for details.

The deconvolution method we use to derive unbiased estimators in Section 3 is applicable because, as explained below, the requirement that an estimator be unbiased can be expressed in terms of a convolution.

[2] Technically, $1/q$ is not a tempered distribution either, but only because it is poorly behaved at 0. This will be addressed in Section 4.

[3] To see that Definition 3 implies Definition 6 for absolutely integrable functions, note that the $e^{-2\pi ixy}$ term does not depend on the function f, so swapping the order of integration by Fubini's theorem immediately gives us the equality in Definition 6.

▶ **Definition 7.** *The convolution of functions f and g is*

$$(f * g)(x) = \int_{-\infty}^{\infty} f(z)g(x-z)dz.$$

Critically, the Fourier transform of a convolution is the product of the convolved functions' Fourier transforms.

▶ **Theorem 8** ([14] section 7.1 property c)**.**

$$\widehat{(f * g)}(y) = \hat{f}(y) \cdot \hat{g}(y).$$

We will also need the following theorem to derive unbiased estimators for the case of Laplace noise.

▶ **Theorem 9** ([14] section 7.8 Example 5)**.** *For any tempered distribution f, the Fourier transform of its k^{th} derivative $f^{(k)}$ is $\widehat{f^{(k)}} = (2\pi iy)^k \hat{f}$.*

With the query q and its privacy-preserving noisy estimate $\tilde{q}$, we say that an estimator g is unbiased for $f(q)$ if

$$f(q) = \mathbb{E}[g(\tilde{q})|q]. \tag{1}$$

By conditioning on the true query value, q, we treat the database as fixed. Our estimators, then, are unbiased with respect to the randomness in the 0-centered noise being added for privacy. Throughout the rest of this paper, all expectations are conditional on q unless otherwise noted and we suppress the extra conditioning notation so that $\mathbb{E}[.] \equiv \mathbb{E}[.|q]$.

Let the noise added for privacy be independent of the database and denote its PDF by r. The deconvolution method, as seen, for example, on page 185 of [15], starts by noting that if g is unbiased for $f(q)$, then Equation 1 can be reexpressed in terms of a convolution:

$$f(q) = \mathbb{E}[g(\tilde{q})] = \int_{-\infty}^{\infty} g(\tilde{q})r(\tilde{q}-q)d\tilde{q} = (g * r)(q). \tag{2}$$

With the unbiasedness equation in this form, Theorem 8 lets us Fourier-transform both sides to turn the convolution on the right-hand side into a simple multiplication. Finally, we simply solve for the Fourier transform of g in terms of the Fourier transforms of f and r and inverse-Fourier-transform the result. Formally,

$$f(q) = (g * r)(q) \iff \hat{f}(y) = \hat{g}(y)\hat{r}(y) \iff \hat{g}(y) = \frac{\hat{f}(y)}{\hat{r}(y)} \iff g(x) = F^{-1}\left[\frac{\hat{f}(y)}{\hat{r}(y)}\right](x), \tag{3}$$

assuming the existence of all the involved Fourier and inverse Fourier transforms.

3 Unbiased Estimation with Laplace Noise

A standard mechanism for differential privacy perturbs the query with Laplace noise scaled to the global sensitivity of a query, which is the maximum difference between the query values on neighboring databases. That is, to achieve $(\epsilon, 0)$-DP when releasing the value of a query q with global sensitivity Δ, we can simply release $q + \text{Lap}(0, \frac{\Delta}{\epsilon})$ [6].

Our primary contribution is deriving unbiased estimators for functions of q when we only have access to the value of $q + \text{Lap}(0, b)$, for some noise scale parameter b. These estimators are unique (up to their values on a set of measure zero).

▶ **Theorem 10.**
Let $\tilde{q} \sim q + \mathrm{Lap}(0, b)$.

- *For any twice-differentiable function* $f : \mathbb{R} \to \mathbb{R}$ *that is a tempered distribution,* $f(\tilde{q}) - b^2 f''(\tilde{q})$ *is an unbiased estimator for* $f(q)$.
- *For any function* $f : \mathbb{R} \to \mathbb{R}$, *if two estimators* $g_1(\tilde{q})$ *and* $g_2(\tilde{q})$ *are unbiased for* $f(q)$, *then* g_1 *and* g_2 *are equal almost everywhere.*

Proof. See Appendix B of this paper's full version, linked to on the title page. ◀

Some examples of unbiased estimators are given below.

▶ **Example 11.**

1. Any power function $f(q) = q^k$ has unbiased estimator $\tilde{q}^k - b^2 k(k-1)\tilde{q}^{k-2}$. In particular, for $f(q) = cq$ for any constant c, the unbiased estimator is also $c\tilde{q}$. Section 3.1 of [11] derives this estimator in the form of a recursive algorithm. We contribute the closed form here to simplify computation and facilitate intuitive understanding.
2. Within the set of twice differentiable functions f that are tempered distributions, Theorem 10 allows us to determine which functions are unbiased estimators of themselves. When $f(\tilde{q})$ is unbiased for $f(q)$, we have $\mathbb{E}[f(\tilde{q})] = \mathbb{E}[f(\tilde{q}) - b^2 f''(\tilde{q})] \implies \mathbb{E}[f''(\tilde{q})] = 0$. By the second part of Theorem 10 and the unbiasedness of the zero function for zero, this implies $f''(x) = 0$ almost everywhere, so $f(x)$ must be linear. The naive plug-in estimator, then, is biased for any nonlinear function in this class. This highlights the usefulness of Theorem 10.
   We can similarly characterize the f whose unbiased estimators are simply linear transformations of the plug-in estimator – that is, f for which $\mathbb{E}[\alpha f(\tilde{q}) + \beta] = f(q)$ for some $\alpha, \beta \in \mathbb{R}$. By Theorem 10, these functions satisfy $\mathbb{E}[\alpha f(\tilde{q}) + \beta] = \mathbb{E}[f(\tilde{q}) - b^2 f''(\tilde{q})]$, and so satisfy $\alpha f(x) + \beta = f(x) - b^2 f''(x)$ almost everywhere. When $\alpha \neq 1$, solutions to this differential equation take the form[4]

$$f(x) = c_1 e^{\frac{\sqrt{1-\alpha}}{b}x} + c_2 e^{-\frac{\sqrt{1-\alpha}}{b}x} + \frac{\beta}{1-\alpha}.$$

