Flexible Accuracy for Differential Privacy

Aman Bansal
IIT Bombay, India

Rahul Chunduru
IIT Bombay, India

Deepesh Data
UCLA, USA

Manoj Prabhakaran
IIT Bombay, India

Abstract

Differential Privacy (DP) has become a gold standard in privacy-preserving data analysis. While it provides one of the most rigorous notions of privacy, there are many settings where its applicability is limited.

Our main contribution is in augmenting differential privacy with *Flexible Accuracy*, which allows small distortions in the input (e.g., dropping outliers) before measuring accuracy of the output, allowing one to extend DP mechanisms to high-sensitivity functions. We present mechanisms that can help in achieving this notion for functions that had no meaningful differentially private mechanisms previously. In particular, we illustrate an application to differentially private histograms, which in turn yields mechanisms for revealing the support of a dataset or the extremal values in the data. Analyses of our constructions exploit new versatile composition theorems that facilitate modular design.

All the above extensions use our new definitional framework, which is in terms of "lossy Wasserstein distance" – a 2-parameter error measure for distributions. This may be of independent interest.

1 Introduction

In the era of big data, privacy has been a major concern, to the point that recent legislative moves, like General Data Protection Regulation (GDPR) in the European Union, have mandated various measures for ensuring privacy. Further, in the face of a global pandemic that has prompted governments to collect and share individual-level information for epidemiological purposes, debates on privacy-utility trade-offs have been brought to sharper relief. Against this backdrop, mathematical theories of privacy are of great importance. Differential Privacy [Dwork et al., 2006b] is by far the most impactful mathematical framework today for privacy in statistical databases. It has seen large scale adoption in theory and practice, including machine learning applications and large scale commercial implementations (e.g., [Abadi et al., 2016, Borgs et al., 2015, 2018, Erlingsson et al., 2014, Differential Privacy Team, 2017]).

In this work, we make foundational contributions to the area of Differential Privacy (DP), extending its applicability. Our main contribution is the notion of *Flexible Accuracy* – a new framework for measuring the *accuracy* of a mechanism (while retaining the DP framework unaltered for quantifying privacy). This lets us develop new DP mechanisms with non-trivial provable (and empirically demonstrable) accuracy guarantees for high-sensitivity functions.

Motivating Flexible Accuracy (FA). Consider querying a database consisting of integer valued observations – say, ages of patients who recovered from a certain disease – for the maximum value. For the sake of privacy, one may wish to apply a DP mechanism, rather than output the maximum in the data itself. Two possible datasets which differ in only one patient are considered neighbors and a DP mechanism needs to make the outputs on these two samples indistinguishable from each other. However, the function in question is *highly sensitive* – two neighboring datasets can have their maxima differ by as much as the entire range of possible ages[1] – and, as we shall see in our empirical evaluations in Section 6, the various kinds of mechanisms in the literature [McSherry and Talwar, 2007, Bun et al., 2019, Vadhan, 2017, Dwork and Lei, 2009, Nissim et al., 2007, Beimel et al., 2016] do not provide a satisfactory solution.

Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS) 2022, Valencia, Spain. PMLR: Volume 151.

[1] In fact, *all datasets* with low maximum values have high sensitivity *locally*, by considering a neighboring dataset with a single additional data item with a large value.

The difficulty in solving this problem is related to another issue. Consider the problem of reporting a *histogram* (again, say, of patients' ages). Here a standard DP mechanism, of adding a zero-mean Laplace noise to each bar of the histogram is indeed reasonable, as the histogram function has low sensitivity in each bar. Now, note that *maximum can be computed as a function of the histogram*. However, even though the histogram mechanism was sufficiently accurate in the standard sense, the maximum computed from its output is no longer accurate! This is because when a non-zero count is added to a large-valued item with original count 0, the maximum can increase arbitrarily.

Flexible Accuracy (FA) is a relaxed notion of accuracy that lets us address both of the above issues. In particular, it not only enables new DP mechanisms for maximum, but also allows one to derive the mechanism from a new DP mechanism for histograms. We provide a general *composition theorem* that enables such transfer of accuracy guarantees that is not applicable to conventional accuracy measures. The high-level idea of FA is to allow for some *distortion of the input* when measuring accuracy. A good example of distortion is *dropping a few items* from the dataset; note that in this case, *adding* a data item is *not* considered low distortion. Referring back to the example of reporting maximum, given a dataset with a single elderly patient and many young patients, FA with respect to (w.r.t.) this distortion allows a mechanism for maximum to report the maximum age of the younger group.[2]

Flexible accuracy needs to account for errors that can be attributed to distortion of the input (input error), as well as to inaccuracies in the output (output error). To be able to exploit input distortion while retaining privacy, we allow input distortion to be randomized. A side-effect of this is that our measure of output accuracy needs to allow the "correct output" to be randomized (i.e., defined by a distribution), even if we are interested in only deterministic functions. To generalize the conventional *probabilistically approximately correct* (PAC) guarantees to this setting, we introduce a natural, but new quantity called *lossy ∞-Wasserstein distance*. Our final definition of flexible accuracy is a 3-parameter quantity, with one parameter accounting for input distortion, and 2 parameters used for output error measured using lossy ∞-Wasserstein distance.

Our contributions. These are three folds:

• *Definitions:* We present a conceptual enhancement to the framework of DP – *flexible accuracy*; see Definition 3. In founding a solid mathematical formalization of this concept, we define *lossy ∞-Wasserstein distance*; see Definition 1. This extends the classical notion of Wasserstein distance (or Earth Mover Distance), along with several existing notions, such as the PAC guarantee, the total variation distance, etc.

• *Composition Theorems:* We present a composition theorem for flexible accuracy (Theorem 1), which involves identifying new quantities, including *distortion sensitivity* (Definition 4) and *error sensitivity* (Definition 5). We also present a new pre-processing theorem for DP (Theorem 2).

• *Mechanisms:* We give a DP mechanism with FA guarantee for releasing a sanitized histogram (called the Shifted-Truncated Laplace mechanism; see Algorithm 1 and Algorithm 3), which, via our composition theorems, yield DP mechanisms with FA guarantees for *histogram-based statistics* (see Theorem 5). These include several high-sensitivity functions, such as maximum and minimum, support of a set, range, median, maximum margin separator, etc. (we give concrete bounds for max/min and support). Our empirical comparison against state-of-the-art DP mechanisms reveals that apart from the theoretical guarantees we obtain (where none were available till now), our mechanisms compare favorably in terms of accuracy (flexible and otherwise) empirically as well.

The surprising power of flexible accuracy. Consider a sequence of $n+1$ neighboring histograms, such that the first in the sequence has all its n elements in the first bar, and the last one has all elements in the last bar, and the first and the last bars are far away from each other. In any reasonably accurate (flexible or not) mechanism for a histogram-based statistic like max, the answers for these two extremes must be very different with probability almost 1. So, intuitively, there should be some pair of neighbors in this sequence for which the answers should be significantly different with probability at least $1/n$. This seems to preclude obtaining (ϵ, δ)-DP for a small constant ϵ with $\delta \ll 1/n$. Remarkably, this intuition turns out to be wrong! By carefully calibrating the probability of the responses (while also making sure that the responses can be attributed to only dropping a few items – as permitted by flexible accuracy), our mechanism can obtain the following guarantee for the max function:

Informal result for max: Our flexibly-accurate mechanism for max over a bounded range achieves $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\alpha n)}\right)$-DP while incurring an arbitrarily small output error after dropping only αn elements.

The above result gives a trade-off between the privacy guarantee and number of elements dropped. For example: (i) By choosing $\epsilon = \frac{1}{n^{1/4}}$ and $\alpha = \frac{1}{\sqrt{n}}$, our

[2] Of course, it is not obvious what should determine which items should be dropped and with what probability. This will be the subject of our new mechanisms.

mechanism is $(\frac{1}{n^{1/4}}, e^{-\Omega(n^{1/4})})$-DP while dropping only $O(\sqrt{n})$ elements. (ii) By choosing ϵ to be a small constant (say, 0.1) and say, $\alpha = \frac{\log^2 n}{n}$, our mechanism is $(0.1, n^{-\Omega(\log n)})$-DP while dropping only $O(\log^2 n)$ elements. See Appendix H.4 for several other parameter choices that are of interest.

Significance of the New Mechanisms. Traditional DP literature has largely not addressed functions like the maximum function, $f_{\max}$, due to the very high sensitivity of such functions: When the database has entries from $[0, B]$, the sensitivity of $f_{\max}$ is B.[3] The same holds for other functions like a "thresholded maximum" $\max_k$, which outputs the maximum value that appears at least k times in the database. With FA, *for the first time, we provide DP mechanisms for such functions, with meaningful worst-case accuracy guarantees.* We emphasize that we retain the *standard* definition of (ϵ, δ)-DP, and achieve strong parameters for it (see above). Further, the additional dimension of inaccuracy that we allow – namely, input distortion – is in line with what applications like (robust) Machine Learning often anticipate and tolerate.

Related work. DP, defined by Dwork et al. [2006b] has developed into a highly influential framework for providing formal privacy guarantees; see [Dwork and Roth, 2014] for more details. The notion of flexible accuracy we define is motivated by the difficulty in handling outliers in the data. Some of the work leading to DP explicitly attempts to address the privacy of outliers [Chawla et al., 2005a,b], as did some of the later works within the DP framework [Dwork and Lei, 2009, Bun et al., 2019, Thakurta and Smith, 2013]. These results rely on having a distribution over the data, or responding only when the answer is a "stable value". Blum et al. [2013] introduced the notion of *usefulness*, that is motivated by similar limitations of DP as those which motivated flexible accuracy, but as explained later, is less generally applicable. Incidentally, Wasserstein distance has been used in privacy mechanisms in the Pufferfish framework [Kifer and Machanavajjhala, 2014, Song et al., 2017], but assuming a data distribution.

Several DP mechanisms for histograms are available with a variety of accuracy guarantees, as discussed in Section 6. While these mechanisms do not claim any accuracy guarantees for functions computed from histograms, on specific data distributions and for some of these mechanisms, we see that FA can be used to empirically capture meaningful accuracy guarantees.

[3] The sensitivity of a real-valued function $f : \mathcal{X} \to \mathbb{R}$ is defined by $\Delta_f := \max_{\mathbf{x},\mathbf{x}' \in \mathcal{X}: \mathbf{x} \sim \mathbf{x}'} |f(\mathbf{x}) - f(\mathbf{x}')|$. In the case of $f_{\max}$, there are neighboring databases $\mathbf{x}, \mathbf{x}'$, where $\mathbf{x}$ has all the inputs as 0 and $\mathbf{x}'$ has $n-1$ inputs as 0 but one input is B, so, $\Delta_{f_{\max}} = B$.

2 Lossy Wasserstein Distance

Central to the formalization of all the results in this work is a new notion of distance between distributions over a metric space, that we call *lossy Wasserstein distance*. Lossy Wasserstein distance generalizes the notion of Wasserstein distance [Villani, 2008], or Earth Mover Distance, which is the minimum cost of transporting probability mass ("earth") of one distribution to make it match the other. We shall use the "infinity norm" version, where the cost paid is the maximum distance any mass is transported.

Formally, consider a metric space $(\Omega, \mathfrak{d})$, where Wasserstein distance can be defined. For example, one may consider $\Omega = \mathbb{R}^n$ and the metric $\mathfrak{d}$ being an ℓ_p-metric. For $\gamma \in [0, 1]$, and distributions P, Q over the metric space $(\Omega, \mathfrak{d})$,[4] we define $\Phi^\gamma(P, Q)$, the set of *γ-lossy couplings of P and Q*, as consisting of joint distributions ϕ over Ω^2 with its first marginal ϕ_1 and second marginal ϕ_2 such that $\Delta(\phi_1, P) + \Delta(\phi_2, Q) \leq \gamma$, where $\Delta(P, Q) := \frac{1}{2}\int_\Omega |P(\omega) - Q(\omega)|\, \mathrm{d}\omega$ denotes the total variation distance between P and Q.

Definition 1 (γ-Lossy ∞-Wasserstein Distance)**.** Let P and Q be two distributions over a metric space $(\Omega, \mathfrak{d})$. For $\gamma \in [0, 1]$, the γ-lossy ∞-Wasserstein distance between P and Q is defined as:

$$W^\infty_\gamma(P, Q) = \inf_{\phi \in \Phi^\gamma(P,Q)} \sup_{(x,y) \leftarrow \phi} \mathfrak{d}(x, y). \tag{1}$$

For simplicity, we write $W^\infty(p, q)$ to denote $W^\infty_0(p, q)$. We remark that while our definition of W^∞_γ uses a worst case notion of distance (as signified by ∞), there is an analogous average case definition, that may be of independent interest. We define this in Appendix A.2.

Now we show that the Lossy ∞-Wasserstein distance generalizes some existing notions.

• **Generalizing the PAC guarantee:** The PAC guarantee states that a randomized quantity G is, except with some small probability γ, within an approximation radius β of a desired *deterministic* quantity f: i.e., $\Pr_{g \leftarrow G}[\mathfrak{d}(f, g) > \beta] \leq \gamma$. For example, when G takes values over $\mathbb{R}$, $\mathfrak{d}$ can be the standard difference metric over $\mathbb{R}$, i.e., $\mathfrak{d}(f, g) = |f - g|$. Representing f by a point distribution F_f, this can be equivalently written as $W^\infty_\gamma(F_f, G) \leq \beta$, where the underlying metric is $\mathfrak{d}$; see Lemma 6 in Appendix A.3 for a proof of this.

• **Generalizing the total variation distance:** It

[4] We will use upper case letters (P, Q, X, Y, etc.) to denote random variables (r.v.), as well as the probability distributions associated with them.

can be shown that $W_\gamma^\infty(P,Q) = 0$ iff $\Delta(P,Q) \le \gamma$; see Lemma 7 in Appendix A.3 for a proof of this.

Triangle inequality for lossy Wasserstein distance. W_γ^∞ satisfies the following triangle inequality.

Lemma 1. *For distributions P, Q, and R over a metric space $(\Omega, \mathfrak{d})$ and for all $\gamma_1, \gamma_2 \in [0,1]$, we have*

$$W_{\gamma_1+\gamma_2}^\infty(P,R) \le W_{\gamma_1}^\infty(P,Q) + W_{\gamma_2}^\infty(Q,R). \quad (2)$$

We can easily prove Lemma 1 for the special case when $\gamma_1 = \gamma_2 = 0$ using standard tools from [Villani, 2008]; see Lemma 3 in Appendix A.1 for a proof. However, proving Lemma 1 in its full generality requires a significantly more involved proof; see Appendix A.1.

3 Flexible Accuracy (FA)

The high-level idea of FA is to allow for some *distortion of the input* before measuring accuracy. We would like to define "natural" distortions of a database, that are meaningful for the function in question. For many functions, removing a few data points (say, outliers) would be a natural distortion, while for others, perturbing the data points (or a combination of both) is more natural. Note that *adding* new entries – even just one – is often not a reasonable distortion. Therefore, distortion is generally defined not using a metric over databases, but a *quasi-metric* (which is not required to be symmetric) with range $\mathbb{R}_{\ge 0} \cup \{\infty\}$, where ∞ indicates that one database cannot be distorted into another one. As we shall need distortion measure between two distributions in our accuracy guarantees and also for defining distortion and error sensitivities, it will be useful to extend the distortion measure to distributions. This can be done in same way as W^∞, but w.r.t. a quasi-metric rather than a metric.

Definition 2 (Measure of Distortion)**.** A *measure of distortion* on a set $\mathcal{X}$ is a function $\partial : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\ge 0} \cup \{\infty\}$ which forms a quasi-metric over $\mathcal{X}$. We define $\widehat{\partial}$ as the extension of ∂ to distribution, which maps a pair of distributions P, Q over $\mathcal{X}$ to a real number as

$$\widehat{\partial}(P,Q) := \inf_{\phi \in \Phi^0(P,Q)} \sup_{(x,y)\leftarrow\phi} \partial(x,y).$$

It is easy to verify that if ∂ is a quasi-metric, so is $\widehat{\partial}$. We prove this in Lemma 13 in Appendix K.

Examples of measures of distortion. We formally define three measures of distortion: ∂_{drop} for dropping elements, ∂_{move} for perturbing/moving elements, and $\partial_{\text{drmv}}^\eta$ for a combination of dropping and moving elements. These are defined when each element in $\mathcal{X}$ is a finite multiset over a ground set $\mathcal{G}$. Formally, $\mathbf{x} \in \mathcal{X}$ is a function $\mathbf{x} : \mathcal{G} \to \mathbb{N}$ that outputs the multiplicity of each element of $\mathcal{G}$ in $\mathbf{x}$. Due to lack of space, we define ∂_{move} and $\partial_{\text{drmv}}^\eta$ in Appendix I and only define ∂_{drop} here. For finite $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, we define $\partial_{\text{drop}}(\mathbf{x}, \mathbf{x}')$ as

$$\begin{cases} \frac{\sum_{g\in\mathcal{G}} \mathbf{x}(g) - \mathbf{x}'(g)}{\sum_{g\in\mathcal{G}} \mathbf{x}(g)} & \text{if } \forall g \in \mathcal{G}, \mathbf{x}(g) \ge \mathbf{x}'(g), \\ \infty & \text{otherwise.} \end{cases} \quad (3)$$

$\partial_{\text{drop}}(\mathbf{x}, \mathbf{x}')$ measures the fraction of elements in $\mathbf{x}$ that are to be dropped to get $\mathbf{x}'$ (unless $\mathbf{x}'$ cannot be derived thus). It is easy to see that ∂_{drop} is a quasi-metric.[5]

Most of the results in this paper are derived w.r.t. ∂_{drop}, but they can also be extended to $\partial_{\text{drmv}}^\eta$; see Appendix I for the extension.

Defining flexible accuracy. Informally, flexible accuracy with distortion α guarantees that on an input $\mathbf{x}$, a mechanism shall produce an output that corresponds to $f(\mathbf{x}')$ for some $\mathbf{x}'$ such that $\partial(\mathbf{x}, \mathbf{x}') \le \alpha$. In addition to such input distortion, we may allow the output to be also probably approximately correct, with an approximation error parameter β and an error probability parameter γ. Formally, the probabilistic approximation guarantee of the output is given as a bound of β on a γ-lossy ∞-Wasserstein distance.

Definition 3 ((α,β,γ)-accuracy)**.** Let ∂ be a measure of distortion on a set $\mathcal{X}$ and $f : \mathcal{X} \to \mathcal{Y}$ be a randomized function such that $\mathcal{Y}$ admits a metric. A mechanism $\mathcal{M}$ is (α,β,γ)-*accurate for f w.r.t. ∂*, if

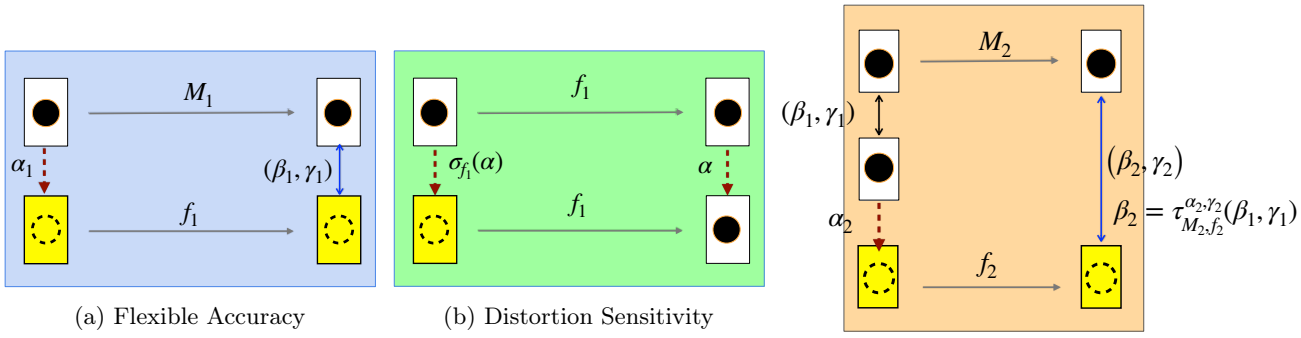
$$\sup_{\mathbf{x}\in\mathcal{X}} \inf_{X' : \widehat{\partial}(\mathbf{x}, \mathfrak{p}_{X'}) \le \alpha} W_\gamma^\infty(\mathcal{M}(\mathbf{x}), f(X')) \quad \le \quad \beta. \quad (4)$$

In other words, for each $x \in \mathcal{X}$, there is a r.v. X' satisfying $\widehat{\partial}(\mathbf{x}, \mathfrak{p}_{X'}) \le \alpha$ (i.e., $\partial(\mathbf{x}, \mathbf{x}') \le \alpha$ for all $\mathbf{x}' \in \text{support}(X')$) such that $W_\gamma^\infty(\mathcal{M}(\mathbf{x}), f(X')) \le \beta$.

See Figure 1a for an illustration of flexible accuracy using a pebbling game.