   Using Euler's formula, we can see that tempered distributions in this class include functions of the form $f(x) = c\cos(ux)$ and $f(x) = c\sin(ux)$. Nonetheless, this is still a rather restricted class of functions.

▶ Remark 12. When the function f is not twice differentiable but is a tempered distribution, an analog of Theorem 10 holds. This relies on the use of an alternative notion of the derivative that applies to all tempered distributions – the *distributional derivative*. For background on this derivative concept, see Appendix A of this paper's full version.

In this case, we still have $\mathbb{E}[f(\tilde{q})] - \mathbb{E}[b^2 f''(\tilde{q})] = f(q)$, but the distributional derivative f'' is a tempered distribution which is not a function. This does not give us an unbiased estimator, but instead we can rearrange to obtain the bias, as a function of q, of the naive plug-in estimator $f(\tilde{q})$:

$$\mathbb{E}[f(\tilde{q})] - f(q) = \mathbb{E}[b^2 f''(\tilde{q})] = \frac{b}{2}\int_{-\infty}^{\infty} f''(q+x)e^{-|x|/b}dx. \tag{4}$$

[4] The case where $\alpha = 1$ and $\beta = 0$ is dealt with above.

For example, let $f(x) = |x|$, $f'(x) = -1$ for $x \leq 0$ and 1 for $x > 0$ (the discontinuity at 0 is irrelevant since $\{0\}$ is a set of measure 0). Then $f''(x) = 2\delta(x)$, where δ is the Dirac delta function (defined in Example 29 in Appendix A of this paper's full version). The bias is simply $\frac{b}{2} \cdot 2e^{-|q|/b} = be^{-|q|/b}$.

Whether or not f is twice differentiable, Equation 4 suggests the intuition that the plug-in estimator will have greater bias when f has greater curvature near the true query value q.

4 Extension to Functions that are not Tempered Distributions

If the function $q \mapsto f(q)$ is not a tempered distribution, we can often bound the domain such that it is continuous and twice-differentiable in that domain. That is, suppose we know a priori that $q \geq L$ for some lower bound $L \in \mathbb{R}$. This is often the case in differential privacy, as DP-protected queries are commonly sums of nonnegative variables. Likewise, counts can often be lower bounded by 1. Then, suppose we replace the function f with some function

$$\tilde{f}(q) = \begin{cases} f(q) & q \geq L \\ h(q) & q < L \end{cases}, \tag{5}$$

where $h(L) = f(L)$ and h is twice differentiable with $h'(L) = f'(L)$ and $h''(L) = f''(L)$. The function $\tilde{f}$ is thus twice differentiable. Assuming that $h(q)$ and $f(q)$ and their derivatives grow no faster than a polynomial as, respectively, $q \to -\infty$ and $q \to \infty$, $\tilde{f}$ is a tempered distribution, as well. We can then apply Theorem 10 to get an unbiased estimator of $\tilde{f}$, i.e.

$$\mathbb{E}\left[\tilde{f}(\tilde{q}) - b^2\tilde{f}''(\tilde{q})\right] = \tilde{f}(q). \tag{6}$$

With the assumption that $q \geq L$, we have $\tilde{f}(q) = f(q)$, making this estimator unbiased for $f(q)$, as well.

▶ **Example 13.** For $f(q) = \frac{1}{q}$ and $L = 1$, we need to find some function h such that $h(1) = 1, h'(1) = -1, h''(1) = 2$. An example of such a function is $h(q) = 1-(q-1)+(q-1)^2$. We can generically use polynomials for h whenever f grows at most polynomially as $q \to \infty$ and is twice differentiable for $q \geq L$.

We now focus on optimizing this method over polynomial extensions for a particular function of interest: $f(q) = 1/q$.

4.1 Unbiased Estimation for $f(q) = 1/q$

We have shown that it is possible to construct a function that permits unbiased estimation as long as it is twice differentiable on some domain that the true query value is known to be in, and, if this domain is unbounded, as long as the function does not grow too quickly. In this section, we show how to optimally choose the function h in the above construction.

We restrict ourselves to polynomial functions h for two reasons. First, the solution among polynomials of fixed degree is efficiently computable. Second, when the optimal function h is a twice continuously differentiable tempered distribution, polynomials can approximate this function arbitrarily well, in the sense that the expected squared error of the polynomial-based estimator can be made arbitrarily close to optimal. This follows from Theorem 14.