Flexible accuracy generalizes existing accuracy definitions. It should be noted that FA is not a completely disparate notion but a more generalized form of the standard accuracy guarantees. In particular: (i) As mentioned in Section 2, $(0,\beta,\gamma)$-accuracy already extends the PAC guarantees. For example, the Laplace mechanism (see [Dwork and Roth, 2014, Chapter 3]) for a function $f : \mathcal{X} \to \mathbb{R}$ that achieves ϵ-DP is $(0, \frac{\nabla_f}{\epsilon}\ln(1/\gamma), \gamma)$-accurate for any $\gamma > 0$, where ∇_f is the sensitivity of f. (ii) Blum et al. [2013] introduced *usefulness* to measure accuracy w.r.t. a "perturbed"

[5] While showing that ∂_{drop} is a quasi-metric is trivial, it is not always so with other measures of distortion; in particular, showing that $\partial_{\text{drmv}}^\eta$ is a quasi-metric is highly non-trivial; see Appendix I.1.



(a) Flexible Accuracy (b) Distortion Sensitivity (c) Error Sensitivity

Figure 1: An illustration of the flexible accuracy, distortion sensitivity, and error sensitivity. Dotted arrows indicate closeness in terms of distortion between histograms (or distributions thereof), and the solid two-sided arrows indicate closeness in terms of the lossy Wasserstein distance. Each figure shows the corresponding guarantee (accuracy, error sensitivity or distortion sensitivity) as a pebbling game: The white boxes with black pebbles correspond to given histograms, and the yellow boxes indicate histograms that are guaranteed to exist, such that the given closeness relations hold. This allows those boxes to be pebbled. Accuracy guarantee of $M_2 \circ M_1$ is derived by first applying the pebbling rule of accuracy of M_1 (to obtain the purple pebbles), then that of the error sensitivity of M_2 (to get the pink pebbles) and finally using the pebbling rule of the distortion sensitivity of f_1 to pebble the remaining yellow box.

function. While adequate for the function classes they considered (half-space queries, range queries etc.), it is not applicable to queries like maximum. FA generalizes usefulness (see Appendix B).

As we show later, FA lets us develop DP mechanisms for highly sensitive functions (*e.g.*, max), for which existing DP mechanisms offered only limited, if not vacuous, guarantees.

4 Flex. Accuracy Under Composition

In order to give our composition theorem for flexible accuracy, we need to define two new sensitivity notions: *distortion sensitivity* for a function and *error sensitivity* for a mechanism.

Distortion sensitivity. When we compose two flexibly accurate mechanisms M_1 and M_2 for $f_1 : A \to B$ and $f_2 : B \to C$, respectively, to obtain the flexible accuracy guarantee of $M_2 \circ M_1$ for $f_2 \circ f_1 : A \to C$, we would like to attribute all the distortion made in A and B (for measuring the output error of M_1 and M_2, respectively) to the distortion in A. This requires transferring the input distortion from B back into A, and the notion of distortion sensitivity allows us to quantify this. Informally, distortion sensitivity of a function f (denoted by σ_f) captures the amount of distortion required in the domain of f to capture a certain amount of distortion in the codomain of f.

Definition 4 (Distortion sensitivity)**.** Let $f : A \to B$ be a randomized function where B admits Wasserstein distance. Let ∂_1, ∂_2 be measures of distortion on A, B, respectively. Then, the *distortion-sensitivity* of f w.r.t. (∂_1, ∂_2) is defined as the function $\sigma_f : \mathbb{R}_{\geq 0} \cup \{\infty\} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ given by

$$\sigma_f(\alpha) = \sup_{\substack{x,Y: \\ \widehat{\partial}_2(f(x),\mathfrak{p}_Y) \leq \alpha}} \inf_{\substack{X: \\ f(X)=Y}} \widehat{\partial}_1(x, \mathfrak{p}_X) \qquad (5)$$

where $x \in A$, and the random variables X and Y are distributed over A and B, respectively. Above, infimum over an empty set is defined to be ∞.

See Figure 1b for an illustration of distortion sensitivity using a pebbling game.

• *Distortion sensitivity at* $\alpha = 0$. It is easy to verify that for any randomized map f, we have $\sigma_f(0) = 0$.

• *Distortion sensitivity of deterministic bijective functions:* When $f : A \to B$ is a deterministic and bijective map, then we can simplify the expression of σ_f as follows (see Appendix C.1 for more details):

$$\sigma_f(\alpha) = \sup_{\substack{x \in A, y \in B: \\ \partial_2(f(x),y) \leq \alpha}} \partial_1(x, f^{-1}(y)). \qquad (6)$$

In particular, if $f : A \to A$ is an identity function and $\partial_1 = \partial_2$, then we have $\sigma_f(\alpha) \leq \alpha$. Many of our flexibly accurate mechanisms in this paper are given for the identity function over the space of histograms.

• *A relaxed definition of distortion sensitivity.* σ_f may not exist for many functions, though we didn't encounter such situations in our paper. To accommodate more functions, we define its relaxation in (18)

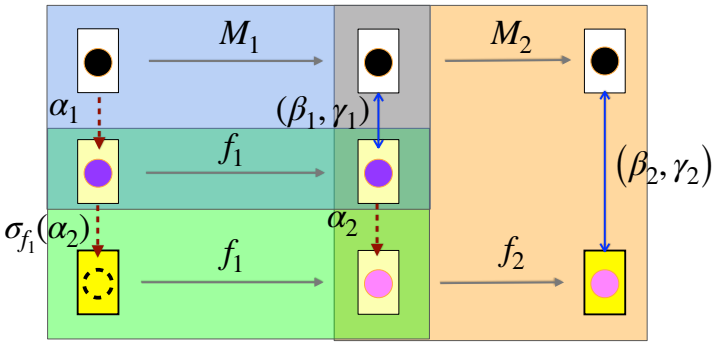


Figure 2: An illustration of the composition theorem, Theorem 1. Accuracy guarantee of $M_2 \circ M_1$ is derived by first applying the pebbling rule of accuracy of M_1 (to obtain the purple pebbles), then that of the error sensitivity of M_2 (to get the pink pebbles), and finally using the pebbling rule of the distortion sensitivity of f_1 to pebble the remaining yellow box. The final parameters are $\alpha = \alpha_1 + \sigma_{f_1}(\alpha_2)$, $\beta = \tau^{\alpha_2,\gamma_2}_{M_2,f_2}(\beta_1,\gamma_1)$, and $\gamma = \gamma_2$.

in Appendix C.1. All the results in this paper can be extended to work with this relaxed definition.

Error sensitivity. When we compose two flexibly accurate mechanisms M_1, M_2, first we measure the output error of M_1 on input x in terms of $W^{\infty}_{\gamma_1}(M_1(x), f_1(X'))$, where X' is a distortion of the *same* x on which we run the mechanism M_1, and then measure the output error of the composed mechanism by $W^{\infty}_{\gamma}(M_2(M_1(x)), f_2(Y))$, where Y is a distortion of $f_1(X')$. This is not directly computable from the FA guarantees of M_2 because the input (distribution) $f(X')$ that we distort is *not the same* as the input (distribution) $M_1(x)$ that we run M_2 on. The error sensitivity generalizes the measure of accuracy (output error) of a flexible accurate mechanism when the input (distribution) to the mechanism is not the same as the input (distribution) that we distort, but they are at a bounded distance from each other (as measured in terms on the lossy ∞-Wasserstein distance). More details on the motivation are given in Appendix C.2.

Definition 5 (Error sensitivity)**.** Let $\mathcal{M} : B \to C$ be a mechanism for $f : B \to C$, where B, C admit Wasserstein distance. Let $\widehat{\partial}$ be a measure of distortion on B. Then, for $\alpha_2, \gamma_2 \geq 0$, the error-sensitivity $\tau^{\alpha_2,\gamma_2}_{\mathcal{M},f} : \mathbb{R}_{\geq 0} \times [0,1] \to \mathbb{R}_{\geq 0}$ of $\mathcal{M}$ w.r.t. f is defined as:

$$\tau^{\alpha_2,\gamma_2}_{\mathcal{M},f}(\beta_1,\gamma_1) = \sup_{(X,X')\in\mathcal{U}} \inf_{Y\in\mathcal{V}} W^{\infty}_{\gamma_2}(\mathcal{M}(X), f(Y)), \quad (7)$$

where $\mathcal{U} = \{(X,X') : W^{\infty}_{\gamma_1}(\mathfrak{p}_X, \mathfrak{p}_{X'}) \leq \beta_1\}$ and $\mathcal{V} = \{Y : \widehat{\partial}(X', Y) \leq \alpha_2\}$.

See Figure 1c for an illustration of error sensitivity using a pebbling game.

Remark 1. *As mentioned earlier, the notion of error sensitivity generalizes the definition of flexible accuracy. In other words, if a mechanism $\mathcal{M}$ for computing a function f is (α,β,γ)-accurate, then $\beta = \tau^{\alpha,\gamma}_{\mathcal{M},f}(0,0)$.*

We can simplify the $\tau^{\alpha_2,\gamma_2}_{\mathcal{M},f}(\beta_1,\gamma_1)$ expression in some special cases that arise in our applications in Section 5; we discuss these after stating Theorem 1 next.

Composition theorem for flexible accuracy.

Our composition theorem for flexible accuracy is given below, and we prove it in Appendix D.

Theorem 1 (Flexible Accuracy Composition)**.** *Let $\mathcal{M}_1 : A \to B$ and $\mathcal{M}_2 : B \to C$ be mechanisms, respectively, with $(\alpha_1,\beta_1,\gamma_1)$-accuracy for $f_1 : A \to B$ and $\tau_{\mathcal{M}_2,f_2}$ error sensitivity for $f_2 : B \to C$, w.r.t. measures of distortion ∂_1, ∂_2 defined on A, B and metrics $\mathfrak{d}_1, \mathfrak{d}_2$ defined on B, C, respectively. Suppose f_1, α_2 are such that $\sigma_{f_1}(\alpha_2)$ is finite. Then, for any $\alpha_2 \geq 0$ and $\gamma_2 \in [0,1]$, the mechanism $\mathcal{M}_2 \circ \mathcal{M}_1 : A \to C$ is (α,β,γ)-accurate for $f_2 \circ f_1$ w.r.t. ∂_1 and $\mathfrak{d}_2$, where $\alpha = \alpha_1 + \sigma_{f_1}(\alpha_2)$, $\beta = \tau^{\alpha_2,\gamma_2}_{\mathcal{M}_2,f_2}(\beta_1,\gamma_1)$, and $\gamma = \gamma_2$.*

An illustration of how the composition theorem works is given as a pebbling game in Figure 2.

Simplified error sensitivity in spacial cases. Theorem 1 requires computing/bounding the error sensitivity of $\mathcal{M}_2$ in order to compute the FA parameter β of $\mathcal{M}_2 \circ \mathcal{M}_1$. Now we show that the expression of error sensitivity can be simplified in some important special cases: (i) **When $\mathcal{M}_1, f_1$ are deterministic maps and $\mathcal{M}_1$ is $(0,\beta_1,0)$-accurate.** In this case, it suffices to take the supremum in (7) over $x, x' \in \mathcal{B}$ such that $\mathfrak{d}_{\mathcal{B}}(x,x') \leq \beta_1$; see Appendix C.3 for more details. This setting arises when we compute the FA parameters of our bucketed histogram mechanism $\mathcal{M}_{\text{BucHist}} = \mathcal{M}_{\text{STLap}} \circ \mathcal{M}_{\text{buc}}$ (Algorithm 3) while proving Theorem 4. (ii) **When $\mathcal{M}_2, f_2$ are deterministic maps and $\mathcal{M}_1$ is $(\alpha_1,\beta_1,0)$-accurate and $\alpha_2 = \gamma_2 = 0$.** This setting arises in the case of histogram-based-statistics in Section 5.2 with $\beta_2 = 0$. In this case, the expression for the error sensitivity can be much simplified as shown in the lemma below, which we prove in Appendix C.4.

Lemma 2. *Let $\mathcal{M} : \mathcal{B} \to \mathcal{C}$ be a deterministic mechanism for a deterministic function $f : \mathcal{B} \to \mathcal{C}$. Then, for any $\beta_1 \geq 0$, we have*

$$\tau^{0,0}_{\mathcal{M},f}(\beta_1, 0) = \sup_{\substack{x,x'\in\mathcal{A}:\\ \mathfrak{d}_{\mathcal{B}}(x,x')\leq\beta_1}} \mathfrak{d}_{\mathcal{C}}(\mathcal{M}(x), f(x')). \quad (8)$$

4.1 A Pre-Processing Theorem for DP

A simple but very useful result in differential privacy (we formally define DP in Appendix E) is the "post-processing" theorem for DP, which states that if $\mathcal{M}_1$ is (ϵ, δ)-DP, then for any mechanism $\mathcal{M}_2$, the composed mechanism $\mathcal{M}_2 \circ \mathcal{M}_1$ would remain (ϵ, δ)-DP. We prove a "pre-processing" theorem for DP (complementing the post-processing theorem for DP), which states that if $\mathcal{M}_2$ is private, then so would $\mathcal{M}_2 \circ \mathcal{M}_1$ be, provided that $\mathcal{M}_1$ is neighborhood-preserving; see Definition 9 in Appendix E for the definition of neighborhood-preserving mechanism.

We prove the following theorem in Appendix E.

Theorem 2 (Differential Privacy under Pre-Processing)**.** *Let $\mathcal{M}_1 : A \to B$ and $\mathcal{M}_2 : B \to C$ be any two mechanisms. If $\mathcal{M}_1$ is neighborhood-preserving w.r.t. neighborhood relations $\sim_A$ and $\sim_B$ over A and B, respectively, and $\mathcal{M}_2$ is (ϵ, δ)-DP w.r.t. $\sim_B$, then $\mathcal{M}_2 \circ \mathcal{M}_1 : A \to C$ is (ϵ, δ)-DP w.r.t. $\sim_A$.*

We will require Theorem 2 to establish the DP guarantee of our $\mathcal{M}_{\text{BucHist}}$ mechanism (Algorithm 3).

5 Mechanisms with Flexible Accuracy

In this section, we propose and analyze concrete mechanisms for several important functions w.r.t. the distortion ∂_{drop}. We extend these results to other measures of distortion in Appendix I.

5.1 Histograms with Flexible Accuracy

We will derive our new histogram mechanism by solving a simpler Boolean task of privately reporting whether a given set is empty or not.

Private mechanism for determining whether a given set is empty or not. For this, the only input distortion we are allowed is to drop some elements – i.e., we cannot report an empty set as non-empty. Since we seek to limit the extent of distortion, let us add a constraint that if a set has q or more elements, then with probability 1 (or very close to 1), we should report the set as being non-empty. Let p_k denote the probability that a set of size $k \in [0, q]$ is reported as being non-empty, so that $p_0 = 0$ and $p_q = 1$.

For our scheme to be (ϵ, δ)-DP, we require

$$p_k \le p_{k+1} e^{\epsilon} + \delta, \qquad (1 - p_k) \le (1 - p_{k+1}) e^{\epsilon} + \delta,$$
$$p_{k+1} \le p_k e^{\epsilon} + \delta, \qquad (1 - p_{k+1}) \le (1 - p_k) e^{\epsilon} + \delta,$$

for $0 \le k < q$, with boundary conditions $p_0 = 0$ and $p_q = 1$. We are interested in simultaneously reducing ϵ and δ subject to the above constraints. The pareto-optimal (ϵ, δ) turn out to be given by $\delta\left(\frac{e^{(q/2)\epsilon}-1}{e^{\epsilon}-1}\right) = \frac{1}{2}$, with corresponding values of p_k being given by

$$p_k = \begin{cases} \delta\left(\frac{e^{k\epsilon}-1}{e^{\epsilon}-1}\right) & \text{if } k \le q/2, \\ 1 - p_{q-k} & \text{otherwise .} \end{cases} \tag{9}$$

Towards a private mechanism for histograms. To generalize this Boolean mechanism to a full-fledged histogram mechanism, we reinterpret it. In a histogram mechanism, where again, the distortion allowed in the input is to only drop elements, we can add a *negative noise* to the count in each "bar" of the histogram. (If the reduced count is negative, we report it as 0.) We seek a noise function such that the probability of the reported count being 0 (when the actual count is $k \in [0, q]$) is the same as that of the above mechanism reporting that a set of size k is empty, i.e., the probability of adding a noise $\nu \le -k$ should be $1 - p_k$. That is, if the noise distribution is given by the density function σ, we require that

$$\int_{-q}^{-k} \sigma(t) \cdot dt = 1 - p_k \quad \text{and} \quad \sigma(t) = 0 \text{ for } t \notin [-q, 0].$$

Substituting p_k from (9), and then differentiating w.r.t. k, we obtain the following expression for $\sigma(t)$:

$$\sigma(t) = \begin{cases} \frac{1}{1-e^{-\epsilon q/2}} \text{Lap}(t \mid -\frac{q}{2}, \frac{1}{\epsilon}), & \text{if } t \in [-q, 0], \\ 0, & \text{otherwise,} \end{cases}$$

where Lap is the Laplace distribution with mean $-\frac{q}{2}$ and scale parameter $1/\epsilon$.[6] We call $\sigma(t)$ the truncated Laplace distribution, which is equal to the (normalized) Laplace distribution with mean $-\frac{q}{2}$ and scale parameter $\frac{1}{\epsilon}$ when $t \in [-q, 0]$; otherwise equal to zero.

The Shifted-Truncated Laplace mechanism for releasing histograms with flexible accuracy. Our final histogram mechanism is derived by adding the noise $\sigma(t)$ from above with appropriate parameter q to each bar of the histogram, followed by rounding to the nearest integer (or to 0, if it is negative). Before describing the mechanism, we need some notation.

Datasets can be abstractly represented by multi-sets, and each element in the multi-set belongs to a ground set $\mathcal{G}$. Formally, a multi-set $\mathbf{x}$ over the ground set $\mathcal{G}$ is a function $\mathbf{x} : \mathcal{G} \to \mathbb{N}$ that outputs the multiplicity of elements in $\mathcal{G}$. The *size* and *support* of $\mathbf{x}$ are defined as $|\mathbf{x}| := \sum_{i \in \mathcal{G}} \mathbf{x}(i)$ and $\text{support}(\mathbf{x}) :=$

[6] The Laplace distribution over $\mathbb{R}$, with *scaling parameter* $b > 0$ and mean μ, is defined by the density function $\text{Lap}(x|\mu, b) := \frac{1}{2b} e^{\frac{-|x-\mu|}{b}}$ for all $x \in \mathbb{R}$. We denote a r.v. that is distributed according to the Laplace distribution with the scaling parameter b and mean 0 by $\text{Lap}(b)$.

Algorithm 1 Shifted and Truncated Laplace Mechanism, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$

Input: A histogram, $\mathbf{x} : \mathcal{G} \to \mathbb{N}$.
Output: A histogram, $\mathbf{y} : \mathcal{G} \to \mathbb{N}$.

1: **for all** $g \in \mathcal{G}$ **do**
2: $z_g \leftarrow \pi_q$, where $q := \tau|\mathbf{x}|$ and $\pi_q(z) = \begin{cases} \frac{1}{1-e^{-\epsilon q/2}} \text{Lap}(z \mid -\frac{q}{2}, \frac{1}{\epsilon}) & \text{if } z \in [-q, 0], \\ 0 & \text{otherwise.} \end{cases}$
3: $\mathbf{y}(g) := \max(0, \lfloor \mathbf{x}(g) + z_g \rceil)$
4: **end for**
5: $\mathbf{y}(g) := \max(0, \lfloor \mathbf{x}(g) + z_g \rceil)$.
6: Return $\mathbf{y}$.

$\{i \in \mathcal{G} : \mathbf{x}(i) \neq 0\}$, respectively. We shall be interested in finite-sized multi-sets, which we refer to as histograms. We denote the domain of all histograms over $\mathcal{G}$ by $\mathcal{H}_{\mathcal{G}}$. For DP, the standard notion of neighborhood among histograms is defined as $\mathbf{x} \sim_{\text{hist}} \mathbf{x}'$ iff $\sum_{i \in \mathcal{G}} |\mathbf{x}(i) - \mathbf{x}'(i)| \leq 1$. Later, we shall also require $\mathcal{G}$ to be a metric space, endowed with a metric $\mathfrak{d}$.

We describe our Shifted-Truncated Laplace for the identity function (denoted by $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}} : \mathcal{H}_{\mathcal{G}} \to \mathcal{H}_{\mathcal{G}}$) in Algorithm 1. It simply *decreases* the multiplicity of each element by adding a bounded quantity sampled from the truncated Laplace distribution. The following theorem states the privacy and FA guarantees achieved by $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$, and we prove it in Appendix F.

Theorem 3. *On inputs $\mathbf{x}$ of size n, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ from Algorithm 1 satisfies the following guarantees:*
- *Privacy: For any ϵ, τ such that $\epsilon\tau n \geq 2$, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)}\right)$-DP w.r.t. $\sim_{\text{hist}}$.*
- *Flexible accuracy: If $|\text{support}(\mathbf{x})| \leq t$, then for any $\epsilon > 0$, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $(\tau t, 0, 0)$-accurate for the identity function, w.r.t. the distortion measure ∂_{drop}.*

Remark 2. *There are many choices of ϵ, τ for which we get favorable privacy parameters in Theorem 3. For instance, choosing $\epsilon = \frac{1}{\sqrt{\tau n}}$ gives that $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $\left(\frac{1}{\sqrt{\tau n}}, \frac{e^{-\Omega(\sqrt{\tau n})}}{\sqrt{\tau n}}\right)$-DP, provided τ is such that $\sqrt{\tau n} \geq 2$. Note that τ is the maximum overall fraction of elements we drop from each bar of the histogram. For example, by choosing $\tau = \frac{1}{n^{1/2}}$, we get that $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $\left(\frac{1}{n^{1/4}}, \frac{e^{-\Omega(n^{1/4})}}{n^{1/4}}\right)$-DP and $(\frac{t}{n^{1/4}}, 0, 0)$-accurate. See Appendix H.4 for more parameter choices and discussion.*

Remark 2 shows that the privacy parameters of $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ improve as the database size $|\mathbf{x}|$ grows, while dropping only a small number of elements, provided that the support size t is not too large. To handle larger supports, this mechanism can be composed with a simple fixed width w bucketing mechanism.

Algorithm 2 Bucketing Mechanism, $\mathcal{M}^{w,[0,B)}_{\text{buc}}$

Input: A histogram $\mathbf{x}$ over $[0, B)$.
Output: A histogram $\mathbf{y}$ over $S = \{w(i - \frac{1}{2}) : i \in [t], t = \lceil \frac{B}{w} \rceil\}$, and $|\mathbf{y}| = |\mathbf{x}|$.

1: For all $s \in S$, set $\mathbf{y}(s) := \sum_{g : g - s \in [\frac{-w}{2}, \frac{w}{2})} \mathbf{x}(g)$
2: Return $\mathbf{y}$

Bucketed, Shifted-Truncated Laplace mechanism. In order to explain the idea behind our bucketing mechanism, for simplicity, we consider the ground set $\mathcal{G} = [0, B)$.[7] We divide the interval $[0, B)$ into $t = \lceil \frac{B}{w} \rceil$ sub-intervals (buckets) of length w, and map each input point to the center of the nearest sub-interval (bucket). This mapping of input points to the nearest bucket introduces error in the output space, and the value of w depends on the amount of error we want to tolerate in the output space. Our bucketing mechanism $\mathcal{M}^{w,[0,B)}_{\text{buc}}$ and the final bucketed-histogram mechanism $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$ are presented in Algorithm 2 and Algorithm 3, respectively.

Algorithm 3 BucketHist Mechanism, $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$

Input: A histogram $\mathbf{x}$ over $[0, B)$.
Output: A histogram $\mathbf{y}$ over $[0, B)$.

1: $w := 2\beta$, $t := \lceil \frac{B}{w} \rceil$, $\tau := \alpha/t$
2: Return $\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}} \circ \mathcal{M}^{w,[0,B)}_{\text{buc}}(\mathbf{x})$

Since $\mathcal{M}^{w,[0,B)}_{\text{buc}}$ introduces error in the output space, we need a metric over $\mathcal{H}_{[0,B)}$ to analyze its flexible accuracy. We use the following natural metic $\mathfrak{d}_{\text{hist}}$ over $\mathcal{H}_{[0,B)}$, which is defined as $\mathfrak{d}_{\text{hist}}(\mathbf{y}, \mathbf{y}') := W^{\infty}(\frac{\mathbf{y}}{|\mathbf{y}|}, \frac{\mathbf{y}'}{|\mathbf{y}'|})$. Here, $\frac{\mathbf{y}}{|\mathbf{y}|}$ is treated as a probability distribution and the underlying metric for W^{∞} is the standard distance metric over $\mathbb{R}$.

We prove the following theorem in Appendix G.

Theorem 4. *On inputs of size n, $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$ is $(\alpha, \beta, 0)$-accurate for the identity function, w.r.t. the distortion measure ∂_{drop} and metric $\mathfrak{d}_{\text{hist}}$. Furthermore, for any $\epsilon > 0$, and $\tau = \alpha(\frac{2\beta}{B})$, if $\epsilon\tau n \geq 2$, then $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$ is $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)}\right)$-DP.*

5.2 Histogram-Based-Statistics

Theorem 4 provides a powerful tool to obtain a DP mechanism for *any deterministic* histogram-based-statistic $f_{\text{HBS}} : \mathcal{H}_{[0,B)} \to \mathcal{A}$, simply by defining

$$\mathcal{M}^{\alpha,\beta,[0,B)}_{f_{\text{HBS}}} = f_{\text{HBS}} \circ \mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}. \tag{10}$$

[7] We also present the general results for $\mathcal{G} = [0, B)^d$ in Appendix J. Also see Remark 5 in Appendix I.

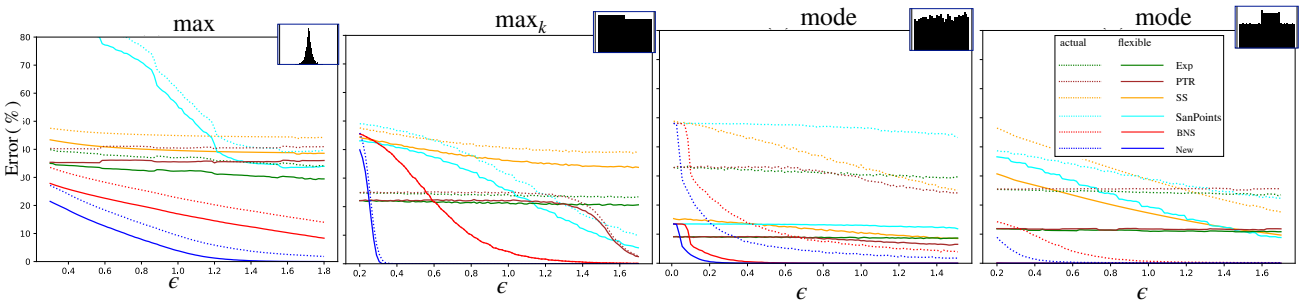


Figure 3: For each evaluation, a typical histogram used is shown in inset. Functions evaluated are $\max_k(\mathbf{x}) := \max\{i \mid \mathbf{x}(i) \geq k\}$, $\max := \max_1$, and $\text{mode}(\mathbf{x}) := \arg\max_i \mathbf{x}(i)$. For max, we used histogram of 10,000 items drawn i.i.d. from a Cauchy distribution with median 45 and scale 4, restricted to 100 bars, with the last 10 set to empty bars. For $\max_{500}$, we used step histogram with 100 bars, with two steps (height $\times$ width): $[540 \times 50, 490 \times 50]$. For mode ($3^{\text{rd}}$ figure), we used histogram of 30 bars, each bar has height drawn from i.i.d. Poisson with mean 250. For mode (4^{th} figure), we used noisy step histogram, with steps $[130 \times 120, 200 \times 5, 185 \times 85, 190 \times 10, 130 \times 80]$. All cases use best parameters available for ϵ shown in the x-axis and $\delta = 2^{-20}$. Errors are shown in the y-axis as a percentage of the full range. Actual error (dotted lines) and flexible errors (solid lines) allowing ∂_{drop} distortion of 0.005 are plotted. While our mechanism (New) yields low errors (sometimes without distortion), none of the other mechanisms achieve this consistently (even with distortion).

To analyze the flexible accuracy of $\mathcal{M}_{f_{\text{HBS}}}$, we define the *metric sensitivity* function of f_{HBS}.

Definition 6. The *metric sensitivity* of a histogram-based-statistic $f_{\text{HBS}} : \mathcal{H}_{[0,B)} \to \mathcal{A}$, is given by $\Delta_{f_{\text{HBS}}} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, in terms of a metric $\mathfrak{d}_{\mathcal{A}}$ over $\mathcal{A}$,

$$\Delta_{f_{\text{HBS}}}(\beta) = \sup_{\substack{\mathbf{x},\mathbf{x}' \in \mathcal{H}_{[0,B)}: \\ \mathfrak{d}_{\text{hist}}(\mathbf{x},\mathbf{x}') \leq \beta}} \mathfrak{d}_{\mathcal{A}}(f_{\text{HBS}}(\mathbf{x}), f_{\text{HBS}}(\mathbf{x}')). \quad (11)$$

We prove the following theorem in Appendix H.1.

Theorem 5. *On inputs of size n, $\mathcal{M}^{\alpha,\beta,[0,B)}_{f_{\text{HBS}}}$ is $(\alpha, \Delta_{f_{\text{HBS}}}(\beta), 0)$-accurate for f_{HBS} w.r.t. distortion ∂_{drop} and metric $\mathfrak{d}_{\mathcal{A}}$. Furthermore, for any $\epsilon > 0$, and $\tau = \alpha(\frac{2\beta}{B})$, if $\epsilon\tau n \geq 2$, then $\mathcal{M}^{\alpha,\beta,[0,B)}_{f_{\text{HBS}}}$ is $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)}\right)$-DP.*

Applications to high-sensitivity functions. Theorem 5 has direct applications to functions that have high sensitivity (defined w.r.t. the neighborhood relation $\sim$), but low metric sensitivity. We point out two such examples: maximum and the support functions.

Maximum function: We define $f_{\max}$ for histograms over real numbers as $f_{\max}(\mathbf{x}) := \max\{g : \mathbf{x}(g) > 0\}$. It can be shown that, the metric sensitivity of $f_{\max}$ is $\Delta_{f_{\max}}(\beta) \leq \beta$. Substituting this in Theorem 5, we get the result for $f_{\max}$. See Appendix H.2 for more details.

Support function: We define f_{supp} (or simply support) for histograms over real numbers as $f_{\text{supp}}(\mathbf{x}) := \{g : \mathbf{x}(g) > 0\}$. To measure accuracy, we use a metric $\mathfrak{d}_{\text{supp}}$ over the set of finite subsets of $\mathbb{R}$: for finite subsets $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}$, define $\mathfrak{d}_{\text{supp}}(\mathcal{U}, \mathcal{V}) := \max\left\{\max_{u\in\mathcal{U}} \min_{v\in\mathcal{V}} |u - v|, \max_{v\in\mathcal{V}} \min_{u\in\mathcal{U}} |v - u|\right\}$. $\mathfrak{d}_{\text{supp}}$ measures the farthest that a point in one of the sets is from any point on the other set. It can be shown that, the metric sensitivity of f_{supp} is $\Delta_{f_{\text{supp}}}(\beta) \leq \beta$. Substituting $\Delta_{f_{\text{HBS}}}(\beta) \leq \beta$ in Theorem 5, we get the same result for f_{supp}. See Appendix H.3 for more details.

6 Empirical Evaluation

We empirically compare our basic mechanism $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ (Algorithm 1) against competing mechanisms, for accuracy on a few histogram-based statistics computed on it. The considered mechanisms are: Exponential Mechanism [McSherry and Talwar, 2007], Propose-Test-Release Mechanism [Dwork and Lei, 2009], Smooth-sensitivity Mechanism [Nissim et al., 2007], Stability-Based Sanitized Histogram [Bun et al., 2019], and Choosing-Based Histogram [Beimel et al., 2016]. We describe these all briefly in Appendix L.2.

We plot average errors (actual and flexible), on different histograms in Figure 3. We emphasize that *for these functions*, the other mechanisms do not offer any *worst-case guarantees* (with or without flexible accuracy), while we do. We provide more details and additional comparisons in Appendix L.

Acknowledgements

The work of Deepesh Data was supported in part by NSF grants #2007714 and #2139304. The work of Manoj Prabhakaran was supported in part by the Joint Indo-Israel Project DST/INT/ISR/P-16/2017 and the Ramanujan Fellowship of Dept. of Science and Technology, India.

References

M. Abadi, A. Chu, I. Goodfellow, B. McMahan, I. Mironov, K. Talwar, and L. Zhang. Deep learning with differential privacy. In *CCS*, pages 308–318, 2016.

A. Beimel, K. Nissim, and U. Stemmer. Private learning and sanitization: Pure vs. approximate differential privacy. *Theory of Computing*, 12(1):1–61, 2016.

A. Blum, K. Ligett, and A. Roth. A learning theory approach to noninteractive database privacy. *J. ACM*, 60(2):12:1–12:25, 2013.

C. Borgs, J. T. Chayes, and A. D. Smith. Private graphon estimation for sparse graphs. In *NIPS*, pages 1369–1377, 2015.

C. Borgs, J. T. Chayes, A. D. Smith, and I. Zadik. Revealing network structure, confidentially: Improved rates for node-private graphon estimation. In *FOCS*, pages 533–543, 2018.

M. Bun, K. Nissim, and U. Stemmer. Simultaneous private learning of multiple concepts. *Journal of Machine Learning Research*, 20:94:1–94:34, 2019.

S. Chawla, C. Dwork, F. McSherry, A. D. Smith, and H. Wee. Toward privacy in public databases. In *TCC*, pages 363–385, 2005a.

S. Chawla, C. Dwork, F. McSherry, and K. Talwar. On privacy-preserving histograms. In *UAI*, 2005b.

A. Differential Privacy Team. Learning with privacy at scale. *Apple Machine Learning Journal*, December 2017. Available online: https://machinelearning.apple.com/docs/learning-with-privacy-at-scale/appledifferentialprivacysystem.pdf.

C. Dwork and J. Lei. Differential privacy and robust statistics. In *STOC*, pages 371–380, 2009.

C. Dwork and A. Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, Aug 2014. ISSN 1551-305X.

C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In *EUROCRYPT*, pages 486–503, 2006a.

C. Dwork, F. McSherry, K. Nissim, and A. D. Smith. Calibrating noise to sensitivity in private data analysis. In *TCC*, pages 265–284, 2006b.

Ú. Erlingsson, V. Pihur, and A. Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *CSS*, pages 1054–1067, 2014.

Ú. Erlingsson, V. Feldman, I. Mironov, A. Raghunathan, S. Song, K. Talwar, and A. Thakurta. Encode, shuffle, analyze privacy revisited: Formalizations and empirical evaluation. *arXiv preprint arXiv:2001.03618*, 2020.

D. Kifer and A. Machanavajjhala. Pufferfish: A framework for mathematical privacy definitions. *ACM Trans. Database Syst.*, 39(1):3:1–3:36, 2014.

F. McSherry and K. Talwar. Mechanism design via differential privacy. In *FOCS*, pages 94–103, 2007.

K. Nissim, S. Raskhodnikova, and A. D. Smith. Smooth sensitivity and sampling in private data analysis. In *STOC*, pages 75–84, 2007.

S. Song, Y. Wang, and K. Chaudhuri. Pufferfish privacy mechanisms for correlated data. In *SIGMOD*, pages 1291–1306, 2017.

A. Thakurta and A. D. Smith. Differentially private feature selection via stability arguments, and the robustness of the lasso. In *COLT*, volume 30, pages 819–850, 2013.

S. P. Vadhan. The complexity of differential privacy. In *Tutorials on the Foundations of Cryptography*, pages 347–450. Springer International Publishing, 2017.

C. Villani. *Optimal transport: old and new*. Springer Verlag, 2008.

Supplementary Material

A Details Omitted from Section 2

In all our proofs in this paper, when dealing with infimum/supremum (for example, in the definitions of the lossy Wasserstein distance, measure of distortion, distortion and error sensitivities, etc.), for simplicity, we assume that the infimum/supremum is always achieved; all our proofs can be easily extended to work without this assumption by taking appropriate limits when working with infinitesimal quantities.

A.1 Lossy ∞-Wasserstein Distance

As mentioned in Section 2, first we prove Lemma 1 for the special case when $\gamma_1 = \gamma_2 = 0$. This is easy and can be shown using standard tools from Villani [2008]. Then, we will prove Lemma 1 in its full generality, which requires a significantly more involved proof.

Lemma 3 (Lemma 1 at $\gamma_1 = \gamma_2 = 0$)**.** *For distributions P, Q, and R over a metric space $(\Omega, \mathfrak{d})$, we have*

$$W^\infty(P, R) \leq W^\infty(P, Q) + W^\infty(Q, R).$$

Proof. Let $\phi_2 \in \Phi(P, Q)$ and $\phi_3 \in \Phi(Q, R)$ denote the optimal couplings for $W^\infty(P, Q)$ and $W^\infty(Q, R)$, respectively, i.e., $W^\infty(P, Q) = \sup_{\substack{(x,y):\\ \phi_2(x,y)\neq 0}} \mathfrak{d}(x, y)$ and $W^\infty(Q, R) = \sup_{\substack{(y,z):\\ \phi_3(y,z)\neq 0}} \mathfrak{d}(y, z)$. It follows from the Gluing Lemma Villani [2008] that we can find a coupling ϕ' over $\Omega \times \Omega \times \Omega$ such that the projection of ϕ' onto its first two coordinates is equal to ϕ_2 and its last two coordinates is equal to ϕ_3. Let ϕ_1 denote the projection of ϕ' onto its first and the third coordinates. Note that $\phi_1 \in \Phi(P, R)$, but it may not be an optimal coupling for $W^\infty(P, R)$. Now the triangle inequality follows from the following set of inequalities:

$$\begin{aligned}
W^\infty(P, R) &= \inf_{\phi\in\Phi(P,R)} \sup_{\substack{(x,z):\\ \phi(x,z)\neq 0}} \mathfrak{d}(x, z) \leq \sup_{\substack{(x,z):\\ \phi_1(x,z)\neq 0}} \mathfrak{d}(x, z) = \sup_{\substack{(x,y,z):\\ \phi'(x,y,z)\neq 0}} \mathfrak{d}(x, z) \\
&\overset{(a)}{\leq} \sup_{\substack{(x,y,z):\\ \phi'(x,y,z)\neq 0}} \mathfrak{d}(x, y) + \mathfrak{d}(y, z) \\
&= \sup_{\substack{(x,y,z):\\ \phi'(x,y,z)\neq 0}} \mathfrak{d}(x, y) + \sup_{\substack{(x,y,z):\\ \phi'(x,y,z)\neq 0}} \mathfrak{d}(y, z) \\
&= \sup_{\substack{(x,y):\\ \phi_2(x,y)\neq 0}} \mathfrak{d}(x, y) + \sup_{\substack{(y,z):\\ \phi_3(y,z)\neq 0}} \mathfrak{d}(y, z) \\
&= W^\infty(P, Q) + W^\infty(Q, R),
\end{aligned}$$

where (a) follows from the fact that $\mathfrak{d}$ is a metric, and so it satisfies the triangle inequality. □

Now we prove Lemma 1 in its full generality, and along the way derive useful properties about lossy Wasserstein distance, that may be of independent interest. For convenience, we rewrite Lemma 1 below.

Lemma (Restating Lemma 1)**.** *For distributions P, Q, and R over a metric space $(\Omega, \mathfrak{d})$ and for all $\gamma_1, \gamma_2 \in [0, 1]$, we have*

$$W^\infty_{\gamma_1+\gamma_2}(P, R) \leq W^\infty_{\gamma_1}(P, Q) + W^\infty_{\gamma_2}(Q, R). \tag{12}$$

The following lemma is crucial to proving Lemma 1.

Lemma 4. *Let P and Q be any two distributions over a metric space $(\Omega, \mathfrak{d})$. If $W^\infty_\gamma(P, Q) = \beta$, then for all $\gamma_1 \in [0, \gamma]$, there exist distributions P' and Q' s.t. $\Delta(P, P') \leq \gamma_1$, $\Delta(Q, Q') \leq \gamma - \gamma_1$, and $W^\infty(P', Q') = \beta$.*

Proof of Lemma 4. Let P and Q be any two distributions over a metric space $(\Omega, \mathfrak{d})$. Let us assume that the optimal $W^\infty_\gamma(P, Q)$ $(= \beta)$ is obtained at the joint distribution ϕ_{opt}. Let the first and the second marginal

distributions of ϕ_{opt} be P_{opt} and Q_{opt}, respectively. Let $\Delta(P, P_{opt}) = \gamma_{opt}$, which implies that $\Delta(Q, Q_{opt}) \leq \gamma - \gamma_{opt}$. Define a function $R_{opt} : \Omega \to \mathbb{R}$ as $R_{opt}(\omega) := P_{opt}(\omega) - P(\omega)$ for all $\omega \in \Omega$. Clearly, $\int_\Omega R_{opt}(\omega)\,\mathrm{d}\omega = 0$ and $\int_\Omega |R_{opt}(\omega)|\,\mathrm{d}\omega = 2\gamma_{opt}$.

In the discussion below, we shall take a general $\gamma_1 \in [0, \gamma_{opt})$ and construct distributions P' and Q' s.t. $\Delta(P, P') \leq \gamma_1$, $\Delta(Q, Q') \leq \gamma - \gamma_1$, and $W^\infty(P', Q') = \beta$, as required in the conclusion of Lemma 4. We can show a similar result for the other case also when $\gamma_1 \in (\gamma_{opt}, \gamma]$ (by swapping the roles of P and Q in the above as well as in the argument below). This will complete the proof of Lemma 4.

Define a function $R' : \Omega \to \mathbb{R}$ as $R'(\omega) := \frac{\gamma_1}{\gamma_{opt}} R_{opt}(\omega)$. For any $\omega \in \Omega$, let $P'(\omega) = P(\omega) + R'(\omega)$. After substituting the value of $R_{opt}(\omega) = P_{opt}(\omega) - P(\omega)$, we get $P'(\omega) = \frac{\gamma_1}{\gamma_{opt}} P_{opt}(\omega) + \left(1 - \frac{\gamma_1}{\gamma_{opt}}\right) P(\omega)$. Since P' is a convex combination of two distributions, it is also a valid distribution. It is easy to see that $\Delta(P, P') = \gamma_1$. Define a joint distribution ϕ' as follow: for every $(x, y) \in \Omega \times \Omega$, define

$$\phi'(x, y) := \begin{cases} \phi_{opt}(x, y) \frac{P'(x)}{P_{opt}(x)} & \text{if } P_{opt}(x) > 0 \\ P'(x)\delta(x - y) & \text{otherwise} \end{cases}$$

where $\delta(\cdot)$ is the Dirac delta function. It follows from the definition that $\int_\Omega \phi'(x, y)\,\mathrm{d}y = P'(x)$, i.e., the first marginal of ϕ' is $P'(\cdot)$. This also implies that ϕ' is a valid joint distribution because (i) $\phi'(x, y) \geq 0$ for all $(x, y) \in \Omega \times \Omega$, and (ii) $\int_{\Omega\times\Omega} \phi'(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_\Omega P'(x)\,\mathrm{d}x = 1$.