▶ **Theorem 14** (Polynomial approximation). *Let $L \in \mathbb{R}$ and let $f : [L, \infty) \to \mathbb{R}$ be twice differentiable and a tempered distribution. Let μ be a probability measure such that the integrals $\int_L^\infty f(q)e^{(L-q)/b}d\mu(q)$ and $\int_L^\infty e^{(L-q)/b}d\mu(q)$ exist and are finite. With $w : (-\infty, L] \to \mathbb{R}$, let $\tilde{f}[w]$ denote the function*

$$\tilde{f}[w](q) = \begin{cases} f(q) & q \geq L \\ w(q) & q < L \end{cases} \tag{7}$$

Let $h : (-\infty, L] \to \mathbb{R}$ be an arbitrary twice continuously differentiable tempered distribution that satisfies $h(L) = f(L)$, $h'(L) = f'(L)$ and $h''(L) = f''(L)$. Denote the estimator $g \equiv \tilde{f}[h] - b^2\tilde{f}[h]''$ and denote its expected squared error by

$$\alpha \equiv \int_{-\infty}^{\infty}\int_{L}^{\infty} (g(x) - f(q))^2 \frac{1}{2b} e^{-|x-q|/b} d\mu(q)dx. \tag{8}$$

There exists a sequence of polynomials $(p_K)_{K=1}^{\infty}$ over $(\infty, L]$ that satisfy $p_K(L) = f(L)$, $p_K'(L) = f'(L)$ and $p_K''(L) = f''(L)$ such that the sequence of associated estimators $g_K \equiv \tilde{f}[p_K] - b^2\tilde{f}[p_K]''$ satisfies

$$\lim_{K\to\infty} \int_{-\infty}^{\infty}\int_{L}^{\infty} (g_K(x) - f(q))^2 \frac{1}{2b} e^{-|x-q|/b} d\mu(q)dx = \alpha. \tag{9}$$

See Appendix C of this paper's full version for a proof.

Now, letting h be a polynomial, suppose our estimator is $g(x) = \tilde{f}(x) - b^2\tilde{f}''(x)$ for $\tilde{f}$ defined in Equation 5. For our error metric, we consider the estimator's expected squared error, with the expectation taken over both the privacy noise and prior beliefs about q, reflected in the probability measure μ. We define our estimator as the solution to the following constrained optimization problem:

$$\begin{aligned}
&\min_{\tilde{f}} \int_{-\infty}^{\infty}\int_{L}^{\infty} \left(\tilde{f}(x) - b^2\tilde{f}''(x) - f(q)\right)^2 \frac{1}{2b} e^{-|x-q|/b} d\mu(q)dx \\
&= \int_{L}^{\infty}\int_{L}^{\infty} \left(f(x) - b^2 f''(x) - f(q)\right)^2 \frac{1}{2b} e^{-|x-q|/b} d\mu(q)dx \\
&+ \min_{g} \int_{-\infty}^{L}\int_{L}^{\infty} (g(x) - f(q))^2 \frac{1}{2b} e^{(x-q)/b} d\mu(q)dx \\
&\text{subject to } h(L) = f(L), h'(L) = f'(L), \text{ and } h''(L) = f''(L).
\end{aligned} \tag{10}$$

Since the first double integral is constant with respect to h, optimizing this error metric is equivalent to optimizing

$$\min_{g} \int_{-\infty}^{L}\int_{L}^{\infty} (g(x) - f(q))^2 \frac{1}{2b} e^{(x-q)/b} d\mu(q)dx, \tag{11}$$

subject to the same constraints.

For simplicity, we shall now treat g as a function with domain $(-\infty, L]$, as that is the only region on which we are optimizing, so $g(x) = \sum_{i=0}^{k} a_i x^i$. There is a one-to-one correspondence between polynomials $g(x)$ and polynomials $h(x) = \sum_{i=0}^{k} b_i x^i$ where $a_i = b_i - b^2(i+2)(i+1)b_{i+2}$. Thus, we are considering extensions of $\tilde{f}(q)$ where the part to the left of the lower bound L is a polynomial.

▶ **Theorem 15.** *For any positive integer k, any real number $L \in \mathbb{R}$, and any function f which is twice differentiable on $[L, \infty)$, there is an algorithm that runs in time poly(k) which computes the polynomial g that minimizes*

$$\int_{-\infty}^{L}\int_{L}^{\infty} (g(x) - f(q))^2 \frac{1}{2b} e^{(x-q)/b} d\mu(q)dx$$

over polynomials of degree k, satisfying the constraints $g(x) = h(x) - b^2h''(x)$, $h(L) = f(L), h'(L) = f'(L)$, and $h''(L) = f''(L)$.

Proof. See Appendix C.1 of this paper's full version, linked to on the title page. ◀

▶ **Corollary 16.** *Provided that the optimal choice of h is a twice continuously differentiable tempered distribution, there exists an efficient algorithm to approximate the optimal unbiased estimator of $f(q)$ given $q + Z$ for $Z \sim \text{Lap}(0, b)$ and the prior knowledge that $q \geq L$.*

This follows immediately from the fact that optimizing g also optimizes $\tilde{f}(q)$.

Note that this result can be easily extended to the cases where we have only an upper bound or both an upper and lower bound. If we only have an upper bound, everything works out exactly the same as if we only have a lower bound. If we have both, suppose we know that $q \in [L, U]$ and define

$$\tilde{f}(q) = \begin{cases} h_0(q) & q > U, \\ f(q) & U \geq q \geq L, \\ h_1(q) & q < L. \end{cases} \tag{12}$$

Then the expected error (with the expectation over both the privacy noise and the prior on q) is

$$\begin{aligned} &\mathbb{E}_{q,x\sim q+Lap(0,b)}\left[\left(\tilde{f}(x) - b^2\tilde{f}''(x) - f(q)\right)^2\right] \\ &= \int_{-\infty}^{\infty}\int_{L}^{U}\left(\tilde{f}(x) - b^2\tilde{f}''(x) - f(q)\right)^2 \frac{1}{2b}e^{-|x-q|/b}d\mu(q)dx. \end{aligned} \tag{13}$$

Just like before, the error incurred by $\tilde{f}(x)$ on $L \leq x \leq U$ is not affected by our choice of functions. Thus, we wish to compute

$$\begin{aligned} &\min_{h,h'}\int_{-\infty}^{L}\int_{L}^{U}\left(h_1(x) - b^2h_1''(x) - f(q)\right)^2 \frac{1}{2b}e^{-|x-q|/b}d\mu(q)dx \\ &\quad + \int_{U}^{\infty}\int_{L}^{U}\left(h_0(x) - b^2h_0''(x) - f(q)\right)^2 \frac{1}{2b}e^{-|x-q|/b}d\mu(q)dx. \end{aligned} \tag{14}$$

Since there is no interaction between h_0 and h_1, we can minimize these integrals independently in the same way as above.

5 Numerical Results: Application to Mean Queries

In this section, we illustrate the utility of our results by numerically comparing two mechanisms designed to return unbiased estimates of the sample size n and the mean m of an attribute $c \in [0, 1]$ in the database D. Sample sizes are published alongside any reported means in most research applications, making this a realistic use case. One mechanism, M_U, returns an unbiased estimate of the mean using the results from Section 4.1. The other, M_{SS}, uses the unbiased mean mechanism from [12] (see their Theorem D.6 and proof). To the best of our knowledge, this is the only published unbiased mechanism for means when the sample size is not treated as known. Both mechanisms use the Laplace mechanism with privacy budget ϵ_1 to obtain the noisy sample size $\tilde{n}$. Each mechanism then allocates a separate privacy budget ϵ_2 to obtain a noisy mean. Both mechanisms have a total privacy budget of $\epsilon_0 = \epsilon_1 + \epsilon_2$.

Denote attribute c of record r by $r.c$ and let $g(\tilde{q}; k, L)$ be the unbiased estimator of $1/q$ from Section 4.1 with the generic query q and polynomial extension of order k for $q \leq L$. Algorithm 1 lays out M_U. This algorithm applies the Laplace mechanism to the sum $s \equiv \sum_{r\in D} r.c$ and forms an unbiased estimate $\tilde{m}_U$ of the mean $m = s/n$ by multiplying the noisy sum $\tilde{s}$ by $g(\tilde{n}; k, L)$. This is unbiased for m as long as $n \geq L$.

For $n \geq L$, the variance of Algorithm 1 is

$$\mathbb{V}[\tilde{m}_U] = (\mathbb{E}[\tilde{s}]^2 + \mathbb{V}[\tilde{s}])(\mathbb{E}[g(\tilde{n}; k, L)]^2 + \mathbb{V}[g(\tilde{n}; k, L)]) - \mathbb{E}[\tilde{s}]^2\,\mathbb{E}[g(\tilde{n}; K, L)]^2 \tag{15}$$

$$= (s^2 + 2/\epsilon_1^2)(\mathbb{E}[1/n]^2 + \mathbb{V}[g(\tilde{n}; k, L)]) - s^2/n^2. \tag{16}$$

For this section's numerical results, we calculate $\mathbb{V}[g(\tilde{n}; k, L)]$ numerically. Finally, we note that it is straightforward to show that all moments of $\tilde{m}_U$ exist and are finite.

Algorithm 1 M_U

procedure MeanPostprocess(Database D, sample size privacy budget ϵ_1, mean privacy budget ϵ_2, polynomial order k, sample size lower bound L)
  $n \leftarrow \sum_{r \in D} 1$
  $\tilde{n} \leftarrow \mathrm{Lap}(n, 1/\epsilon_1)$
  $s \leftarrow \sum_{r \in D} r.c$
  $\tilde{s} \leftarrow \mathrm{Lap}(s, 1/\epsilon_2)$
  $\tilde{v} \leftarrow g(\tilde{n}; k, L)$ (Unbiased estimator of $1/n$ for $n \geq L$)
  $\tilde{m}_U \leftarrow \tilde{s}\tilde{v}$
  output $(\tilde{n}, \tilde{m}_U)$
end procedure

Let T_3 denote a random variable distributed according to a standard t distribution with 3 degrees of freedom. M_{SS} is laid out in Algorithm 2. This algorithm first forms a version m_{SS} of the mean query that simply equals 1 if $n = 0$. It then scales the noise variable T_3 in proportion to an upper bound on the query's smooth sensitivity [12]. The final noisy mean $\tilde{m}_{SS}$ is obtained by simply adding the scaled noise variable to m_{SS}.

The scaling factor for the noise is $\tau \max(e^{-\beta(n-1)}, 1/\max(n, 1))$, where τ and β satisfy $\epsilon_2 = 4\beta + 2/(\sqrt{3}\tau)$. The standard t distribution with d degrees of freedom has variance $d/(d-2)$ giving $\tilde{m}_{SS}$ a variance of

$$\mathbb{V}[m_{SS}] = 3\tau^2 \max(e^{-\beta(n-1)}, 1/\max(n, 1))^2. \tag{17}$$

Because the t distribution is symmetric, this mechanism is unbiased for s/n as long as $n \geq 1$. Unlike $\tilde{m}_U$, however, the third and higher moments of $\tilde{m}_{SS}$ are infinite or do not exist. This is due to the t distribution's very fat tails and implies that $\tilde{m}_{SS}$ is more liable than $\tilde{m}_U$ to produce extreme outliers.

Algorithm 2 M_{SS}.

procedure MeanSmoothSens(Database D, sample size privacy budget ϵ_1, mean privacy budget ϵ_2, noise parameters β, τ satisfying $\epsilon_2 = 4\beta + 2/(\sqrt{3}\tau)$)
  $n \leftarrow \sum_{r \in D} 1$
  $\tilde{n} \leftarrow \mathrm{Lap}(n, 1/\epsilon_1)$
  $s \leftarrow \sum_{r \in D} r.c$
  $m_{SS} \leftarrow$ **if** $\boldsymbol{n \geq 1}$**:** $\boldsymbol{s/n}$ **else: 1**
  $\tilde{m}_{SS} \leftarrow m_0 + T_3 \cdot \tau \max(e^{-\beta(n-1)}, 1/\max(n, 1))$
  output $(\tilde{n}, \tilde{m}_{SS})$
end procedure

In our numerical evaluation of these mechanisms, we fix $\epsilon_1 = \epsilon_2 = m = .5$ for both mechanisms. For M_{SS}, we follow [12] in setting $\beta = \epsilon_2/12$ and $\tau = \sqrt{3}/\epsilon_2$. For M_U, we set $L = 1$ so that both mechanisms are unbiased for $n \geq 1$ and set $k = 10$.