Let the second marginal of ϕ' be Q'. We show in Claim 2 in Appendix A.1.1 that $\Delta(Q, Q') \leq \gamma - \gamma_1$.

The only thing left to prove is to show that $W^\infty(P', Q') = \beta$ for the above constructed P' and Q'. First, we show $W^\infty(P', Q') \geq \beta$ and then show $W^\infty(P', Q') \leq \beta$.

- **Showing $W^\infty(P', Q') \geq \beta$:** This follows from the following claim, which we prove in Appendix A.1.1.

  Claim 1. *For distributions P and Q over a metric space $(\Omega, \mathfrak{d})$ and $\gamma \in [0, 1]$, we have*

  $$W^\infty_\gamma(P, Q) \quad = \quad \inf_{\substack{\hat{P}, \hat{Q}: \\ \Delta(P,\hat{P}) + \Delta(Q,\hat{Q}) \leq \gamma}} W^\infty(\hat{P}, \hat{Q}). \tag{13}$$

  Now, since P', Q' satisfy $\Delta(P, P') + \Delta(Q, Q') \leq \gamma$, we have $W^\infty_\gamma(P, Q) \leq W^\infty(P', Q')$. Since $W^\infty_\gamma(P, Q) = \beta$, we have shown that $W^\infty(P', Q') \geq \beta$.
- **Showing $W^\infty(P', Q') \leq \beta$:** For the sake of contradiction, let us assume that $W^\infty(P', Q') > \beta$. Then there is a pair $(x, y) \in \Omega^2$ such that $\phi'(x, y) > 0$ and $\mathfrak{d}(x, y) > \beta$. This implies that $\phi_{opt}(x, y) = 0$, because, otherwise, we would have $W^\infty_\gamma(P, Q) > \beta$, which contradicts our hypothesis that $W^\infty_\gamma(P, Q) = \beta$. So, we know that $\phi'(x, y) > 0$ and $\phi_{opt}(x, y) = 0$. From the definition of ϕ', this is only possible if $P_{opt}(x) = 0$ and $P'(x)\delta(x - y) > 0$. This can happen only if $x = y$, but this implies $\mathfrak{d}(x, y) = 0 \leq \beta$, which is a contradiction. Hence $W^\infty(P', Q') \leq \beta$.

This completes the proof of Lemma 4. □

Now we are ready to prove Lemma 1.

Proof of Lemma 1. Let $W^\infty_{\gamma_1}(P, Q) = \beta_1$ and $W^\infty_{\gamma_2}(Q, R) = \beta_2$. It follows from Lemma 4 that there exists a distribution P' such that $\Delta(P, P') \leq \gamma_1$ and $W^\infty(P', Q) = \beta_1$. Similarly, there exists a distribution R' such that $\Delta(R, R') \leq \gamma_2$ and $W^\infty(Q, R') = \beta_2$. Using these, we have from Lemma 3 that $W^\infty(P', R') \leq \beta_1 + \beta_2$.

Now, the result follows from the following set of inequalities.

$$W^\infty_{\gamma_1+\gamma_2}(P, R) \overset{\text{(d)}}{=} \inf_{\substack{\hat{P}, \hat{R}: \\ \Delta(P,\hat{P}) + \Delta(R,\hat{R}) \leq \gamma_1+\gamma_2}} W^\infty(\hat{P}, \hat{R}) \overset{\text{(e)}}{\leq} W^\infty(P', R') \leq \beta_1 + \beta_2 = W^\infty_{\gamma_1}(P, Q) + W^\infty_{\gamma_2}(Q, R),$$

where (d) follows from Claim 1 and (e) follows because P', R' satisfy $\Delta(P, P') + \Delta(R, R') \leq \gamma_1 + \gamma_2$.

This concludes the proof of Lemma 1. □

A.1.1 Omitted Details from the Proof of Lemma 4

Claim 2. $\Delta(Q, Q') \leq \gamma - \gamma_1$.

Proof. The claim follows from the following set of inequalities.

$$\begin{aligned}
\Delta(Q, Q') &\leq \Delta(Q, Q_{opt}) + \Delta(Q_{opt}, Q') \\
&\leq \gamma - \gamma_{opt} + \frac{1}{2}\int_{\Omega} |Q_{opt}(y) - Q'(y)| \, \mathrm{d}y && \text{(Since } \Delta(Q, Q_{opt}) \leq \gamma - \gamma_{opt}\text{)} \\
&= \frac{1}{2}\int_{\Omega} \left| \int_{\Omega} \phi_{opt}(x, y) \, \mathrm{d}x - \int_{\Omega} \phi'(x, y) \, \mathrm{d}x \right| \mathrm{d}y + (\gamma - \gamma_{opt}) \\
&\leq \frac{1}{2}\int_{\Omega}\int_{\Omega} |\phi_{opt}(x, y) - \phi'(x, y)| \, \mathrm{d}x \, \mathrm{d}y + (\gamma - \gamma_{opt})
\end{aligned}$$

Define $\Omega_1 := \{x \in \Omega : P_{opt}(x) > 0\}$ and $\overline{\Omega}_1 := \Omega \setminus \Omega_1$. Since $P_{opt}(x) = 0$ for all $x \in \overline{\Omega}_1$ and $P_{opt}()$ is the first marginal of ϕ_{opt}, we have that $\phi_{opt}(x, y) = 0$ for all $x \in \overline{\Omega}_1$ and $y \in \Omega$. Now, continuing from above, we get

$$\begin{aligned}
\Delta(Q, Q') &\leq \frac{1}{2}\int_{\Omega}\int_{x \in \Omega_1} |\phi_{opt}(x, y) - \phi'(x, y)| \, \mathrm{d}x \, \mathrm{d}y + \frac{1}{2}\int_{\Omega}\int_{x \in \overline{\Omega}_1} |\phi_{opt}(x, y) - \phi'(x, y)| \, \mathrm{d}x \, \mathrm{d}y + (\gamma - \gamma_{opt}) \\
&= \frac{1}{2}\int_{\Omega}\int_{\Omega_1} \phi_{opt}(x, y) \left| 1 - \frac{P'(x)}{P_{opt}(x)} \right| \mathrm{d}x \, \mathrm{d}y + \frac{1}{2}\int_{\Omega}\int_{\overline{\Omega}_1} |\phi'(x, y)| \, \mathrm{d}x \, \mathrm{d}y + (\gamma - \gamma_{opt}) \\
&= \frac{1}{2}\int_{\Omega_1} \left| 1 - \frac{P'(x)}{P_{opt}(x)} \right| \mathrm{d}x \int_{\Omega} \phi_{opt}(x, y) \, \mathrm{d}y + \frac{1}{2}\int_{\Omega}\int_{\overline{\Omega}_1} P'(x)\delta(x - y) \, \mathrm{d}x \, \mathrm{d}y + (\gamma - \gamma_{opt}) \\
&\qquad \text{(Since } \phi'(x, y) = P'(x)\delta(x - y) \text{ for } x \in \overline{\Omega}_1\text{)} \\
&= \frac{1}{2}\int_{\Omega_1} |P_{opt}(x) - P'(x)| \, \mathrm{d}x + \frac{1}{2}\int_{\overline{\Omega}_1} P'(x) \, \mathrm{d}x + (\gamma - \gamma_{opt}) \\
&\qquad \text{(Since } \textstyle\int_{\Omega} \phi_{opt}(x, y) \, \mathrm{d}y = P_{opt}(x) \text{ and } \int_{\Omega} \delta(x - y) \, \mathrm{d}y = 1 \text{ for any } x\text{)} \\
&= \frac{1}{2}\int_{\Omega_1} |P_{opt}(x) - P'(x)| \, \mathrm{d}x + \frac{1}{2}\int_{\overline{\Omega}_1} |P_{opt}(x) - P'(x)| \, \mathrm{d}x + (\gamma - \gamma_{opt}) \\
&\qquad \text{(Since } P_{opt}(x) = 0 \text{ whenever } x \in \overline{\Omega}_1\text{)} \\
&= \frac{1}{2}\int_{\Omega} |P_{opt}(x) - P'(x)| \, \mathrm{d}x + (\gamma - \gamma_{opt}) \\
&\overset{(a)}{=} \frac{1}{2}\int_{\Omega} \left| \left(1 - \frac{\gamma_1}{\gamma_{opt}}\right) R_{opt}(x) \right| \mathrm{d}x + (\gamma - \gamma_{opt}) \\
&= \frac{(\gamma_{opt} - \gamma_1)}{2\gamma_{opt}} \int_{\Omega} |R_{opt}(x)| \, \mathrm{d}x + (\gamma - \gamma_{opt}) \\
&= \gamma_{opt} - \gamma_1 + (\gamma - \gamma_{opt}) && \text{(Since } \textstyle\int_{\Omega} |R_{opt}(\omega)| \, \mathrm{d}\omega = 2\gamma_{opt}\text{)} \\
&= \gamma - \gamma_1 && (14)
\end{aligned}$$

Here (a) follows because for every $x \in \Omega$, we have $P_{opt}(x) - P'(x) = R_{opt}(x) + P(x) - P'(x) = R_{opt}(x) - R'(x) = R_{opt}(x) - \frac{\gamma_1}{\gamma_{opt}} R_{opt}(x)$. □

Claim (Restating Claim 1)**.** *For distributions P and Q over a metric space $(\Omega, \mathfrak{d})$ and $\gamma \in [0, 1]$, we have*

$$W^{\infty}_{\gamma}(P, Q) \quad = \quad \inf_{\substack{\hat{P}, \hat{Q}: \\ \Delta(P, \hat{P}) + \Delta(Q, \hat{Q}) \leq \gamma}} W^{\infty}(\hat{P}, \hat{Q}).$$

Proof. This claim simply follows by by viewing the infimum set in the definition of γ-Lossy ∞-Wasserstein distance differently.

$$W^{\infty}_{\gamma}(P, Q) \quad \overset{(a)}{=} \inf_{\phi \in \Phi^{\gamma}(P, Q)} \max_{(x, y) \leftarrow \phi} \mathfrak{d}(x, y)$$

$$\overset{(b)}{=} \inf_{\substack{\hat{P},\hat{Q}:\\ \Delta(P,\hat{P})+\Delta(Q,\hat{Q})\le\gamma}} \inf_{\phi\in\Phi^0(\hat{P},\hat{Q})} \max_{(x,y)\leftarrow\phi} \mathfrak{d}(x,y)$$

$$\overset{(c)}{=} \inf_{\substack{\hat{P},\hat{Q}:\\ \Delta(P,\hat{P})+\Delta(Q,\hat{Q})\le\gamma}} W^\infty(\hat{P},\hat{Q}).$$

where (a) follows from the definition of γ-Lossy ∞-Wasserstein distance; (b) trivially holds by viewing the infimum set differently; in (c) we substituted the definition of W_∞; and (d) follows because P', Q' satisfies $\Delta(P,P') + \Delta(Q,Q') \le \gamma$. □

A.2 Average Version of Lossy Wasserstein Distance

Our definition of W_θ^∞ uses a worst case notion of distance. Many of the results using this notion have analogues using an average case version. We formally present this definition below, as it may be of interest elsewhere.

Definition 7 (θ-Lossy Average Wasserstein Distance)**.** Let P and Q be two probability distributions over a metric space $(\Omega, \mathfrak{d})$, and let $\theta \in [0,1]$. The *θ-lossy average Wasserstein distance* between P and Q is defined as:

$$W_\theta(P,Q) = \inf_{\phi\in\Phi^\theta(P,Q)} \mathop{\mathbb{E}}_{(x,y)\leftarrow\phi}[\mathfrak{d}(x,y)]. \tag{15}$$

The following lemma relates lossy average Wasserstein and lossy ∞-Wasserstein distances.

Lemma 5. *For any two distributions P, Q, and $0 \le \beta' < \beta \le 1$,*

$$W_\beta(P,Q) \le W_\beta^\infty(P,Q) \le \frac{W_{\beta'}(P,Q)}{(\beta-\beta')}.$$

Proof. Clearly from the definitions, $W_\beta(P,Q) \le W_\beta^\infty(P,Q)$.

Suppose $W_{\beta'}(P,Q) = \gamma$ and $\phi \in \Phi^{\beta'}(P,Q)$ is an optimal coupling that realizes this. Then, in ϕ, the total mass that is transported more than a distance γ' is at most γ/γ' and the total mass that is lost is at most β'. By choosing to simply not transport this mass at all, one loses $\beta' + \gamma/\gamma'$ mass, but no mass is transported more than a distance γ'. Choosing $\gamma' = \gamma/(\beta-\beta')$ this upper bound on loss is β, and hence this modified coupling shows that $W_\beta^\infty(P,Q) \le \gamma'$. □

A.3 γ-Lossy ∞-Wasserstein Distance Generalizes Existing Notions

Lemma 6. *Let $(\Omega, \mathfrak{d})$ be a metric space. Let F_f be a point distribution on some $f \in \Omega$ and G be a distribution over Ω. Then for any $\gamma \in [0,1]$ and $\beta \ge 0$, we have*

$$W_\gamma^\infty(F_f, G) \le \beta \iff \Pr_{g\leftarrow G}[\mathfrak{d}(f,g) > \beta] \le \gamma.$$

Proof. We show both the directions below.

- **Only if part (⇒):** Suppose $W_\gamma^\infty(F_f,G) \le \beta$. It follows from Lemma 4 that there exists a distribution G' such that $\Delta(G',G) \le \gamma$ and $W^\infty(F_f,G') \le \beta$. Since F_f is a point distribution, all couplings $\phi \in \Phi^0(F_f,G')$ will be such that $\phi_1 = F_f$ and $\phi_2 = G'$, which implies that $W^\infty(F_f,G') = \sup_{g'\leftarrow G'} \mathfrak{d}(f,g') \le \beta$. Now we show that, together with $\Delta(G',G) \le \gamma$, this implies $\Pr_{g\leftarrow G}[\mathfrak{d}(f,g) > \beta] \le \gamma$:

$$\begin{aligned}\Pr_{g\leftarrow G}[\mathfrak{d}(f,g)>\beta] &= \underbrace{\Pr_{g\leftarrow G}[\mathfrak{d}(f,g)>\beta \mid g\in\text{support}(G')]}_{=0} \Pr_{g\leftarrow G}[g\in\text{support}(G')] \\ &\quad + \underbrace{\Pr_{g\leftarrow G}[\mathfrak{d}(f,g)>\beta \mid g\notin\text{support}(G')]}_{\le 1} \Pr_{g\leftarrow G}[g\notin\text{support}(G')] \\ &\le \Pr_{g\leftarrow G}[g\notin\text{support}(G')]\end{aligned}$$

$$
\begin{aligned}
&= \int_{g\in\Omega:\ p_G(g)>0\ \&\ p_{G'}(g)=0} p_G(g)dg \\
&= \int_{g\in\Omega:\ p_G(g)>0\ \&\ p_{G'}(g)=0} (p_G(g)-p_{G'}(g))dg \\
&\overset{(a)}{\leq} \int_{g\in\Omega:\ p_G(g)>p_{G'}(g)} (p_G(g)-p_{G'}(g))dg \\
&\overset{(b)}{=} \Delta(G,G') \leq \gamma,
\end{aligned}
$$

where (a) follows because $\{g \in \Omega : p_G(g) > 0 \ \&\ p_{G'}(g) = 0\} \subseteq \{g \in \Omega : p_G(g) > p_{G'}(g)\}$, and (b) follows from the reasoning given below.
Define $\Omega_G^+ := \{g \in \Omega : p_G(g) > p_{G'}(g)\}$ and $\Omega_G^- := \{g \in \Omega : p_G(g) < p_{G'}(g)\}$. Since $\int_{g\in\Omega} p_G(g)dg = \int_{g\in\Omega} p_{G'}(g)dg$, it follows that $\int_{g\in\Omega_G^+}(p_G(g)-p_{G'}(g))dg = \int_{g\in\Omega_G^-}(p_{G'}(g)-p_G(g))dg$. Substituting this in the definition of $\Delta(G,G')$, we get $\Delta(G,G') = \int_{g\in\Omega_G^+}(p_G(g)-p_{G'}(g))dg$.

- **If part ($\Leftarrow$):** Suppose $\Pr_{g\leftarrow G}[\mathfrak{d}(f,g) > \beta] \leq \gamma$. Let $\Omega' = \{g \in \Omega : \mathfrak{d}(f,g) \leq \beta\}$ and G' be a distribution supported on Ω' such that $p_{G'}(g) = \frac{1}{\eta} p_G(g)$ when $g \in \Omega'$, otherwise $p_{G'}(g) = 0$. Here $\eta = \int_{g\in\Omega'} p_G(g)dg \geq (1-\gamma)$ is the normalizing constant. First we show that $\Delta(G,G') \leq \gamma$.

$$
\begin{aligned}
\Delta(G,G') &= \frac{1}{2}\int_{g\in\Omega} |p_{G'}(g)-p_G(g)|dg \\
&= \frac{1}{2}\int_{g\in\Omega'} |p_{G'}(g)-p_G(g)|dg + \frac{1}{2}\int_{g\in\Omega\setminus\Omega'} p_G(g)dg && \text{(Since } p_{G'}(g)=0 \text{ when } g\in\Omega\setminus\Omega'\text{)} \\
&= \frac{1}{2}\int_{g\in\Omega'} p_G(g)(\frac{1}{\eta}-1)dg + \frac{1}{2}\int_{g\in\Omega\setminus\Omega'} p_G(g)dg \\
&= \frac{1}{2}(\frac{1}{\eta}-1)\eta + \frac{1}{2}(1-\eta) && \text{(Since } \textstyle\int_{g\in\Omega'} p_G(g)dg = \eta\text{)} \\
&= 1-\eta \leq \gamma. && (16)
\end{aligned}
$$

Now define a joint distribution ϕ, whose first marginal is the point distribution F_f and the second marginal is G', which implies that $\sup_{(x,y)\leftarrow\phi} \mathfrak{d}(x,y) = \sup_{g'\in\Omega'} \mathfrak{d}(f,g')$. It follows from the argument above that $\phi \in \Phi^\gamma(F_f, G)$, which implies that $W_\gamma^\infty(F_f,G) \leq \sup_{(x,y)\leftarrow\phi} \mathfrak{d}(x,y) = \sup_{g'\in\Omega'} \mathfrak{d}(f,g') \leq \beta$, where the last inequality is by definition of Ω'. Hence, we get $W_\gamma^\infty(F_f,G) \leq \beta$.

This completes the proof of Lemma 6. □

Lemma 7. *For any two distributions P, Q over a metric space $(\Omega, \mathfrak{d})$ and $\gamma \in [0,1]$, we have*

$$W_\gamma^\infty(P,Q) = 0 \quad \Longleftrightarrow \quad \Delta(P,Q) \leq \gamma.$$

Proof. We show both the directions below.

- **Only if part ($\Rightarrow$):** Suppose $W_\gamma^\infty(P,Q) = 0$. This implies that there exists a joint distribution $\phi \in \Phi^\gamma(P,Q)$ such that $\sup_{(x,y)\leftarrow\phi} \mathfrak{d}(x,y) = 0$. Since $\mathfrak{d}$ is a metric, this implies that for all $(x,y) \leftarrow \phi$, we have $x = y$. Hence, the first marginal ϕ_1 and the second marginal ϕ_2 of ϕ are equal, which implies that $\Delta(\phi_1,P) + \Delta(\phi_2,Q) \leq \gamma$. Then, by triangle inequality and that $\phi_1 = \phi_2$, we get $\Delta(P,Q) \leq \gamma$.
- **If part ($\Leftarrow$):** Suppose $\Delta(P,Q) \leq \gamma$. Define a joint distribution $\phi := P \times P$. Since $\phi_1 = \phi_2 = P$, we have $\phi \in \Phi^\gamma(P,Q)$. This, by definition, implies $W_\gamma^\infty(P,Q) \leq \sup_{(x,y)\leftarrow\phi} \mathfrak{d}(x,y)$. Since both the marginals of ϕ are the same, we have $\mathfrak{d}(x,y) = 0$ for every $(x,y) \leftarrow \phi$. This, by the non-negativity of $W_\gamma^\infty(P,Q)$, gives $W_\gamma^\infty(P,Q) = 0$.

□

B Details Omitted from Section 3 – Usefulness [Blum et al., 2013] vs. Flexible Accuracy

To express accuracy guarantees of their mechanisms, Blum et al. [2013] introduced a notion of (β, γ, ψ)-*usefulness* that parallels (α, β, γ)-accuracy, except that ψ measures perturbation of the function rather than input distortion. Note that this is a reasonable notion for the function classes they considered (half-space queries, range queries etc.), but it is not applicable to queries like maximum.