With these settings, we compare the standard deviations (SDs) of the two mechanisms' mean estimates for a range of sample sizes. Because the mechanisms are unbiased, this is equivalent to their root mean squared error. The sample size estimates of both mechanisms are the same, so we do not report their properties.

Figures 1 and 2 plot the mechanisms' SDs as functions of n, with Figure 2 zooming in on larger values of n for clarity. It is immediately clear that the M_U has a larger SD than M_{SS} for $n < 13$, and that this pattern reverses for larger n. For $n \leq 19$, however, both mechanisms have SDs greater than 1, making both unfit for most purposes at these sample sizes, given that the mean m has the domain [0,1].

Figure 3 shows the relative SD - that is, the ratio $SD(M_{SS})/SD(M_U)$ - as a function of n. For $n \geq 13$, the relative SD rises to a peak near 15 before settling down to an apparent constant of about 1.9 for $n \geq 115$.

Ultimately, M_U appears to be the better mechanism for this setting; for any sample size where either mechanism returns useful results, M_U has a substantially lower SD. The thinner tails of M_U also recommend it as the better choice.

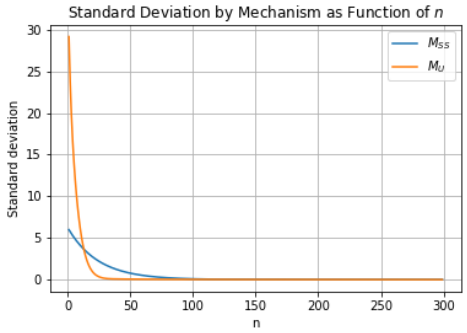

Figure 1 Standard deviations of the mechanisms M_{SS} and M_U for a mean of n records in [0,1]. The mean is fixed at .5 and the mechanisms have a privacy budget of $\epsilon_2 = .5$.

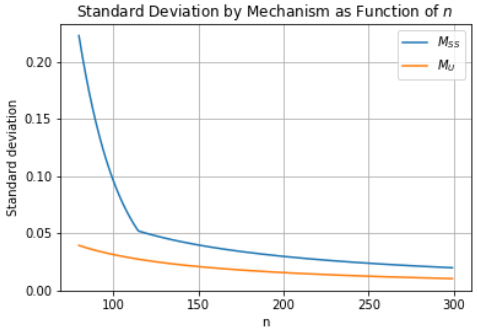

Figure 2 Same as Figure 1, zoomed in to larger values of n (see the horizontal axis endpoints). Standard deviations of the mechanisms M_{SS} and M_U for a mean of n records in [0,1]. The mean is fixed at .5 and the mechanisms have a privacy budget of $\epsilon_2 = .5$.

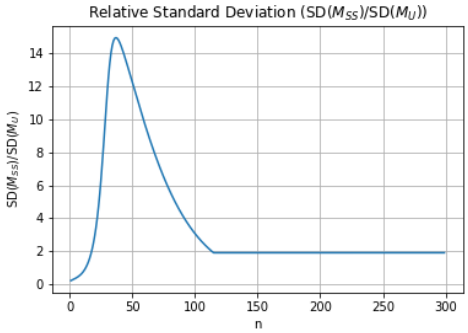

Figure 3 Standard deviation of M_{SS} divided by the standard deviation of M_U for a mean of n records in [0,1]. The mean is fixed at .5 and the mechanisms have a privacy budget of $\epsilon_2 = .5$.

6 Application to Slowly Scaling PRDP

In this section, we use our estimators to develop versions of unbiased privacy mechanisms from [8] that enjoy stronger privacy guarantees. Our estimators allow us to do so while maintaining the mechanisms' unbiasedness.

$(\epsilon, 0)$-DP guarantees the upper bound ϵ on the privacy loss between any pair of neighboring databases. [8] develops mechanisms for a related privacy concept, per-record DP (PRDP) [13]. PRDP generalizes $(\epsilon, 0)$-DP by letting the privacy loss bound be a function of the record on which a given pair of neighbors differs. Semantically, this allows different records to have different levels of protection.

Denote a record in the database D by r and denote r's attribute c by $r.c \in [0, \infty)$. PRDP was originally motivated by the need to protect data used to compute the sum query $q(D) = \sum_{r.c \in D} r.c$. Because the domain of c is unbounded, this sum can change by an arbitrarily large amount when a record is added or deleted. That is, the sum's global sensitivity is infinite. This prevents commonly used privacy mechanisms, such as the Laplace mechanism, from providing a differential privacy guarantee with finite ϵ.

The traditional fix for this is to clip attribute c to lie in a bounded set before taking the sum. $(\epsilon, 0)$-DP can then be guaranteed by perturbing the sum with noise scaled in proportion to the width of the clipped data's domain. Unfortunately, when the sum is dominated by a small number of large outliers, the outliers typically need to be clipped to drastically smaller values to preserve a reasonable balance of privacy loss and noise variance. This can induce catastrophic bias, rendering the clipped sums essentially useless. One might expect to see this type of behavior with income data, for example.

PRDP allows us to take a finer look at the privacy-utility tradeoff by recognizing that, even though outliers may suffer extreme privacy loss, the rest of the dataset may still enjoy strong privacy protections. Intuitively, a particular record's privacy loss is proportional only to the amount by which the addition or deletion of *that record* can change the query. Queries may be highly sensitive to the presence of outliers while being relatively insensitive to typical records, leading different records to have different levels of privacy loss. The reassurance that the vast majority of the data may enjoy strong privacy guarantees whether or not the data is clipped may allow a data curator to reasonably decide against clipping if the resulting bias outweighs the enhanced privacy protection for a small number of records.

Below, we define PRDP.

▶ **Definition 17** (P-Per-Record Differential Privacy (P-PRDP) [13, 8]). *Let $\ominus$ denote the symmetric set difference. The mechanism M satisfies per-record differential privacy with the policy function P (P-PRDP) if, for any record r; any pair of neighboring databases D, D' such that $D \ominus D' = \{r\}$; and any measurable set of possible outcomes S, we have*

$$\Pr[\mathcal{M}(D) \in S] < e^{P(r)} \cdot \Pr[\mathcal{M}(D') \in S].$$

Ensuring strong privacy protection corresponds to ensuring that P is, in some sense, small. $(\epsilon, 0)$-DP is recovered by making the constant privacy guarantee $P(r) = \epsilon$ for all r, and strong privacy protection follows from a small ϵ. We cannot always make a guarantee this strong. Take the example where we want to publish a sum query on an unbounded attribute c that we are unwilling to clip. In this case, the privacy loss of the mechanisms that we will consider here is growing in $r.c$. Even though we cannot prevent $P(r)$ from growing without bound in $r.c$, we can use mechanisms for which the growth rate is slow. [8] call such mechanisms "slowly scaling." A slowly growing P narrows the gap in privacy losses between records with large and small values of c, letting a data curator more easily provide a desired level of protection for the bulk of the data without compromising too much on the privacy of outliers.

[8] introduces slowly scaling mechanisms, called transformation mechanisms, that work by adding Gaussian noise to a concave transformation f of the query (plus offset term) $q(D) + a$ and then feeding the noisy value of $f(q(D) + a)$ to an estimator g of $q(D)$. By adding Gaussian noise, these mechanisms satisfy per-record zero-concentrated DP (PRzCDP), which is a weaker privacy guarantee than PRDP. PRzCDP relates to zero-concentrated DP [4, 7] in the same way that PRDP relates to ϵ-DP. The use of Gaussian noise also allowed [8] to draw on existing unbiased estimators from [16] to make their mechanism unbiased for a variety of transformation functions f.

Using the unbiased estimators from Theorem 10, we strengthen the transformation mechanisms to provide PRDP guarantees by adding Laplace, rather than Gaussian, noise, and we do so without losing the mechanism's unbiasedness. Algorithm 3 lays out our transformation mechanism.

Algorithm 3 PRDP Transformation Mechanism.

procedure TransformationPrivatizeLap(Private query answer $q(D)$, offset parameter a,
scale parameter b, transformation function $f : [a, \infty) \to \mathcal{F} \subseteq \mathbb{R}$, estimator $g : \mathcal{F} \to \mathcal{G} \subseteq \mathbb{R}$)
  $v \leftarrow f(q(D) + a)$
  $\tilde{v} \leftarrow \text{Lap}(v, b)$
  $\tilde{S} \leftarrow g(\tilde{v})$
  output $\tilde{S}$
end procedure

To obtain the PRDP guarantee of Algorithm 3, we first need to define the per-record sensitivity [8], a record-specific analog of the global sensitivity.

▶ **Definition 18** (Per-Record Sensitivity [8]). *The per-record sensitivity of the univariate, real-valued query q for record r is*

$$\Delta(r) \equiv \sup_{D, D' \text{ such that } D \ominus D' = \{r\}} |q(D) - q(D')|.$$

Theorem 19 gives our most generic result on the PRDP guarantees of Algorithm 3. Theorems 19 and 20, their proofs, as well as Algorithm 3 are minimally modified from their analogs in [8], which use Gaussian, rather than Laplace noise. This is to facilitate the interested reader's comparison of our results with theirs.

▶ **Theorem 19** (PRDP Guarantee for Transformation Mechanisms). *Assume the query value $q(D) \in [0,\infty)$; the offset parameter $a \in \mathbb{R}$; the noise scale parameter $b \in (0,\infty)$; the transformation function $f : [a,\infty) \to \mathcal{F} \subseteq \mathbb{R}$ is concave and strictly increasing; and the estimator $g : \mathcal{F} \to \mathcal{G} \subseteq \mathbb{R}$. Denote by $\Delta_f(r)$ the per-record sensitivity of the query $f(q(D)+a)$, as defined in Definition 18. Algorithm 3$(q(D), a, b, f, g)$ satisfies P-PRDP for $P(r) = \frac{\Delta_f(r)}{b}$.*

See Appendix D of this paper's full version (linked to on the title page) for the proof.

Probably the most common query encountered in applications of formal privacy, even as a component of other, larger queries, is the sum query. We now use the above result to derive the policy function for the transformation mechanism applied to a sum query.

▶ **Theorem 20** (Privacy of Transformation Mechanisms for Sum Query). *Let the assumptions of Theorem 19 hold, and further assume $a \geq 0$ and $r.c \geq 0$ for all records r. For the sum query $q(D) = \sum_{r\in D} r.c$, the per-record sensitivity of $f(q(D)+a)$ is $\Delta_f(r) = f(r.c+a) - f(a)$ and Algorithm 3$(q(D), a, b, f, g)$ satisfies PRDP with the policy function $P(r) = [f(r.c+a) - f(a)]/b$.*

See Appendix D for the proof.