Flexible accuracy generalizes the notion of usefulness. Firstly, mechanisms which are $(\beta, \gamma, 0)$-useful are $(0, \beta, \gamma)$-accurate (in [Blum et al., 2013], such mechanisms were given for interval queries). But even general usefulness can be translated to flexible accuracy generically, by redefining the function to have an extra input parameter that specifies perturbation. Further, the specific (β, γ, ψ)-useful DP mechanism of Blum et al. [2013] for half-space counting queries – with data points on a unit sphere, and the perturbation of the function corresponded to rotating the half-space by ψ radians – is (ψ, β, γ)-accurate for the same functions, w.r.t. the distortion ∂_{move}. This is because, the rotation of the half-space can be modeled as moving all the points on the unit sphere by a distance of at most ψ.

C Details Omitted from Section 4

C.1 More Details about Distortion Sensitivity

Distortion sensitivity of deterministic bijective functions. When $f : A \to B$ is a deterministic and bijective map, then for every x, Y such that $\widehat{\partial}_2(f(x), \mathfrak{p}_Y) \le \alpha$, there is only one choice of X for which $f(X) = Y$ holds, that is $X = f^{-1}(Y)$. Since for any point $x \in A$ and distribution P over A, we have $\widehat{\partial}_1(x, P) = \sup_{x' \in \text{support}(P)} \partial_1(x, x')$, it follows that

$$\sigma_f(\alpha) \quad = \quad \sup_{\substack{x, Y: \\ \widehat{\partial}_2(f(x), \mathfrak{p}_Y) \le \alpha}} \widehat{\partial}_1(x, \mathfrak{p}_{f^{-1}(Y)}) \quad = \quad \sup_{\substack{x \in A, y \in B: \\ \partial_2(f(x), y) \le \alpha}} \partial_1(x, f^{-1}(y)). \tag{17}$$

In particular, if $f : A \to A$ is an identity function and $\partial_1 = \partial_2$, then we have $\sigma_f(\alpha) \le \alpha$. Many of our mechanisms in this paper for which we derive flexible accuracy guarantees are given for the identity function over the space of histograms; see, for example, the result for our basic histogram mechanism (Theorem 3), the bucking mechanism (see Claim 6), and their composition (Theorem 4), etc.

A relaxed definition of distortion sensitivity. Though the distortion sensitivity is bounded in many circumstances (including all the applications we consider in this paper); however, due to the strict requirement of having an X such that $f(X) = Y$ (under infimum) in its definition, it may be infinite in other situations where this condition cannot be satisfied. To accommodate more functions, we can relax the definition with more parameters $\theta \in [0, 1]$, $\omega \ge 0$ as follows:

$$\sigma_f^{\gamma, \omega}(\alpha) = \sup_{\substack{x, Y: \\ \widehat{\partial}_2(f(x), \mathfrak{p}_Y) \le \alpha}} \inf_{\substack{X: \\ W_\gamma^\infty(f(X), \mathfrak{p}_Y) \le \omega}} \widehat{\partial}_1(x, \mathfrak{p}_X). \tag{18}$$

All the results in this paper can be extended to work with this more general definition of distortion sensitivity.

C.2 More Details about Error Sensitivity

Motivation. Suppose we want to compose an $(\alpha_1, \beta_1, \gamma_1)$-accurate mechanism M_1 for $f_1 : A \to B$ with another flexibly accurate mechanism M_2 for $f_2 : B \to C$ to obtain flexible accuracy guarantee of the composed mechanism $M_2 \circ M_1$ for $f_2 \circ f_1 : A \to C$. For this, on any input $x \in A$, first we measure the output error of M_1 on input x in terms of $W_{\gamma_1}^\infty(M_1(x), f_1(X'))$, where X' is an α_1-distortion of the *same* x on which we run the mechanism M_1; see (4). Now, for composition, we need to run M_2 on $M_1(x)$ and distort $f_1(X')$ to obtain another r.v. Y, and the output error of the composed mechanism is given by $W_\gamma^\infty(M_2(M_1(x)), f_2(Y))$. The problem here is that since the input (distribution) $f(X')$ that we distort is *not the same* as the input (distribution) $M_1(x)$ that we run M_2 on, we cannot directly obtain the output error guarantee of the composed mechanism from that of M_2.

Therefore, we need a way to generalize the measure of accuracy (output error) of a flexibly accurate mechanism when the input (distribution) to the mechanism is not the same as the input (distribution) that we distort, but they are at a bounded distance from each other (as measured in terms on the lossy ∞-Wasserstein distance). The notion of error sensitivity formalizes this intuition. Informally, it captures the sensitivity of the output error of a flexibly accurate mechanism in such situations.

C.3 More Details about Composition Theorem for Flexible Accuracy

Theorem 1 requires computing/bounding the error sensitivity of $\mathcal{M}_2$ in order to compute the flexible accuracy parameter β of $\mathcal{M}_2 \circ \mathcal{M}_1$. Now we show that the expression of error sensitivity can be simplified in some important special cases.

• **When $\mathcal{M}_1, f_1$ are deterministic maps and $\mathcal{M}_1$ is $(0, \beta_1, 0)$-accurate.** This setting arises when we compute the flexible accuracy parameters of our bucketed histogram mechanism $\mathcal{M}_{\text{BucHist}} = \mathcal{M}_{\text{STLap}} \circ \mathcal{M}_{\text{buc}}$ (Algorithm 3) while proving Theorem 4.

In this case, for any $x \in \mathcal{A}$, both $\mathcal{M}_1(x), f_1(x)$ are point distributions. This means that in order to compute the error sensitivity of $\mathcal{M}_2$, we only need to take the supremum in (7) over point distributions $\mathfrak{p}_x, \mathfrak{p}_{x'}$ over $\mathcal{B}$ (where $\mathfrak{p}_x, \mathfrak{p}_{x'}$ can be thought of being supported on $x := \mathcal{M}_1(x)$ and $x' := f_1(x)$, respectively) such that $W^\infty(\mathfrak{p}_x, \mathfrak{p}_{x'}) \le \beta_1$. Since $W^\infty(\mathfrak{p}_x, \mathfrak{p}_{x'}) = \mathfrak{d}_{\mathcal{B}}(x, x')$, we only need to take the supremum in (7) over $x, x' \in \mathcal{B}$ such that $\mathfrak{d}_{\mathcal{B}}(x, x') \le \beta_1$.

• **When $\mathcal{M}_2, f_2$ are deterministic maps and $\mathcal{M}_1$ is $(\alpha_1, \beta_1, 0)$-accurate and $\alpha_2 = \gamma_2 = 0$.** This setting arises in the case of histogram-based-statistics (denoted by a deterministic function f_{HBS}) in Section 5.2, in which we use the composed mechanism $f_{\text{HBS}} \circ \mathcal{M}_{\text{BucHist}}$ for computing f_{HBS}, where $\mathcal{M}_{\text{BucHist}}$ is our final histogram mechanism that is $(\alpha, \beta, 0)$-accurate (see Theorem 4) and f_{HBS} (as a mechanism) is $(0, 0, 0)$-accurate for computing f_{HBS}.

Upon substituting these parameters in (7), the expression for the error sensitivity reduces to computing $\tau^{0,0}_{\mathcal{M}_2, f_2}(\beta_1, 0) = \sup_{X, X' : W^\infty(\mathfrak{p}_X, \mathfrak{p}_{X'}) \le \beta_1} W^\infty(\mathcal{M}_2(X), f_2(X'))$, which can be simplified further as shown in the lemma below, which we prove in Appendix C.4.

C.4 Proof of Lemma 2

For convenience, we write the lemma statement below.

Lemma (Restating Lemma 2)**.** *Let $\mathcal{M} : \mathcal{B} \to \mathcal{C}$ be a deterministic mechanism for a deterministic function $f : \mathcal{B} \to \mathcal{C}$. Then, for any $\beta_1 \ge 0$, we have*

$$\tau^{0,0}_{\mathcal{M},f}(\beta_1, 0) \quad = \sup_{\substack{X, X' : \\ W^\infty(\mathfrak{p}_X, \mathfrak{p}_{X'}) \le \beta_1}} W^\infty(\mathcal{M}(X), f(X')) \quad = \sup_{\substack{x, x' \in \mathcal{A} : \\ \mathfrak{d}_{\mathcal{B}}(x, x') \le \beta_1}} \mathfrak{d}_{\mathcal{C}}(\mathcal{M}(x), f(x')).$$

Proof. The first equality follows from the definition of error sensitivity. We only need to prove the second equality.

- *LHS $\ge$ RHS:* This is the easy part.

$$\sup_{\substack{X, X' : \\ W^\infty(\mathfrak{p}_X, \mathfrak{p}_{X'}) \le \beta_1}} W^\infty(\mathcal{M}(X), f(X')) \quad \ge \sup_{\substack{x, x' \in \mathcal{B} : \\ W^\infty(\mathfrak{p}_x, \mathfrak{p}_{x'}) \le \beta_1}} W^\infty(\mathcal{M}(x), f(x')) \quad = \sup_{\substack{x, x' \in \mathcal{B} : \\ \mathfrak{d}_{\mathcal{B}}(\mathfrak{p}_x, \mathfrak{p}_{x'}) \le \beta_1}} \mathfrak{d}_{\mathcal{C}}(\mathcal{M}(x), f(x')),$$

  where the inequality holds because considering only point distributions restricts the set over which we take supremum and the equality holds because the ∞-Wasserstein distance between any two point distributions in any metric is just the distance between the points on which the distributions are supported in that metric.
- *LHS $\le$ RHS:* Consider any two distributions $\mathfrak{p}_X, \mathfrak{p}_{X'}$ over $\mathcal{B}$ s.t. $W^\infty(\mathfrak{p}_X, \mathfrak{p}_{X'}) \le \beta$. Let ϕ_1 be the optimal coupling between $\mathfrak{p}_X, \mathfrak{p}_{X'}$ such that

$$W^\infty(\mathfrak{p}_X, \mathfrak{p}_{X'}) \quad = \sup_{(\mathbf{x}, \mathbf{x}') \leftarrow \phi_1} \mathfrak{d}_{\mathcal{B}}(\mathbf{x}, \mathbf{x}') \le \beta_1.$$

Using $\phi_1, \mathcal{M}, f$, we define a joint distribution ϕ_2 over $\mathcal{C} \times \mathcal{C}$ as follows: For any $\mathbf{a}, \mathbf{b} \in \mathcal{C}$, define

$$\phi_2(\mathbf{a}, \mathbf{b}) \quad := \sum_{\substack{\mathbf{x},\mathbf{x}': \\ \mathcal{M}(\mathbf{x})=\mathbf{a}, f(\mathbf{x}')=\mathbf{b}}} \phi_1(\mathbf{x}, \mathbf{x}').$$

It can be verified that $\phi_2 \in \Phi(\mathcal{M}(X), f(X'))$, i.e., ϕ_2 is a valid coupling between $\mathcal{M}(X), f(X')$. Now

$$W^{\infty}(\mathcal{M}(X), f(X')) \leq \sup_{(\mathbf{a},\mathbf{b}) \leftarrow \phi_2} \mathfrak{d}_{\mathcal{C}}(\mathbf{a}, \mathbf{b}) = \sup_{(\mathbf{x},\mathbf{x}') \leftarrow \phi_1} \mathfrak{d}_{\mathcal{C}}(\mathcal{M}(\mathbf{x}), \mathbf{f}(\mathbf{x}')) \leq \sup_{\substack{\mathbf{x},\mathbf{x}' \in \mathcal{B}: \\ \mathfrak{d}_{\mathcal{B}}(\mathbf{x},\mathbf{x}') \leq \beta_1}} \mathfrak{d}_{\mathcal{C}}(\mathcal{M}(\mathbf{x}), \mathbf{f}(\mathbf{x}')),$$

where the last inequality holds because $\{(\mathbf{x}, \mathbf{x}') : (\mathbf{x}, \mathbf{x}') \leftarrow \phi_1\} \subseteq \{(\mathbf{x}, \mathbf{x}') : \mathfrak{d}_{\mathcal{B}}(\mathbf{x}, \mathbf{x}') \leq \beta_1\}$.
Note that the RHS of the last inequality does not depend on X, X'. So, taking supremum over all distributions X, X' such that $W^{\infty}(\mathfrak{p}_X, \mathfrak{p}_{X'}) \leq \beta_1$ gives the required result.

This completes the proof of Lemma 2. □

D Proof of Theorem 1 – Composition Theorem for Flexible Accuracy

The following lemma will be useful in proving Theorem 1. It translates the definition of distortion sensitivity (Definition 4) to apply to distortion of input distributions.

Lemma 8. *Suppose $f : A \to B$ has distortion sensitivity σ_f w.r.t. (∂_1, ∂_2). For all r.v.s X_0 over A and Y over B such that $\widehat{\partial}_2(f(X_0), \mathfrak{p}_Y) \leq \alpha$ for some $\alpha \geq 0$, there must exist a r.v. X over A such that $Y = f(X)$ and $\widehat{\partial}_1(\mathfrak{p}_{X_0}, \mathfrak{p}_X) \leq \sigma_f(\alpha)$, provided $\sigma_f(\alpha)$ is finite.*

Proof. Fix random variables X_0 over A and Y over B such that $\widehat{\partial}_2(f(X_0), \mathfrak{p}_Y) \leq \alpha$. Let ϕ be an optimal coupling that achieves the infimum in the definition of $\widehat{\partial}_2(f(X_0), \mathfrak{p}_Y)$, i.e.,

$$\widehat{\partial}_2(f(X_0), \mathfrak{p}_Y) = \sup_{(u,y) \leftarrow \phi} \partial_2(u, y) \leq \alpha. \tag{19}$$

For each $x_0 \in \text{support}(X_0)$, consider the conditional distribution $\phi_{x_0} = \phi | \{X_0 = x_0\}$. Clearly, the first marginal of ϕ_{x_0} is a point distribution supported at $f(x_0)$. Let its second marginal be denoted by $\mathfrak{p}_{Y_{x_0}}$. First we show that for each $x_0 \in \text{support}(X_0)$, we have $\widehat{\partial}_2(f(x_0), \mathfrak{p}_{Y_{x_0}}) \leq \alpha$.

$$\widehat{\partial}_2(f(x_0), \mathfrak{p}_{Y_{x_0}}) = \inf_{\phi \in \Phi^0(f(x_0), \mathfrak{p}_{Y_{x_0}})} \sup_{(u,y) \leftarrow \phi} \partial_2(u, y) \leq \sup_{(u,y) \leftarrow \phi_{x_0}} \partial_2(u, y) \overset{(a)}{\leq} \sup_{(u,y) \leftarrow \phi} \partial_2(u, y) \overset{(b)}{\leq} \alpha.$$

Here (a) follows from the fact that $\text{support}(\phi_{x_0}) \subseteq \text{support}(\phi)$ and (b) follows from (19). Thus for each $x_0 \in \text{support}(X_0)$, we have $\widehat{\partial}_2(f(x_0), \mathfrak{p}_{Y_{x_0}}) \leq \alpha$. Since $\sigma_f(\alpha)$ is finite, by the definition of σ_f, there exist a r.v. X_{x_0} such that

$$Y_{x_0} = f(X_{x_0}), \tag{20}$$

$$\widehat{\partial}_1(x_0, \mathfrak{p}_{X_{x_0}}) \leq \sigma_f(\alpha). \tag{21}$$

Define $X = \sum_{x_0 \in \text{support}(X_0)} \mathfrak{p}_{X_0}(x_0) X_{x_0}$. Now we show that $Y = f(X)$ and $\widehat{\partial}_1(\mathfrak{p}_{X_0}, \mathfrak{p}_X) \leq \sigma_f(\alpha)$.

- **Showing $Y = f(X)$:** Note that $Y = \sum_{x_0 \in \text{support}(X_0)} \mathfrak{p}_{X_0}(x_0) Y_{x_0}$ and $f(X) = \sum_{x_0 \in \text{support}(X_0)} \mathfrak{p}_{X_0}(x_0) f(X_{x_0})$. Now the claim follows because because $Y_{x_0} = f(X_{x_0})$ for each $x_0 \in \text{support}(X_0)$ (from (20)).
- **Showing $\widehat{\partial}_1(\mathfrak{p}_{X_0}, \mathfrak{p}_X) \leq \sigma_f(\alpha)$:** For each $x_0 \in \text{support}(X_0)$, let ψ_{x_0} be the optimal coupling that achieves the infimum in the definition of $\widehat{\partial}_1(x_0, \mathfrak{p}_{X_{x_0}})$. That is, for each x_0, $\psi_{x_0} \in \Phi^0(x_0, \mathfrak{p}_{X_{x_0}})$ and $\widehat{\partial}_1(x_0, \mathfrak{p}_{X_{x_0}}) = \sup_{(a,b) \leftarrow \psi_{x_0}} \partial_1(a, b)$. Let ψ be defined by $\psi(a, b) = \mathfrak{p}_{X_0}(x_0) \psi_{x_0}(a, b)$. It is easy to verify that $\psi \in \Phi^0(\mathfrak{p}_{X_0}, \mathfrak{p}_X)$. Further,

$$\widehat{\partial}_1(\mathfrak{p}_{X_0}, \mathfrak{p}_X) \leq \sup_{(a,b) \leftarrow \psi} \partial_1(a, b) = \sup_{x_0 \leftarrow \mathfrak{p}_{X_0}} \sup_{(a,b) \leftarrow \psi_{x_0}} \partial_1(a, b) = \sup_{x_0 \leftarrow \mathfrak{p}_{X_0}} \widehat{\partial}_1(x_0, \mathfrak{p}_{X_{x_0}}) \leq \sigma_f(\alpha),$$

where the last inequality follows from (21).

This completes the proof of Lemma 8. □

Now we prove Theorem 1, which is essentially formalizing the pictorial proof given in Figure 2. For convenience, we rewrite the statement of Theorem 1 below.

Theorem (Restating Theorem 1)**.** *Let $\mathcal{M}_1 : A \to B$ and $\mathcal{M}_2 : B \to C$ be mechanisms, respectively, with $(\alpha_1, \beta_1, \gamma_1)$-accuracy for $f_1 : A \to B$ and $\tau_{\mathcal{M}_2, f_2}$ error sensitivity for $f_2 : B \to C$, w.r.t. measures of distortion ∂_1, ∂_2 defined on A, B and metrics $\mathfrak{d}_1, \mathfrak{d}_2$ defined on B, C, respectively. Suppose f_1, α_2 are such that $\sigma_{f_1}(\alpha_2)$ is finite. Then, for any $\alpha_2 \geq 0$ and $\gamma_2 \in [0, 1]$, the mechanism $\mathcal{M}_2 \circ \mathcal{M}_1 : A \to C$ is (α, β, γ)-accurate for the function $f_2 \circ f_1$ w.r.t. ∂_1 and $\mathfrak{d}_2$, where $\alpha = \alpha_1 + \sigma_{f_1}(\alpha_2)$, $\beta = \tau^{\alpha_2, \gamma_2}_{\mathcal{M}_2, f_2}(\beta_1, \gamma_1)$, and $\gamma = \gamma_2$.*

Proof. For a given element $x \in A$, since $\mathcal{M}_1$ is $(\alpha_1, \beta_1, \gamma_1)$-accurate mechanism for f_1, we have from Definition 3 that there exists a r.v. X' such that

$$\widehat{\partial}_1(x, \mathfrak{p}_{X'}) \leq \alpha_1, \tag{22}$$
$$W^{\infty}_{\gamma_1}(f_1(X'), \mathcal{M}_1(x)) \leq \beta_1. \tag{23}$$

Now, applying the mechanism $\mathcal{M}_2$ on $\mathcal{M}_1(x)$, we incur an overall error of at most $\tau^{\alpha_2, \gamma_2}_{\mathcal{M}_2, f_2}(\beta_1, \gamma_1)$ to the output of function f_2 over a distorted input (see Definition 5). Therefore, there exists a r.v. Y^* such that,

$$\widehat{\partial}_2(f_1(X'), \mathfrak{p}_{Y^*}) \leq \alpha_2, \tag{24}$$
$$W^{\infty}_{\gamma_2}(f_2(Y^*), \mathcal{M}_2(\mathcal{M}_1(x))) \leq \tau^{\alpha_2, \gamma_2}_{\mathcal{M}_2, f_2}(\beta_1, \gamma_1). \tag{25}$$

Since $\sigma_{f_1}(\alpha_2)$ is finite (by assumption), it follows from (24) and Lemma 8 that there exists a r.v. X over A such that

$$\widehat{\partial}_1(\mathfrak{p}_{X'}, \mathfrak{p}_X) \leq \sigma_{f_1}(\alpha_2), \tag{26}$$
$$Y^* = f_1(X). \tag{27}$$

Since ∂_1 is a quasi-metric, it follows that $\widehat{\partial}_1$ is also a quasi-metric; see Lemma 13 in Appendix K for a proof. This, together with (22) and (26), implies that

$$\widehat{\partial}_1(x, \mathfrak{p}_X) \leq \alpha_1 + \sigma_{f_1}(\alpha_2). \tag{28}$$

Substituting $Y^* = f_1(X)$ from (27) into (25) gives

$$W^{\infty}_{\gamma_2}(f_2(f_1(X)), \mathcal{M}_2(\mathcal{M}_1(x))) \leq \tau^{\alpha_2, \gamma_2}_{\mathcal{M}_2, f_2}(\beta_1, \gamma_1). \tag{29}$$

(28) and (29) imply that $\mathcal{M}_2 \circ \mathcal{M}_1$ is (α, β, γ)-accurate for $f_2 \circ f_1$ w.r.t. the distortion measure ∂_1 on A and metric $\mathfrak{d}_2$ on C, where $\alpha = \alpha_1 + \sigma_{f_1}(\alpha_2)$, $\beta = \tau^{\alpha_2, \gamma_2}_{\mathcal{M}_2, f_2}(\beta_1, \gamma_1)$, and $\gamma = \gamma_2$.