Critically, the policy function from Theorem 20 grows in $r.c$ more slowly when the transformation function f grows more slowly. In the case where $a = 0$ and $f(x) = \sqrt[k]{x}$ for some $k \geq 1$, the policy function is simply $\sqrt[k]{r.c}$. Choosing larger values of k, then, forces the privacy loss to grow more slowly in $r.c$, reducing the gap in privacy losses between records with large and small values of c.

Applying our main results, we can obtain estimators such that the transformation mechanism gives us an unbiased estimate of $q(D)$. In particular, polynomials[5] are twice differentiable functions which are tempered distributions, so the following holds for $f(x) = \sqrt[k]{x}$:

▶ **Corollary 21.** *Given any function f such that f^{-1} satisfies the conditions in Theorem 10, $a \in \mathbb{R}$, $b \in (0,\infty)$, estimator $g : \mathcal{F} \to \mathcal{G} \subseteq \mathbb{R}$, and $r.c \geq 0$ for all records r, there exists an unbiased estimator for $q(D)$ satisfying P-PRDP for $P(r) = [f(r.c+a) - f(a)]/b$.*

Proof. The conditions for Theorem 20 hold, and Theorem 10 gives us a function g such that $\mathbb{E}[g(\tilde{v})] = q(D) + a$ is an unbiased estimator for f^{-1}. Therefore, $g(\tilde{v}) - a$ is an unbiased estimator for $q(D)$. ◀

7 Polynomial Functions under General Noise Distributions

Additive mechanisms other than the Laplace mechanism, such as the discrete Gaussian or the staircase mechanisms, may be preferable in practice due to achieving higher accuracy while having similar privacy loss [5, 10]. In contrast to the Laplace case, these mechanisms may not admit tractable Fourier transforms, and hence unbiased estimators are generally not available in closed form. One exception is when the query of interest is a polynomial in one or many queries.

[5] In the case of polynomials, [11] derived unbiased estimators which could also be used here. The estimator obtained from our Theorem 10 merely simplifies computation in this setting.

Using the following results to obtain an unbiased estimator of a polynomial that approximates a non-polynomial estimand may also allow users to obtain approximately unbiased estimators with great generality.

▶ **Theorem 22.** *Suppose a mechanism takes as input $q \in \mathbb{R}$ and outputs $\tilde{q} = q + Z$ for a random variable Z with at least p finite, publicly known moments. If $f(q)$ is a polynomial in q of degree at most p, there exists an unbiased estimator $g(\tilde{q})$ of $f(q)$, which is itself a polynomial of degree at most p and is available in closed form.*

Proof. Suppose $f(q) = \sum_{n=0}^{p} b_n q^n$. Let us find an unbiased estimator of the form $g(\tilde{q}) = \sum_{n=0}^{p} a_n \tilde{q}^p$. We have

$$g(q+z) = \sum_{n=0}^{p} a_n \sum_{k=0}^{n} \binom{n}{k} q^k z^{n-k}. \tag{18}$$

Denote $\mu_r = \mathbb{E}[z^r]$ and take expectations to obtain

$$\mathbb{E}[g(q+z)] = \sum_{n=0}^{p} a_n \sum_{k=0}^{n} \binom{n}{k} q^k \mu_{n-k} \tag{19}$$

$$= \sum_{n=0}^{p} \binom{n}{0} a_n \mu_n + \left(\sum_{n=1}^{p} \binom{n}{1} a_n \mu_{n-1} \right) q + \left(\sum_{n=2}^{p} \binom{n}{2} a_n \mu_{n-2} \right) q^2 + \cdots \tag{20}$$

For this polynomial to be equal to f, we need $a = (a_0, \ldots, a_p)'$ to solve $Ma = b$, where $b = (b_0, \ldots, b_p)'$ and

$$M = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_p \\ 0 & \binom{1}{1}\mu_0 & \binom{2}{1}\mu_1 & \binom{3}{1}\mu_2 & \cdots & \binom{p}{1}\mu_{p-1} \\ 0 & 0 & \binom{2}{2}\mu_0 & \binom{3}{2}\mu_1 & \cdots & \binom{p}{2}\mu_{p-2} \\ 0 & 0 & 0 & \binom{3}{3}\mu_0 & \cdots & \binom{p}{3}\mu_{p-3} \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \binom{p}{p}\mu_0 \end{pmatrix}. \tag{21}$$

Clearly, M is nondegenerate, and so the desired coefficients a exist and are unique. ◀

We now extend this result to polynomials in multiple (univariate) queries, assuming that the noise variables added to each query are independent. The latter assumption, while seemingly restrictive, is typical for additive noise mechanisms in differential privacy.

▶ **Theorem 23.** *Suppose a mechanism takes as input $(q_1, \ldots, q_m)$ and outputs $(\tilde{q}_1, \ldots, \tilde{q}_m) = (q_1 + Z_1, q_2 + Z_2, \ldots, q_m + Z_m)$ for independent random variables $Z_1, \ldots, Z_m$ with finite, publicly known moments. If $f(q_1, \ldots, q_m)$ is a polynomial in $(q_1, \ldots, q_m)$, there exists an unbiased estimator $g(\tilde{q}_1, \ldots, \tilde{q}_m)$ of $f(q)$, which is itself a polynomial available in closed form.*

Proof. Clearly, it suffices to derive unbiased estimators for $f(q_1, \ldots, q_m) = \prod_{i=1}^{m} q_i^{p_i}$. Let $g_i(\tilde{q}_i)$ be the unbiased estimator of $q_i^{p_i}$ as in Theorem 22 and set $g(\tilde{q}_1, \ldots, \tilde{q}_m) = \prod_{i=1}^{m} g_i(\tilde{q}_i)$. Since $g_1(\tilde{q}_1), \ldots, g_m(\tilde{q}_m)$ are independent random variables, we have

$$\mathbb{E}[g(\tilde{q}_1, \ldots \tilde{q}_m)] = \mathbb{E}\left[\prod_{i=1}^{m} g_i(\tilde{q}_i) \right] = \prod_{i=1}^{m} \mathbb{E}\left[g_i(\tilde{q}_i) \right] = \prod_{i=1}^{m} q_i^{p_i} = f(q_1, \ldots, q_m). \tag{22}$$

◀

8 Conclusions and Future Work

In this work, we have shown how to compute unbiased estimators of twice-differentiable tempered distributions evaluated on privatized statistics with added Laplace noise. In addition, we have proposed a method to extend this result to twice-differentiable functions which are not tempered distributions in a way that achieves approximately optimal expected squared error.