This concludes the proof of Theorem 1. □

E Proof of Theorem 2 – Differential Privacy Under Composition

Differential Privacy. Let $\mathcal{X}$ denote a universe of possible "databases" with a symmetric neighborhood relation $\sim$. In typical applications, two databases $\mathbf{x}$ and $\mathbf{x}'$ are considered neighbors if one is obtained from the other by removing the data corresponding to a single "individual." A *mechanism* $\mathcal{M}$ over $\mathcal{X}$ is an algorithm that takes $\mathbf{x} \in \mathcal{X}$ as input and samples from an output space $\mathcal{Y}$, according to some distribution. We shall denote this distribution by $\mathcal{M}(\mathbf{x})$.

Definition 8 (Differential Privacy [Dwork et al., 2006b,a])**.** A randomized algorithm $\mathcal{M} : \mathcal{X} \to \mathcal{Y}$ is (ϵ, δ)-differentially private (DP), if for all neighboring databases $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ and all measurable subsets $S \subseteq \mathcal{Y}$, we have $\Pr[\mathcal{M}(\mathbf{x}) \in S] \leq e^{\epsilon} \Pr[\mathcal{M}(\mathbf{x}') \in S] + \delta$.

Definition 9 (Neighborhood preserving Mechanism)**.** A mechanism $\mathcal{M} : A \to B$ is *neighborhood preserving* w.r.t. neighborhood relations $\sim_A$ over A and $\sim_B$ over B, if for all $x, y \in A$ s.t. $x \sim_A y$, there exists a pair of jointly distributed random variables (X, Y) s.t. $\mathfrak{p}_X = \mathcal{M}(x)$, $\mathfrak{p}_Y = \mathcal{M}(y)$, and $\Pr[X \sim_B Y] = 1$.

Theorem (Restating Theorem 2)**.** *Let $\mathcal{M}_1 : A \to B$ and $\mathcal{M}_2 : B \to C$ be any two mechanisms. If $\mathcal{M}_1$ is neighborhood-preserving w.r.t. neighborhood relations $\sim_A$ and $\sim_B$ over A and B, respectively, and $\mathcal{M}_2$ is (ϵ, δ)-DP w.r.t. $\sim_B$, then $\mathcal{M}_2 \circ \mathcal{M}_1 : A \to C$ is (ϵ, δ)-DP w.r.t. $\sim_A$.*

Proof. For simplicity, we consider the case when B is discrete. The proof can be generalized to the continuous setting.

Since the mechanism $\mathcal{M}_1$ is neighborhood preserving, for $x, x' \in A$ s.t. $x_1 \sim_A x_2$, there exists a pair of jointly distributed random variables (X_1, X_2) over $B \times B$ s.t, $\mathfrak{p}_{X_1} = \mathcal{M}_1(x)$, $\mathfrak{p}_{X_2} = \mathcal{M}_1(x')$ and $\Pr[X_1 \sim_B X_2] = 1$. So, for all (x_1, x_2) such that $\mathfrak{p}_{X_1,X_2}(x_1, x_2) > 0$, we have $x_1 \sim_B x_2$ and hence, by the (ϵ, δ)-differential privacy of the mechanism $\mathcal{M}_2$, for all subsets $S \subseteq C$, we have,

$$\Pr(\mathcal{M}_2(x_1) \in S) \leq e^\epsilon \Pr(\mathcal{M}_2(x_2) \in S) + \delta.$$

Thus, if $x \sim_A x'$, then for any subset $S \subseteq C$, we have,

$$\begin{aligned}
\Pr[\mathcal{M}_2(\mathcal{M}_1(x)) \in S] &= \sum_{x_1} \mathfrak{p}_{X_1}(x_1) \Pr[\mathcal{M}_2(x_1) \in S] \\
&= \sum_{(x_1,x_2)} \mathfrak{p}_{X_1,X_2}(x_1, x_2) \Pr[\mathcal{M}_2(x_1) \in S] \\
&\leq \sum_{(x_1,x_2)} \mathfrak{p}_{X_1,X_2}(x_1, x_2) \left(e^\epsilon \Pr[\mathcal{M}_2(x_2) \in S] + \delta\right) \\
&= e^\epsilon \left(\sum_{(x_1,x_2)} \mathfrak{p}_{X_1,X_2}(x_1, x_2) \Pr[\mathcal{M}_2(x_2) \in S] \right) + \delta \\
&= e^\epsilon \left(\sum_{x_2} \mathfrak{p}_{X_2}(x_2) \Pr[\mathcal{M}_2(x_2) \in S] \right) + \delta \\
&= e^\epsilon \Pr[\mathcal{M}_2(\mathcal{M}_1(x')) \in S] + \delta
\end{aligned}$$

This completes the proof of Theorem 2. □

F Proof of Theorem 3 – Truncated Laplace Mechanism for Histograms

First, we prove the flexible accuracy part, which is easy, and then we will move on to proving the privacy part, which is more involved than the existing privacy analysis of differentially-private histogram mechanisms. We also note that the requirement of $|\text{support}(\mathbf{x})| \leq t$ is only needed the accuracy result.

Flexible accuracy. Note that the noise added by $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ in each bar of the histogram is bounded by $-q = -\tau|\mathbf{x}|$, which can lead to a drop of at most τ fraction of total number of elements from each bar. Combined with the fact that $|\text{support}(\mathbf{x})| \leq t$, the fraction of the maximum fraction of elements that can be dropped is τt. Hence, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $(\tau t, 0, 0)$-accurate.

Differential privacy. Our proof of the privacy part of Theorem 3 depends on the following lemma.

Lemma 9. *For any $\nu \geq 0, \epsilon > 0$ and on inputs $\mathbf{x}$ s.t. $|\mathbf{x}| \geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\frac{\epsilon\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})}-1}\right)$, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $\left((1+\nu)\epsilon, \frac{e^\epsilon - 1}{2(e^{\epsilon q/2}-1)}\right)$-DP w.r.t. $\sim_{\text{hist}}$, where $q = \tau|\mathbf{x}|$.*

Proof. We shall, in fact, prove that a mechanism that outputs $\hat{\mathbf{y}}$ with $\hat{\mathbf{y}}(i) := \mathbf{x}(i) + z_i$ (without rounding, and without replacing negative values with 0) is already differentially private as desired. Then, since the actual mechanism is a post-processing of this mechanism, it will also be differentially private with the same parameters.

Let $\mathbf{x}$ and $\mathbf{x}'$ be two neighbouring histograms. For simplicity, for every $i \in \mathcal{G}$, define $x_i := \mathbf{x}(i)$ and $x'_i := \mathbf{x}'(i)$. Since $\mathbf{x} \sim \mathbf{x}'$, there exists an $i^* \in \mathcal{G}$ such that $|x_{i^*} - x'_{i^*}| = 1$ and that $x_i = x'_i$ for every $i \in \mathcal{G} \setminus \{i^*\}$. Without

loss of generality, assume that $x_{i^*} = x'_{i^*} + 1$, which implies $|\mathbf{x}| = |\mathbf{x}'| + 1 = n + 1$. Let $q = \tau(n+1)$ and $q' = \tau n$. For simplicity of notation, we will denote support($\mathbf{y}$) by $\mathcal{G}_{\mathbf{y}}$ for any $\mathbf{y} \in \{\mathbf{x}, \mathbf{x}'\}$.

In order to prove the lemma, for every subset $S \subseteq \mathcal{H}_{\mathcal{G}}$, we need to show that

$$\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S] \leq e^{(1+\nu)\epsilon} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in S] + \delta, \tag{30}$$

$$\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in S] \leq e^{(1+\nu)\epsilon} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S] + \delta, \tag{31}$$

where $\delta = \frac{e^{\epsilon}-1}{2(e^{\epsilon q/2}-1)}$. We only prove (30); (31) can be shown similarly.

Fix an arbitrary subset $S \subseteq \mathcal{H}_{\mathcal{G}}$. Since $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ adds independent noise to each bar of the histogram according to $\pi_q(z)$, we have that for every $\mathbf{s} \in \mathcal{H}_{\mathcal{G}}$, we have $\mathfrak{p}_{\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})}(\mathbf{s}) = \prod_{i \in \mathcal{G}_{\mathbf{x}}} \pi_q(s_i - x_i)$ where $s_i = \mathbf{s}(i)$. Thus, we have

$$\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in S] = \int_S \Big[\prod_{i \in \mathcal{G}_{\mathbf{x}}} \pi_q(s_i - x_i) \Big] \, \mathrm{d}\mathbf{s}, \tag{32}$$

$$\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S] = \int_S \Big[\prod_{i \in \mathcal{G}_{\mathbf{x}'}} \pi_{q'}(s_i - x'_i) \Big] \, \mathrm{d}\mathbf{s}. \tag{33}$$

Now, using the fact that $\forall k \neq i^*, x_k = x'_k$ and $x_{i^*} = x'_{i^*} + 1$, we partition S into three disjoint sets:

1. $S_0 := \{\mathbf{s} \in \mathcal{H}_{\mathcal{G}} : s_{i^*} - x'_{i^*} < -q'\} \cup \{\mathbf{s} \in \mathcal{H}_{\mathcal{G}} : 0 < s_{i^*} - x'_{i^*}\}$.
2. $S_1 := \{\mathbf{s} \in \mathcal{H}_{\mathcal{G}} : -q' \leq s_{i^*} - x'_{i^*} < -q' + (1-\tau)\}$.
3. $S_2 := \{\mathbf{s} \in \mathcal{H}_{\mathcal{G}} : -q' + (1-\tau) \leq s_{i^*} - x'_{i^*} \leq 0)\}$.

The proof of (30) is a simple corollary of the following two claims, which we prove in Appendix F.1.

Claim 3. $\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S_0 \cup S_2] \leq e^{(1+\nu)\epsilon} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in S_0 \cup S_2]$, *provided* $n \geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\frac{\epsilon\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})}-1}\right)$.

Claim 4. $\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S_1] \leq \delta$, *where* $\delta = \frac{e^{\epsilon}-1}{2(e^{\epsilon q/2}-1)}$.

The above two claims together imply (30) as follows:

$$\begin{aligned}
\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S] &= \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S_0 \cup S_2] + \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S_1] \\
&\leq e^{(1+\nu)\epsilon} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in S_0 \cup S_2] + \delta \\
&\leq e^{(1+\nu)\epsilon} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in S] + \delta. && \text{(Since } S_0 \cup S_2 \subseteq S\text{)}
\end{aligned}$$

This completes the proof of Lemma 9. □

In Lemma 9, ν is a free variable. By taking $\nu = 0$, we get the following result in Corollary 1. We can also get different guarantees by restricting to $\nu > 0$; see Remark 3 below for this.

Corollary 1. *For any* $\epsilon, \tau, \mathbf{x}$ *such that* $\tau|\mathbf{x}|\epsilon \geq 2$, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ *is* $\left(\epsilon, \frac{e^{\epsilon}-1}{2(e^{\epsilon q/2}-1)}\right)$*-DP w.r.t.* $\sim_{\text{hist}}$, *where* $q = \tau|\mathbf{x}|$.

Proof. Substituting $\nu = 0$ in Lemma 9 gives that when $\mathbf{x}$ satisfies $|\mathbf{x}| \geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\frac{\epsilon\tau}{2}}}{e^{\frac{\epsilon\tau}{2}}-1}\right)$, we have that $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $\left(\epsilon, \frac{e^{\epsilon}-1}{2(e^{\epsilon n\tau/2}-1)}\right)$-DP w.r.t. $\sim_{\text{hist}}$. Now, the corollary follows because $\frac{2}{\epsilon\tau} \geq \frac{2}{\epsilon\tau} e^{-\frac{\epsilon\tau}{2}} \geq \frac{2}{\epsilon\tau} \ln\left(1 + e^{-\frac{\epsilon\tau}{2}}\right) = \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\frac{\epsilon\tau}{2}}}{e^{\frac{\epsilon\tau}{2}}-1}\right)$, where the first inequality uses $x \geq \ln(1+x)$ for $x > 0$. □

Remark 3. *We show in Lemma 10 in Appendix F.1 that by restricting Lemma 9 to* $\nu > 0$, *we can get a weaker condition than what we have in Corollary 1 with a slight increase in the privacy parameter* ϵ. *In particular, we show that for all* $\epsilon, \mathbf{x}$ *such that* $\epsilon\nu \geq \ln\left(1 + \frac{1}{|\mathbf{x}|}\right)$, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ *is* $\left((1+\nu)\epsilon, \frac{e^{\epsilon}-1}{2(e^{\epsilon q/2}-1)}\right)$*-DP w.r.t.* $\sim_{\text{hist}}$. *We can take* $\nu = 1$ *here.*

Now the privacy part of Theorem 3 follows because $q = \tau|\mathbf{x}|$ and $\tau|\mathbf{x}|\epsilon \geq 2$ (note that $\tau|\mathbf{x}|\epsilon$ is typically a much bigger number than 2 as it scales with the size of the dataset), which implies that $\frac{e^{\epsilon}-1}{2(e^{\epsilon q/2}-1)} = \epsilon e^{-\Omega(\epsilon\tau|\mathbf{x}|)}$. Hence, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $(\epsilon, \epsilon e^{-\Omega(\epsilon\tau|\mathbf{x}|)})$-DP.

This completes the proof of Theorem 3.

F.1 Details Omitted from the Proof of Theorem 3

Claim (Restating Claim 3)**.** $\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S_0 \cup S_2] \leq e^{(1+\nu)\epsilon} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in S_0 \cup S_2]$, *provided* $n \geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\frac{\epsilon\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})}-1}\right)$.

Proof. First we show that for $\mathbf{s} \in S_0 \cup S_2$, we have, $\pi_{q'}(s_{i^*} - x'_{i^*}) \leq e^{(1+\nu)\epsilon}\pi_q(s_{i^*} - x_{i^*})$, provided $n \geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\frac{\epsilon\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})}-1}\right)$, and then we show how this implies the result.

For $\mathbf{s} \in S_0$, $\pi_{q'}(s_{i^*} - x'_{i^*}) = 0$ so the inequality trivially holds. For $\mathbf{s} \in S_2$, both $\pi_{q'}(s_{i^*} - x'_{i^*}) > 0$ and $\pi_q(s_{i^*} - x_{i^*}) > 0$; hence, we will be done if we show that $\frac{\pi_{q'}(s_{i^*}-x'_{i^*})}{\pi_q(s_{i^*}-x_{i^*})} \leq e^{(1+\nu)\epsilon}$. Note that we are given the following inequality:

$$n \geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\epsilon\frac{\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})}-1}\right),$$

which can be rewritten as (which we show in Claim 5 after this proof):

$$\ln\left(\frac{1-e^{-\epsilon\frac{\tau(n+1)}{2}}}{1-e^{-\epsilon\frac{\tau n}{2}}}\right) \leq \epsilon(\nu + \frac{\tau}{2}). \tag{34}$$

By substituting $q = \tau(n+1)$ and $q' = \tau n$, (34) is equivalent to

$$\frac{1}{\epsilon}\ln\left(\frac{1-e^{-\epsilon\frac{q}{2}}}{1-e^{-\epsilon\frac{q'}{2}}}\right) + (1 - \frac{\tau}{2}) \leq 1 + \nu.$$

This, using the triangle inequality, implies that

$$\frac{1}{\epsilon}\ln\left(\frac{1-e^{-\epsilon\frac{q}{2}}}{1-e^{-\epsilon\frac{q'}{2}}}\right) + \left|s_{i^*} - x_{i^*} + \frac{q}{2}\right| - \left|s_{i^*} - x_{i^*} + \frac{q}{2} + (1 - \frac{\tau}{2})\right| \leq 1 + \nu.$$

Putting $q' = q - \tau$ and $x'_{i^*} = x_{i^*} - 1$, we get

$$\frac{1}{\epsilon}\ln\left(\frac{1-e^{-\epsilon\frac{q}{2}}}{1-e^{-\epsilon\frac{q'}{2}}}\right) + \left|s_{i^*} - x_{i^*} + \frac{q}{2}\right| - \left|s_{i^*} - x'_{i^*} + \frac{q'}{2}\right| \leq 1 + \nu.$$

By taking exponents of both sides, this is equivalent to showing

$$\frac{(1-e^{-\epsilon\frac{q}{2}})}{(1-e^{-\epsilon\frac{q'}{2}})} \frac{e^{-\epsilon|s_{i^*} - x'_{i^*} + \frac{q'}{2}|}}{e^{-\epsilon|s_{i^*} - x_{i^*} + \frac{q}{2}|}} \leq e^{(1+\nu)\epsilon}$$

By substituting the values of $\pi_q(s_{i^*} - x_{i^*})$ and $\pi_{q'}(s_{i^*} - x'_{i^*})$, this can be equivalently written as

$$\frac{\pi_{q'}(s_{i^*} - x'_{i^*})}{\pi_q(s_{i^*} - x_{i^*})} \leq e^{(1+\nu)\epsilon}. \tag{35}$$

Now we show $\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S_0 \cup S_2] \leq e^{(1+\nu)\epsilon} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in S_0 \cup S_2]$. Recall that $\mathcal{G}_{\mathbf{x}} = \text{support}(\mathbf{x})$ for any histogram $\mathbf{x} \in \mathcal{H}_{\mathcal{G}}$.

$$\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}') \in S_0 \cup S_2] = \int_{S_0 \cup S_2} \Big[\prod_{i \in \mathcal{G}_{\mathbf{x}'}} \pi_{q'}(s_i - x'_i)\Big] \, \mathbf{ds}$$

$$
\begin{aligned}
&= \int_{S_0 \cup S_2} \Big[\prod_{i \in \mathcal{G}_{\mathbf{x}'}: i \neq i^*} \pi_{q'}(s_i - x'_i) \Big] \pi_{q'}(s_{i^*} - x'_{i^*}) \, \mathrm{d}\mathbf{s} \\
&\leq \int_{S_0 \cup S_2} \Big[\prod_{i \in \mathcal{G}_{\mathbf{x}}: i \neq i^*} \pi_q(s_i - x_i) \Big] e^{(1+\nu)\epsilon} \pi_q(s_{i^*} - x_{i^*}) \, \mathrm{d}\mathbf{s} \\
&\qquad\qquad \text{(Using (35) and that } x_i = x'_i, \forall i \neq i^*) \\
&= e^{(1+\nu)\epsilon} \int_{S_0 \cup S_2} \Big[\prod_{i \in \mathcal{G}_{\mathbf{x}}} \pi_q(s_i - x_i) \Big] \, \mathrm{d}\mathbf{s} \\
&= e^{(1+\nu)\epsilon} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\mathrm{STLap}}(\mathbf{x}) \in S_0 \cup S_2]
\end{aligned}
$$

This completes the proof of Claim 3. □

Claim 5.

$$
n \geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1 - e^{-\epsilon\frac{\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})} - 1}\right) \quad \iff \quad \ln\left(\frac{1 - e^{-\epsilon\frac{\tau(n+1)}{2}}}{1 - e^{-\epsilon\frac{\tau n}{2}}}\right) \leq \epsilon(\nu + \frac{\tau}{2}).
$$

Proof. We will start with the RHS and show that it is equivalent to the LHS.

$$
\begin{aligned}
\frac{1 - e^{-\epsilon\frac{\tau(n+1)}{2}}}{1 - e^{-\epsilon\frac{\tau n}{2}}} &\leq e^{\epsilon(\nu+\frac{\tau}{2})} \\
\iff 1 - e^{-\epsilon\frac{\tau(n+1)}{2}} &\leq e^{\epsilon(\nu+\frac{\tau}{2})} - e^{\epsilon(\nu+\frac{\tau}{2})} e^{-\epsilon\frac{\tau n}{2}} \\
\iff 1 - e^{-\epsilon\frac{\tau n}{2}} e^{-\epsilon\frac{\tau}{2}} &\leq e^{\epsilon(\nu+\frac{\tau}{2})} - e^{\epsilon(\nu+\frac{\tau}{2})} e^{-\epsilon\frac{\tau n}{2}} \\
\iff e^{-\epsilon\frac{\tau n}{2}} \left(e^{\epsilon(\nu+\frac{\tau}{2})} - e^{-\epsilon\frac{\tau}{2}}\right) &\leq e^{\epsilon(\nu+\frac{\tau}{2})} - 1 \\
\iff e^{\epsilon\frac{\tau n}{2}} &\geq \frac{e^{\epsilon(\nu+\frac{\tau}{2})} - e^{-\epsilon\frac{\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})} - 1} \\
\iff e^{\epsilon\frac{\tau n}{2}} &\geq 1 + \frac{1 - e^{-\epsilon\frac{\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})} - 1} \\
\iff n &\geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1 - e^{-\epsilon\frac{\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})} - 1}\right).
\end{aligned}
$$

□

Claim (Restating Claim 4)**.** $\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\mathrm{STLap}}(\mathbf{x}') \in S_1] \leq \frac{e^\epsilon - 1}{2(e^{\epsilon q/2} - 1)}$.