As the Laplace mechanism is simple and commonly used, these results are widely applicable to obtain unbiased statistics for free in postprocessing, which is particularly valuable due to the fact that aggregating unbiased statistics accumulates error more slowly than aggregating biased statistics. We have applied our results to derive a competitive unbiased algorithm for means and to derive unbiased transformation mechanisms for per-record DP mechanisms that enjoy stronger privacy protection than do analogs in previous work. Finally, we have derived an unbiased estimator for polynomials under arbitrary noise distributions with known moments, such as the discrete Gaussian mechanism or the staircase mechanism [5, 10].

We believe this paper opens several avenues for future research. These include the use of the deconvolution method to obtain unbiased estimators for other estimands and noise distributions. We believe a deconvolution method using multivariate Fourier transforms could also be used to obtain unbiased estimators of functions of multivariate queries. Although we did not attempt to optimize the numerical implementation in Section 5 of the integration in Section 4, we believe that an improved implementation could enable the practical use of higher-order polynomial extensions and further reduce error. In Section 7, we developed estimators that are exactly unbiased for polynomials that could approximate other functions of interest. Further work could elaborate on this process, developing concrete procedures for picking the approximating polynomial and deriving bounds on the resulting bias. Finally, future work could attempt to derive noise distributions that are optimal in the sense of minimizing the variances (or other utility metrics) of their unbiased estimators.

References

1 John Abowd, Robert Ashmead, Ryan Cumings-Menon, Simson Garfinkel, Micah Heineck, Christine Heiss, Robert Johns, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, Brett Moran, William Sexton, Matthew Spence, and Pavel Zhuravlev. The 2020 Census Disclosure Avoidance System TopDown Algorithm. *Harvard Data Science Review*, (Special Issue 2), June 24 2022. URL: `https://hdsr.mitpress.mit.edu/pub/7evz361i/release/2`.

2 Daniel Alabi, Audra McMillan, Jayshree Sarathy, Adam Smith, and Salil Vadhan. Differentially private simple linear regression, 2020. `arXiv:2007.05157`.

3 Margaret Beckom, William Sexton, and Anthony Caruso. Researching formal privacy for the Census Bureau's County Business Patterns program, 2023. URL: `https://www.census.gov/data/academy/webinars/2023/differential-privacy-webinar.html`.

4 Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*, pages 635–658. Springer, 2016. `doi:10.1007/978-3-662-53641-4_24`.

5 Clément L. Canonne, Gautam Kamath, and Thomas Steinke. The discrete Gaussian for differential privacy. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc.

6 Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. *Journal of Privacy and Confidentiality*, 7(3):17–51, May 2017. `doi:10.29012/jpc.v7i3.405`.

7 Cynthia Dwork and Guy N Rothblum. Concentrated differential privacy. *arXiv preprint arXiv:1603.01887*, 2016.

8 Brian Finley, Anthony M Caruso, Justin C Doty, Ashwin Machanavajjhala, Mikaela R Meyer, David Pujol, William Sexton, and Zachary Terner. Slowly scaling per-record differential privacy. *arXiv preprint arXiv:2409.18118*, 2024. doi:10.48550/arXiv.2409.18118.

9 Jean Baptiste Joseph Fourier. *The Analytical Theory of Heat*. Cambridge Library Collection - Mathematics. Cambridge University Press, 2009.

10 Quan Geng, Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The staircase mechanism in differential privacy. *IEEE Journal of Selected Topics in Signal Processing*, 9(7):1176–1184, 2015. doi:10.1109/JSTSP.2015.2425831.

11 Quentin Hillebrand, Vorapong Suppakitpaisarn, and Tetsuo Shibuya. Unbiased locally private estimator for polynomials of Laplacian variables. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 741–751, 2023. doi:10.1145/3580305.3599537.

12 Gautam Kamath, Argyris Mouzakis, Matthew Regehr, Vikrant Singhal, Thomas Steinke, and Jonathan Ullman. A bias-variance-privacy trilemma for statistical estimation, 2023. doi:10.48550/arXiv.2301.13334.

13 Jeremy Seeman, William Sexton, David Pujol, and Ashwin Machanavajjhala. Privately answering queries on skewed data via per-record differential privacy. *Proc. VLDB Endow.*, 17(11):3138–3150, 2024. doi:10.14778/3681954.3681989.

14 Gerrit van Dijk. *Distribution Theory*. De Gruyter, Berlin, Boston, 2013. doi:doi:10.1515/9783110298512.

15 V.G. Voinov and M. Nikulin. *Unbiased Estimators and Their Applications Volume 1: Univariate Case*. Kluwer Academic Publishers, Dordrecht, 1993.

16 Yasutoshi Washio, Haruki Morimoto, and Nobuyuki Ikeda. Unbiased estimation based on sufficient statistics. *Bulletin of Mathematical Statistics*, 6:69–93, 1956. URL: https://api.semanticscholar.org/CorpusID:55591271.