Proof. Observe that, for every $\mathbf{s} \in S_1$, we have $-q' \leq s_{i^*} - x'_{i^*} < -q' + (1 - \tau)$. Recall that $\mathcal{G}_{\mathbf{x}'} = \mathrm{support}(\mathbf{x}')$ and $|\mathbf{x}'| = n$. Let $|\mathcal{G}_{\mathbf{x}'}| = t$ for some $t \leq n$, and, for simplicity, assume that $\mathcal{G}_{\mathbf{x}'} = \{1, 2, \ldots, t\}$. For $i \in [t]$, define $S_1(i) := \{\hat{s}_i : \exists \mathbf{s} \in S_1 \text{ s.t. } \hat{s}_i = s_i\}$, which is equal to the collection of the multiplicity of i in the histograms in S_1.

$$
\begin{aligned}
\Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\mathrm{STLap}}(\mathbf{x}') \in S_1] &= \int_{S_1} \Big[\prod_{i=1}^{t} \pi_{q'}(s_i - x'_i) \Big] \, \mathrm{d}\mathbf{s} \\
&= \int_{S_1(1)} \cdots \int_{S_1(i^*)} \cdots \int_{S_1(t)} \Big[\prod_{i=1}^{t} \pi_{q'}(s_i - x'_i) \Big] \, \mathrm{d}s_t \ldots \mathrm{d}s_{i^*} \ldots \mathrm{d}s_1 \\
&= \int_{S_1(i^*)} \pi_{q'}(s_{i^*} - x'_{i^*}) \underbrace{\left(\int_{S_1(1)} \cdots \int_{S_1(t)} \Big[\prod_{i=1: i \neq i^*}^{t} \pi_{q'}(s_i - x'_i) \Big] \, \mathrm{d}s_t \ldots \mathrm{d}s_1 \right)}_{\leq 1} \mathrm{d}s_{i^*} \\
&\leq \int_{S_1(i^*)} \pi_{q'}(s_{i^*} - x'_{i^*}) \, \mathrm{d}s_{i^*} \\
&= \int_{q'}^{q' + (1-\tau)} \pi_{q'}(z) \, \mathrm{d}z \qquad \text{(Since } \forall \mathbf{s} \in S_1, \, (s_{i^*} - x'_{i^*}) \in [-q', -q' + (1 - \tau))\text{)}
\end{aligned}
$$

$$= \frac{e^{(1-\tau)\epsilon}-1}{2(1-e^{-\epsilon q/2})} e^{-\epsilon q/2}$$
$$\leq \frac{e^{\epsilon}-1}{2(e^{\epsilon q/2}-1)}. \qquad \text{(Since } \tau > 0\text{)}$$

This proves Claim 4. □

Lemma 10. *For any $\nu, \epsilon > 0$ and $\mathbf{x}$ such that $\epsilon\nu > \ln\left(1+\frac{1}{|\mathbf{x}|}\right)$, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $\left((1+\nu)\epsilon, \frac{e^{\epsilon}-1}{2(e^{\epsilon q/2}-1)}\right)$-DP w.r.t. $\sim_{\text{hist}}$, where $q = \tau|\mathbf{x}|$.*

Proof. We use Lemma 9 and put a restriction that ν should be > 0. We will analyze the effect of this restriction on the bound of $|\mathbf{x}|$. We restate the bound on $|\mathbf{x}|$ here again for convenience:

$$|\mathbf{x}| \geq \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\epsilon\frac{\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})}-1}\right)$$

It can be easily checked that for any fixed $\epsilon, \nu > 0$, the RHS is a decreasing function of τ. Hence, if we set τ to its minimum value, we get a lower bound on $|\mathbf{x}|$ which is independent of τ. Since this expression is not defined at $\tau = 0$, we will take its one-sided limit as $\tau \to 0^+$, i.e.,

$$\lim_{\tau\to 0^+} \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\epsilon\frac{\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})}-1}\right)$$

We will replace $\frac{\epsilon\tau}{2}$ with l. As $\tau \to 0^+$, $l \to 0^+$, and we get

$$\begin{aligned}
\lim_{\tau\to 0^+} \frac{2}{\epsilon\tau} \ln\left(1 + \frac{1-e^{-\epsilon\frac{\tau}{2}}}{e^{\epsilon(\nu+\frac{\tau}{2})}-1}\right) &= \lim_{l\to 0^+} \frac{1}{l} \ln\left(1 + \frac{1-e^{-l}}{e^{\epsilon\nu+l}-1}\right) \\
&= \lim_{l\to 0^+} \frac{1}{l} \ln\left(1 + \frac{1-e^{-l}}{e^{\epsilon\nu+l}-1}\right)\left(\frac{1-e^{-l}}{e^{\epsilon\nu+l}-1}\right)\left(\frac{e^{\epsilon\nu+l}-1}{1-e^{-l}}\right) \\
&= \lim_{l\to 0^+} \left(\frac{1}{e^{\epsilon\nu+l}-1}\right)\left(\frac{1-e^{-l}}{l}\right)\left(\frac{\ln\left(1+\frac{1-e^{-l}}{e^{\epsilon\nu+l}-1}\right)}{\frac{1-e^{-l}}{e^{\epsilon\nu+l}-1}}\right) \\
&= \frac{1}{e^{\epsilon\nu}-1} \qquad \left(\lim_{x\to 0^+}\tfrac{1-e^{-x}}{x} = 1;\ \lim_{x\to 0^+}\tfrac{\ln(1+x)}{x} = 1\right)
\end{aligned}$$

We have proved that on inputs $\mathbf{x}$ s.t. $|\mathbf{x}| > \frac{1}{e^{\epsilon\nu}-1}$, which is equivalent to the condition that $\epsilon\nu > \ln\left(1+\frac{1}{|\mathbf{x}|}\right)$, $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ is $\left((1+\nu)\epsilon, \frac{e^{\epsilon}-1}{2(e^{\epsilon q/2}-1)}\right)$-DP w.r.t. $\sim_{\text{hist}}$, where $q = \tau|\mathbf{x}|$. □

G Proof of Theorem 4 – Bucketed, Shifted and Truncated Laplace Mechanism

Note that $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}} = \mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}} \circ \mathcal{M}^{w,[0,B)}_{\text{buc}}$, with $w = 2\beta$ and $\tau = \frac{\alpha}{t}$, where $t = \lceil \frac{B}{2\beta} \rceil$. We will use Theorem 2 to show the DP guarantee and Theorem 1 to show the flexible accuracy guarantee of $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$.

Differential privacy. First note that $\mathcal{M}^{w,[0,B)^d}_{\text{buc}}$ is a neighborhood-preserving mechanism w.r.t. the neighborhood relation $\sim_{\text{hist}}$. This follows because adding/removing any one element changes the output of bucketing by at most one element; hence, neighbors remain neighbors after bucketing. Now, since $\mathcal{M}^{w,[0,B)}_{\text{buc}}$ outputs a histogram whose support size is at most $t = \lceil \frac{B}{w} \rceil$, and $\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}}$ on input histograms with support size at most t is $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)}\right)$-differentially private w.r.t. $\sim_{\text{hist}}$, it follows from Theorem 2 that $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$ is also differentially private w.r.t. $\sim_{\text{hist}}$with the same parameters.

Flexible accuracy. We claim the following flexible accuracy guarantee of the bucketing mechanism $\mathcal{M}^{w,[0,B)}_{\text{buc}}(\mathbf{x})$ that is proved in Appendix G.1.

Claim 6. *$\mathcal{M}^{w,[0,B)}_{\text{buc}}$ is $\left(0, \frac{w}{2}, 0\right)$-accurate for the identity function f_{id} over $\mathcal{H}_{[0,B)}$ w.r.t the metric $\mathfrak{d}_{\text{hist}}$.*

Note that when we apply Theorem 3 to compute the flexible accuracy parameters of the composed mechanism $\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}} \circ \mathcal{M}^{w,[0,B)}_{\text{buc}}$, the parameters of the composed mechanism depend on the distortion sensitivity $\sigma_{f_1}(\alpha_2)$ and the error sensitivity of $\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}}$. We compute them below.

• *Distortion sensitivity of f_1:* Since f_1 is the identity function f_{id} over $\mathcal{H}_{[0,B)}$, we have (as noted in the first example in Section **??**) that $\sigma_{f_1}(\alpha_2) \leq \alpha_2$.

• *Error sensitivity of $\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}}$:* Note that the bucketing mechanism $\mathcal{M}^{w,[0,B)}_{\text{buc}} : \mathcal{H}_{[0,B)} \to \mathcal{H}_{[0,B)}$ is a deterministic map, and is $(0, \beta, 0)$-accurate (see Claim 6) for computing the identity function f_{id}, where $\beta = \frac{w}{2}$. As mentioned in the first bullet after the statement of Theorem 1, this implies that when computing the error sensitivity of $\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}}$ (which is required for calculating the output error β of the composed mechanism $\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}} \circ \mathcal{M}^{w,[0,B)}_{\text{buc}}$), we only need to take supremum in (7) over point distributions $\mathbf{x}, \mathbf{x}'$ such that $\mathfrak{d}_{\text{hist}}(\mathbf{x}, \mathbf{x}') \leq \beta$, where $\mathfrak{d}_{\text{hist}}(,)$ is the metric that we use over $\mathcal{H}_{[0,B)}$. In other words, in order to compute the error sensitivity of $\mathcal{M}^{\tau,[0,B)}_{\text{STLap}}$, we only need to bound $\sup_{\mathbf{x},\mathbf{x}':\mathfrak{d}_{\text{hist}}(\mathbf{x},\mathbf{x}')\leq\beta} \inf_{Y:\widehat{\partial}(\mathbf{x}',Y)\leq\alpha} W^{\infty}(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}), \mathfrak{p}_Y)$. We bound this in Lemma 11 below.

Lemma 11. *For any $\alpha, \beta \geq 0$, we have*

$$\tau^{\alpha,0}_{\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}},f_{\text{id}}}(\beta, 0) \quad = \quad \sup_{\substack{\mathbf{x},\mathbf{x}':\\ \mathfrak{d}_{\text{hist}}(\mathbf{x},\mathbf{x}')\leq\beta}} \inf_{\substack{Y:\\ \widehat{\partial}(\mathbf{x}',Y)\leq\alpha}} W^{\infty}(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}), \mathfrak{p}_Y) \leq \beta$$

w.r.t. the distortion ∂_{drop} and the metric $\mathfrak{d}_{\text{hist}}$. Here, input histograms to the mechanism $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ are restricted to t bars and $\tau = \alpha/t$.

Proof. For simplicity, we denote $[0, B)$ by $\mathcal{G}$. For any two histograms $\mathbf{x}, \mathbf{x}' \in \mathcal{H}_{\mathcal{G}}$ such that $\mathfrak{d}_{\text{hist}}(\mathbf{x}, \mathbf{x}') \leq \beta$, we will construct a r.v. Y over $\mathcal{H}_{\mathcal{G}}$ such that $\widehat{\partial}_{\text{drop}}(\mathbf{x}', \mathfrak{p}_Y) \leq \alpha$ and $W^{\infty}(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}), \mathfrak{p}_Y) \leq \beta$. The claim then immediately follows from this. Details follow.

Consider any two histograms $\mathbf{x}, \mathbf{x}' \in \mathcal{H}_{\mathcal{G}}$ such that $\mathfrak{d}_{\text{hist}}(\mathbf{x}, \mathbf{x}') \leq \beta$. Let $\mathfrak{d}_{\mathcal{G}}(\cdot, \cdot)$ denote the underlying metric over $\mathcal{G}$ (consists of t elements) and $|\mathbf{x}|$ denote number of elements in the histogram $\mathbf{x}$. By definition of $\mathfrak{d}_{\text{hist}}(\cdot, \cdot)$, we have $\mathfrak{d}_{\text{hist}}(\mathbf{x}, \mathbf{x}') = W^{\infty}(\frac{\mathbf{x}}{|\mathbf{x}|}, \frac{\mathbf{x}'}{|\mathbf{x}'|})$. Let ϕ be an optimal coupling of $\frac{\mathbf{x}}{|\mathbf{x}|}$ and $\frac{\mathbf{x}'}{|\mathbf{x}'|}$ such that

$$\mathfrak{d}_{\text{hist}}(\mathbf{x}, \mathbf{x}') = W^{\infty}(\frac{\mathbf{x}}{|\mathbf{x}|}, \frac{\mathbf{x}'}{|\mathbf{x}'|}) = \sup_{(a,b)\leftarrow\phi} \mathfrak{d}_{\mathcal{G}}(a, b) \leq \beta. \tag{36}$$

Using ϕ we define a transformation f_ϕ, which, when given a histogram $\mathbf{z}$ that is α-distorted from $\mathbf{x}$, returns $f_\phi(\mathbf{z})$ that is an α-distorted histogram from $\mathbf{x}'$. Recall that for a histogram $\mathbf{x}$ and $a \in \mathcal{G}$, we denote by $\mathbf{x}(a)$ the multiplicity of a in $\mathbf{x}$. Now, for any $b \in [0, B)$, we define $f_\phi(\mathbf{z})(b)$ as follows:

$$f_\phi(\mathbf{z})(b) := |\mathbf{x}'| \sum_{a\in\mathcal{G}} \frac{\mathbf{z}(a)\phi(a, b)}{\mathbf{x}(a)}.$$

The following claim is proved in Appendix G.1.

Claim 7. *For any $\mathbf{x} \in \mathcal{H}_{\mathcal{G}}$, if $\mathbf{z}$ is α-distorted from $\mathbf{x}$, then $f_\phi(\mathbf{z})$ is α-distorted from $\mathbf{x}'$.*

Recall that $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})$ outputs α-distorted histograms from $\mathbf{x}$. This suggests defining a r.v. $Y := f_\phi\left(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})\right)$ over $\mathcal{H}_{\mathcal{G}}$, whose distribution is given as follows:

$$\text{For } \mathbf{y} \in \mathcal{H}_{\mathcal{G}}, \text{ define } \Pr[Y = \mathbf{y}] := \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) \in f^{-1}_\phi(\mathbf{y})],$$

where $f_\phi^{-1}(\mathbf{y}) := \{\mathbf{z} \in \mathcal{H}_\mathcal{G} : f_\phi(\mathbf{z}) = \mathbf{y}\}$ is the inverse mapping of f_ϕ.

In the following two claims (which we prove in Appendix G.1), we show that the above defined Y satisfies $\widehat{\partial}_{\text{drop}}(\mathbf{x}', \mathfrak{p}_Y) \le \alpha$ and $W^\infty(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}), \mathfrak{p}_Y) \le \beta$.

Claim 8. $\widehat{\partial}_{\text{drop}}(\mathbf{x}', \mathfrak{p}_Y) \le \alpha$.

Claim 9. $W^\infty(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}), \mathfrak{p}_Y) \le \beta$.

It follows from Claim 8 and Claim 9 that $\inf_{Y:\widehat{\partial}(\mathbf{x}',Y)\le\alpha} W^\infty(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}), \mathfrak{p}_Y) \le \beta$. Since this holds for any two histograms $\mathbf{x}, \mathbf{x}' \in \mathcal{H}_\mathcal{G}$ such that $\mathfrak{d}_{\text{hist}}(\mathbf{x}, \mathbf{x}') \le \beta$, we have proved Lemma 11. □

Now, applying Theorem 1 to $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}} = \mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}} \circ \mathcal{M}^{t,[0,B)}_{\text{buc}}$, we get that $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$ is $(\alpha, \beta, 0)$-accurate.

This completes the proof of Theorem 4.

G.1 Details Omitted from the Proof of Theorem 4

Claim (Restating Claim 6)**.** *$\mathcal{M}^{w,[0,B)}_{\text{buc}}$ is $\left(0, \frac{w}{2}, 0\right)$-accurate for the identity function f_{id} over $\mathcal{H}_{[0,B)}$ w.r.t the metric $\mathfrak{d}_{\text{hist}}$.*

Proof. Since both f_{id} and $\mathcal{M}^{w,[0,B)}_{\text{buc}}$ are deterministic maps, on any input $\mathbf{x} \in \mathcal{H}_{[0,B)}$, we denote $\mathbf{x}$ (as the output of $f_{\text{id}}(\mathbf{x})$) and $\mathcal{M}^{w,[0,B)}_{\text{buc}}$ as point distributions over $\mathcal{H}_{[0,B)}$. Now, in order to prove the claim, we need to show that $W^\infty(\mathcal{M}^{w,[0,B)}_{\text{buc}}(\mathbf{x}), \mathbf{x}) \le \frac{w}{2}$ holds for any $\mathbf{x} \in \mathcal{H}_{[0,B)}$.

Fix any $\mathbf{x} \in \mathcal{H}_{[0,B)}$ and define $\mathbf{y} := \mathcal{M}^{w,[0,B)}_{\text{buc}}(\mathbf{x})$. Since $\mathbf{x}, \mathbf{y}$ are point distributions and the underlying metric is $\mathfrak{d}_{\text{hist}}$, we have $W^\infty(\mathbf{y}, \mathbf{x}) = \mathfrak{d}_{\text{hist}}(\mathbf{y}, \mathbf{x})$, where $\mathfrak{d}_{\text{hist}}$ is defined as $\mathfrak{d}_{\text{hist}}(\mathbf{y}, \mathbf{x}) = W^\infty(\frac{\mathbf{y}}{|\mathbf{y}|}, \frac{\mathbf{x}}{|\mathbf{x}|})$. Since $\mathbf{y}$ is a deterministic function of $\mathbf{x}$, $W^\infty(\frac{\mathbf{y}}{|\mathbf{y}|}, \frac{\mathbf{x}}{|\mathbf{x}|})$ is upper bounded by the maximum distance any point in $\mathbf{x}$ moves to form $\mathbf{y}$, which is equal to the the maximum distance of the center of a bucket from any point in that bucket, which is $\frac{w}{2}$. □

Claim (Restating Claim 7)**.** *For any $\mathbf{x} \in \mathcal{H}_\mathcal{G}$, if $\mathbf{z}$ is α-distorted from $\mathbf{x}$, then $f_\phi(\mathbf{z})$ is α-distorted from $\mathbf{x}'$.*

Proof. We need to show two things: (i) $f_\phi(\mathbf{z})(b) \le \mathbf{x}'(b)$ holds for every $b \in \mathcal{G}$, and (ii) $\sum_{b\in\mathcal{G}} f_\phi(\mathbf{z})(b) \ge (1-\alpha)\sum_{b\in\mathcal{G}} \mathbf{x}'(b)$. The first condition holds because $\mathbf{z}(a) \le \mathbf{x}(a), \forall a \in \mathcal{G}$ (since $\mathbf{z}$ is α-distorted from $\mathbf{x}$) and that $\sum_{a\in\mathcal{G}} \phi(a,b) = \frac{\mathbf{x}'(b)}{|\mathbf{x}'|}$. For the second condition,

$$\sum_{b\in\mathcal{G}} f_\phi(\mathbf{z})(b) = \sum_{b\in\mathcal{G}} |\mathbf{x}'| \sum_{a\in\mathcal{G}} \frac{\mathbf{z}(a)\phi(a,b)}{\mathbf{x}(a)} = |\mathbf{x}'| \sum_{a\in\mathcal{G}} \frac{\mathbf{z}(a)}{\mathbf{x}(a)} \sum_{b\in\mathcal{G}} \phi(a,b) \overset{\text{(a)}}{=} \frac{|\mathbf{x}'|}{|\mathbf{x}|} \sum_{a\in\mathcal{G}} \mathbf{z}(a) \overset{\text{(b)}}{\ge} (1-\alpha)|\mathbf{x}'|, \tag{37}$$

where (a) follows from $\sum_{b\in\mathcal{G}} \phi(a,b) = \frac{\mathbf{x}(a)}{|\mathbf{x}|}$ and (b) follows because $\mathbf{z}$ is α-distorted from $\mathbf{x}$, which implies that $\sum_{a\in\mathcal{G}} \mathbf{z}(a) = |\mathbf{z}| \ge (1-\alpha)|\mathbf{x}|$. Therefore, $f_\phi(\mathbf{z})$ is α-distorted from $\mathbf{x}'$. □

Claim (Restating Claim 8)**.** $\widehat{\partial}_{\text{drop}}(\mathbf{x}', \mathfrak{p}_Y) \le \alpha$.

Proof. Note that the support of $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})$ is the set of all α-distorted histograms from $\mathbf{x}$. We have shown in Claim 7 that for any $\mathbf{z} \in \mathcal{H}_\mathcal{G}$ such that $\partial_{\text{drop}}(\mathbf{x}, \mathbf{z}) \le \alpha$, we have $\partial_{\text{drop}}(\mathbf{x}', f_\phi(\mathbf{z})) \le \alpha$. This implies that $\sup_{\mathbf{y}\in\text{support}(f_\phi(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})))} \partial_{\text{drop}}(\mathbf{x}', \mathbf{y}) \le \alpha$, which in turn implies that $\widehat{\partial}_{\text{drop}}(\mathbf{x}', \mathfrak{p}_Y) \le \alpha$. □

Claim (Restating Claim 9)**.** $W^\infty(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}), \mathfrak{p}_Y) \le \beta$.

Proof. Define a coupling $\phi_\mathbf{x}$ of $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})$ and $\mathfrak{p}_Y$ over $\mathcal{H}_\mathcal{G} \times \mathcal{H}_\mathcal{G}$ as follows:

$$\phi_\mathbf{x}(\mathbf{z}, \mathbf{y}) := \begin{cases} \Pr[\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}) = \mathbf{z}] & \text{if } \mathbf{y} = f_\phi(\mathbf{z}), \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to verify that the above defined $\phi_{\mathbf{x}}$ is a valid coupling of $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})$ and $\mathfrak{p}_Y$, i.e., its first marginal is equal to $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})$ and the second marginal is equal to $\mathfrak{p}_Y$. Note that $\phi_{\mathbf{x}}(\mathbf{z},\mathbf{y})$ is non-zero only when $\mathbf{y} = f_\phi(\mathbf{z})$. This implies that

$$W^\infty(\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x}), \mathfrak{p}_Y) \le \sup_{(\mathbf{z},\mathbf{y})\leftarrow\phi_{\mathbf{x}}} \mathfrak{d}_{\text{hist}}(\mathbf{z},\mathbf{y}) = \sup_{(\mathbf{z},f_\phi(\mathbf{z}))\leftarrow\phi_{\mathbf{x}}} \mathfrak{d}_{\text{hist}}(\mathbf{z}, f_\phi(\mathbf{z})) \le \beta,$$

where the last inequality follows from Claim 10 (stated and proven below) and using the fact that $\mathbf{z} \sim \mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}(\mathbf{x})$ is α-distorted from $\mathbf{x}$. □

Claim 10. *Let $\mathbf{x}, \mathbf{x}' \in \mathcal{H}_{\mathcal{G}}$ be such that $\mathfrak{d}_{\text{hist}}(\mathbf{x},\mathbf{x}') \le \beta$. Then, for any $\mathbf{z}$ that is α-distorted from $\mathbf{x}$, we have $\mathfrak{d}_{\text{hist}}(\mathbf{z}, f_\phi(\mathbf{z})) \le \mathfrak{d}_{\text{hist}}(\mathbf{x},\mathbf{x}') \le \beta$.*

Proof. Define $\phi'(a,b) = \frac{\mathbf{z}(a)|\mathbf{x}|\phi(a,b)}{\mathbf{x}(a)|\mathbf{z}|}$. For any $a \in \mathcal{G}$, its first marginal is equal to $\sum_{b\in\mathcal{G}} \phi'(a,b) = \frac{\mathbf{z}(a)}{|\mathbf{z}|}$. For any $b \in \mathcal{G}$, its second marginal is equal to $\sum_{a\in\mathcal{G}} \phi'(a,b) = \frac{|\mathbf{x}|}{|\mathbf{z}|}\sum_{a\in\mathcal{G}} \frac{\mathbf{z}(a)}{\mathbf{x}(a)}\phi(a,b) = \frac{|\mathbf{x}|}{|\mathbf{x}'||\mathbf{z}|} f_\phi(\mathbf{z})(b)$. We would like to say that the quantity on the RHS is equal to $\frac{f_\phi(\mathbf{z})(b)}{|f_\phi(\mathbf{z})|}$. We show this as follows: Since $|\mathbf{z}| \ge (1-\alpha)|\mathbf{x}|$, there exists $c \ge 0$ such that $|\mathbf{z}| = (1-\alpha+c)|\mathbf{x}|$. If we put this instead of $|\mathbf{z}| \ge (1-\alpha)|\mathbf{x}|$ in (37), we would get $\sum_{b\in\mathcal{G}} f_\phi(\mathbf{z})(b) = (1-\alpha+c)|\mathbf{x}'|$. With these substitutions, we get $\frac{|\mathbf{x}|}{|\mathbf{x}'||\mathbf{z}|} f_\phi(\mathbf{z})(b) = \frac{f_\phi(\mathbf{z})(b)}{\sum_{b\in\mathcal{G}} f_\phi(\mathbf{z})(b)}$, which implies that the second marginal of ϕ' is equal to $\sum_{a\in\mathcal{G}} \phi'(a,b) = \frac{f_\phi(\mathbf{z})(b)}{|f_\phi(\mathbf{z})|}$ for any $b \in \mathcal{G}$.

This means that $\phi'(a,b)$ is a valid coupling of $\mathbf{z}, f_\phi(\mathbf{z})$. This implies that

$$\mathfrak{d}_{\text{hist}}(\mathbf{z}, f_\phi(\mathbf{z})) = W^\infty(\mathbf{z}, f_\phi(\mathbf{z})) \le \sup_{(a',b')\leftarrow\phi'} \mathfrak{d}_{\mathcal{G}}(a',b') \overset{\text{(c)}}{\le} \sup_{(a',b')\leftarrow\phi} \mathfrak{d}_{\mathcal{G}}(a',b') = \mathfrak{d}_{\text{hist}}(\mathbf{x},\mathbf{x}') \le \beta,$$

where (c) holds because $\text{support}(\phi') \subseteq \text{support}(\phi)$ (by the definition of ϕ'). □

H Omitted Proofs from Section 5.2 – Histogram-Based-Statistics

In this section, we will prove Theorem 5, Corollary 2, and Corollary 3.

H.1 Proof of Theorem 5 – Any Histogram-Based-Statistic

First we show the flexible accuracy and then the differential privacy guarantee of our composed mechanism $\mathcal{M}^{\alpha,\beta,[0,B)}_{f_{\text{HBS}}} = f_{\text{HBS}} \circ \mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$.

Flexible accuracy. Note that f_{HBS} (as a mechanism) for computing f_{HBS} is $(0,0,0)$-accurate, and we have from Theorem 4 that $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$ is $(\alpha,\beta,0)$-accurate for the identity function f_{id} w.r.t. the distortion measure ∂_{drop} and the metric $\mathfrak{d}_{\text{hist}}$. Applying Theorem 1, we get that $\mathcal{M}^{\alpha,\beta,[0,B)}_{f_{\text{HBS}}}$ is $(\alpha + \sigma_{f_{\text{id}}}(0), \tau^{0,0}_{f_{\text{HBS}},f_{\text{HBS}}}(0,\beta), 0)$-accurate. It follows from (8) (by substituting $\mathcal{M} = f_{\text{HBS}}$ as a mechanism for $f = f_{\text{HBS}}$) and the definition of the metric sensitivity (11), that $\tau^{0,0}_{f_{\text{HBS}},f_{\text{HBS}}}(0,\beta) = \Delta_{f_{\text{HBS}}}(\beta)$. We have also noted after (5) that the distortion sensitivity of any randomized function at zero is equal to zero; in particular, $\sigma_{f_{\text{id}}}(0) = 0$. Substituting these in the flexible accuracy parameters of $\mathcal{M}^{\alpha,\beta,[0,B)}_{f_{\text{HBS}}}$, we get that $\mathcal{M}^{\alpha,\beta,[0,B)}_{f_{\text{HBS}}}$ is $(\alpha, \Delta_{f_{\text{HBS}}}(\beta), 0)$-accurate for f_{HBS} w.r.t. distortion ∂_{drop} and metric $\mathfrak{d}_{\mathcal{A}}$.

Differential privacy. Since $\mathcal{M}_{\text{BucHist}}$ is $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)}\right)$-DP, and $\mathcal{M}_{f_{\text{HBS}}}$ is a post-processing of $\mathcal{M}_{\text{BucHist}}$, it follows that $\mathcal{M}_{f_{\text{HBS}}}$ is also differentially private with the same parameters.

This completes the proof of Theorem 5.

H.2 Computing the Maximum or Minimum Element of a Multi-set

We define $f_{\max}$ (or simply max) for histograms over real numbers as $f_{\max}(\mathbf{x}) := \max\{g : \mathbf{x}(g) > 0\}$. Similarly, we can define $f_{\min}$ (or simply min) as $f_{\min}(\mathbf{x}) := \min\{g : \mathbf{x}(g) > 0\}$. We give our result for $f_{\max}$ only; the same result holds for $f_{\min}$ as well.

Corollary 2. *On inputs of size n, $\mathcal{M}_{\max}^{\alpha,\beta,[0,B)}$ is $(\alpha,\beta,0)$-accurate for $f_{\max}$ w.r.t. the distortion ∂_{drop} and the standard distance metric over $\mathbb{R}$. Furthermore, for any $\epsilon > 0$, and $\tau = \alpha(\frac{2\beta}{B})$, if $\epsilon\tau n \geq 2$, then $\mathcal{M}_{\max}^{\alpha,\beta,[0,B)}$ is $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)}\right)$-DP.*

Proof. For any two histograms $\mathbf{y}$, $\mathbf{y}'$, by definition of $\mathfrak{d}_{\text{hist}}(\mathbf{y},\mathbf{y}') = W^{\infty}(\frac{\mathbf{y}}{|\mathbf{y}|}, \frac{\mathbf{y}'}{|\mathbf{y}'|})$ and $f_{\max}$, it follows that $|f_{\max}(\mathbf{y}) - f_{\max}(\mathbf{y}')| \leq \mathfrak{d}_{\text{hist}}(\mathbf{y},\mathbf{y}')$. Using this in (11) implies that $\Delta_{f_{\max}}(\beta) \leq \beta$ for every $\beta \geq 0$. Then, the corollary follows from Theorem 5, with $f_{\text{HBS}} = f_{\max}$. □

We can instantiate Corollary 2 with different parameter settings to achieve favorable privacy-accuracy tradeoffs. See Appendix H.4 for more details.

H.3 Computing the Support of a Multi-set

f_{supp} (or simply support) is defined as $f_{\text{supp}}(\mathbf{x}) := \{g : \mathbf{x}(g) > 0\}$, which maps a multiset to the set that forms its support. To measure accuracy, we use a metric $\mathfrak{d}_{\text{supp}}$ over the set of finite subsets of $\mathbb{R}$: for any two finite subsets $\mathcal{S}_1, \mathcal{S}_2 \subseteq \mathbb{R}$, define

$$\mathfrak{d}_{\text{supp}}(\mathcal{S}_1,\mathcal{S}_2) := \max\left\{\max_{s_1\in\mathcal{S}_1}\min_{s_2\in\mathcal{S}_2}|s_1 - s_2|, \max_{s_2\in\mathcal{S}_2}\min_{s_1\in\mathcal{S}_1}|s_2 - s_1|\right\}.$$

$\mathfrak{d}_{\text{supp}}$ measures the farthest that a point in one of the sets is from any point on the other set. For example, if $s_i^{\min} := \min_{s\in\mathcal{S}_i}\{s\}$ and $s_i^{\max} := \max_{s\in\mathcal{S}_i}\{s\}$ denote the minimum and the maximum elements of the set $\mathcal{S}_i$ (for $i = 1, 2$), respectively, then it can be verified that $\mathfrak{d}_{\text{supp}}(\mathcal{S}_1,\mathcal{S}_2) = \max\{|s_1^{\min} - s_2^{\min}|, |s_1^{\max} - s_2^{\max}|\}$.

Corollary 3. *On inputs of size n, $\mathcal{M}_{\text{supp}}^{\alpha,\beta,[0,B)}$ is $(\alpha,\beta,0)$-accurate for f_{supp} w.r.t. the distortion ∂_{drop} and metric $\mathfrak{d}_{\text{supp}}$. Furthermore, for any $\epsilon > 0$, and $\tau = \alpha(\frac{2\beta}{B})$, if $\epsilon\tau n \geq 2$, then $\mathcal{M}_{\text{supp}}^{\alpha,\beta,[0,B)}$ is $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)}\right)$-DP.*

Proof. Since $\mathfrak{d}_{\text{supp}}(\mathcal{S}_1,\mathcal{S}_2)$ is the difference between the maximum or the minimum elements of $\mathcal{S}_1$ and $\mathcal{S}_2$, it follows that for any two histograms $\mathbf{y}$ and $\mathbf{y}'$, we have $\mathfrak{d}_{\text{supp}}(f_{\text{supp}}(\mathbf{y}), f_{\text{supp}}(\mathbf{y}')) \leq \max\{|f_{\max}(\mathbf{y}) - f_{\max}(\mathbf{y}')|, |f_{\min}(\mathbf{y}) - f_{\min}(\mathbf{y}')|\}$, where $|f_{\max}(\mathbf{y}) - f_{\max}(\mathbf{y}')| \leq \mathfrak{d}_{\text{hist}}(\mathbf{y},\mathbf{y}')$ (from Corollary 2), and similarly, $|f_{\min}(\mathbf{y}) - f_{\min}(\mathbf{y}')| \leq \mathfrak{d}_{\text{hist}}(\mathbf{y},\mathbf{y}')$. Using this in (11) implies that $\Delta_{f_{\text{supp}}}(\beta) \leq \beta$ for every $\beta \geq 0$. Then, the corollary follows from Theorem 5, with $f_{\text{HBS}} = f_{\text{supp}}$. □

We can instantiate Corollary 3 with different parameter settings to achieve favorable privacy-accuracy tradeoffs. See Appendix H.4 for more details.

H.4 Choosing the Parameters

As mentioned in Remark 2 for Theorem 3, there are many choices of ϵ, τ for which we can get favorable privacy, accuracy parameters in Theorems 4, 5, and Corollaries 2, 3. For concreteness, in the following, we illustrate the privacy accuracy trade-off by choosing parameters for the $\mathcal{M}_{\max}^{\alpha,\beta,[0,B)}$ mechanism in Corollary 2; the same result applies to Theorems 4, 5, and Corollary 3 as well.

If we choose $\epsilon = \frac{1}{\sqrt{\tau n}}$ and τ is such that $\frac{1}{\epsilon} = \sqrt{\tau n} \geq 2$, then by dropping only $\alpha n = \frac{1}{\epsilon^2}\frac{2\beta}{B}$ elements from the entire dataset, the mechanism $\mathcal{M}_{\max}^{\alpha,\beta,[0,B)}$ achieves $\left(\frac{1}{\sqrt{\tau n}}, \frac{e^{-\Omega(\sqrt{\tau n})}}{\sqrt{\tau n}}\right)$-differential privacy. If β/B is a small constant (say, $1/100$), which corresponds to perturbing the output by a small constant fraction of the whole range B, then by dropping only $\alpha n = O(\frac{1}{\epsilon^2})$ elements, $\mathcal{M}_{\max}^{\alpha,\beta,[0,B)}$ achieves $(\epsilon, \epsilon e^{-\Omega(\frac{1}{\epsilon})})$-differential privacy. We can set any τ that satisfies $\frac{1}{\epsilon} = \sqrt{\tau n} \geq 2$ in this result. For example,

By setting $\epsilon = \frac{1}{(\log n)^2}$, we get that by dropping only $O((\log n)^4)$ elements from the entire dataset, $\mathcal{M}_{\max}^{\alpha,\beta,[0,B)}$ achieves $(\frac{1}{(\log n)^2}, \frac{n^{-\Omega(\log n)}}{(\log n)^2})$-differential privacy while incurring only a small constant error (of the entire range) in the output.

Note that in the above setting of parameters, we take $\epsilon = \frac{1}{\sqrt{\tau n}}$, which implies that the bound on δ can at best be a small constant for any constant ϵ. This is because $\epsilon\tau n = \sqrt{\tau n} = \frac{1}{\epsilon}$ is a constant, which implies that

$\delta = \epsilon e^{-\Omega(\frac{1}{\epsilon})}$ will be a constant too. Therefore, for getting privacy guarantees with small constant ϵ such that δ (exponentially) decays with n, we will work with the general privacy result of $(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)})$-DP as in Corollary 2. For example,

By setting $\epsilon = 0.1$ and $\tau = \frac{1}{n^c}$ (for any $c \in (0,1)$), we get that by dropping only $\alpha n = \tau n \frac{B}{2\beta} = O(n^{1-c})$ elements from the entire dataset, $\mathcal{M}_{\max}^{\alpha,\beta,[0,B)}$ achieves $(0.1, e^{-\Omega(n^{1-c})})$-differential privacy while incurring only a small constant error (of the entire range) in the output.

For other parameter settings, see the result on page 2 after we stated our informal result for max.

I Further Applications: Beyond ∂_{drop}

Useful variants of Theorem 5 can be obtained with measures of distortion other than ∂_{drop}. In the following, we define two distortions ∂_{move} and $\partial^{\eta}_{\text{drmv}}$, respectively, where ∂_{move} allows moving/perturbing of data points and $\partial^{\eta}_{\text{drmv}}$ allows both dropping and moving.

1. **Perturbing/Moving elements:** For finite $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, we define ∂_{move}, a measure of distortion for moving elements, as follows:
$$\partial_{\text{move}}(\mathbf{x}, \mathbf{y}) = \begin{cases} W^{\infty}(\frac{\mathbf{x}}{|\mathbf{x}|}, \frac{\mathbf{y}}{|\mathbf{y}|}) & \text{if } |\mathbf{x}| = |\mathbf{y}|, \\ \infty & \text{otherwise,} \end{cases} \tag{38}$$
where $\frac{\mathbf{x}}{|\mathbf{x}|}$ (similarly, $\frac{\mathbf{y}}{|\mathbf{y}|}$) is treated as a probability vector of size $|\mathcal{G}|$, indexed by the elements of $\mathcal{G}$; the i'th element of $\frac{\mathbf{x}}{|\mathbf{x}|}$ is equal to $\frac{\mathbf{x}(i)}{|\mathbf{x}|}$. We show in that Claim 11 in Appendix I.1 that ∂_{move} is a metric.
2. **Both dropping and moving elements:** For finite $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, we define $\partial^{\eta}_{\text{drmv}}$, a measure of distortion for both moving and dropping elements, as follows:
$$\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{y}) = \inf_{\mathbf{z}} \left(\partial_{\text{drop}}(\mathbf{x}, \mathbf{z}) + \eta \cdot \partial_{\text{move}}(\mathbf{z}, \mathbf{y}) \right). \tag{39}$$
We show in Claim 12 in Appendix I.1 that $\partial^{\eta}_{\text{drmv}}$ is a quasi-metric.

The following theorem provides the privacy and accuracy guarantees of $\mathcal{M}_{f_{\text{HBS}}}^{\alpha,\beta,[0,B)}$ (defined in (10)) w.r.t. the distortion measure $\partial^{\eta}_{\text{drmv}}$.

Theorem 6. *On inputs of size n, $\mathcal{M}_{f_{\text{HBS}}}^{\alpha,\beta,[0,B)}$ is $(\alpha + \eta\beta, 0, 0)$-accurate for f_{HBS} w.r.t. the distortion measure $\partial^{\eta}_{\text{drmv}}$. Furthermore, for any $\epsilon > 0$, and $\tau = \alpha(\frac{2\beta}{B})$, if $\epsilon\tau n \geq 2$, then $\mathcal{M}_{f_{\text{HBS}}}^{\alpha,\beta,[0,B)}$ is $\left(\epsilon, \epsilon e^{-\Omega(\epsilon\tau n)}\right)$-DP.*

We prove Theorem 6 after the following two remarks.

Remark 4. *This is analogous to Theorem 5, but with the important difference that it does not refer to the metric sensitivity of the function f_{HBS}, and does not even require a metric over its codomain $\mathcal{A}$. This makes this result applicable to complex function families like maximum-margin separators or neural net classifiers. However, the accuracy notion uses a measure of distortion that allows dropping a (small) fraction of the data* and *(slightly) moving all data points, which may or may not be acceptable to all applications.*

Remark 5 (Extending the results from $[0, B)$ to $[0, B)^d$)**.** *Note that the bucketing mechanism $\mathcal{M}_{\text{buc}}^{w,[0,B)}$ and the bucketed-histogram mechanism $\mathcal{M}_{\text{BucHist}}^{\alpha,\beta,[0,B)}$ in Algorithm 3 are given for the ground set $\mathcal{G} = [0, B)$. However, as mentioned in Footnote 7, they can easily be extended to the d-dimensional ground set $\mathcal{G} = [0, B)^d$, and we present the d-dimensional analogues of the above two mechanisms in Appendix J. All our results in Theorem 4, Theorem 5, and Theorem 6 will hold verbatim with these generalized mechanisms, except for the value of τ, which will be replaced by $\tau = \alpha(\frac{2\beta}{B\sqrt{d}})^d$; see Appendix J for a proof of this.*

Proof of Theorem 6. Since $\mathcal{M}_{f_{\text{HBS}}}^{\alpha,\beta,[0,B)}$ is the same mechanism for which the results in Theorem 4 hold, the same privacy results as in Theorem 4 will also hold here. In the rest of this proof, we prove the flexible accuracy part.

Since f_{HBS} is a $(0, 0, 0)$-accurate mechanism for f_{HBS} (which implies that $\Delta_{f_{\text{HBS}}}(0) = 0$), in order to prove the accuracy guarantee of $\mathcal{M}_{f_{\text{HBS}}}^{\alpha,\beta,[0,B)}$, it suffices to show that $\mathcal{M}_{\text{BucHist}}^{\alpha,\beta,[0,B)}$ is $(\alpha + \eta\beta, 0, 0)$-accurate w.r.t. $\partial^{\eta}_{\text{drmv}}$. Note

that $\mathcal{M}^{\alpha,\beta,[0,B)}_{f_{\text{HBS}}} = \mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}} \circ \mathcal{M}^{w,[0,B)}_{\text{buc}}$. On any input $\mathbf{x}$, first we produce an intermediate bucketed output $\mathbf{z} := \mathcal{M}^{w,[0,B)}_{\text{buc}}(\mathbf{x})$ and then produce $\mathbf{y} := \mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}}(\mathbf{z})$ as the final output. We have shown in Claim 6 in the proof of Theorem 4 that the output $\mathbf{z}$ produced by $\mathcal{M}^{w,[0,B)}_{\text{buc}}$ on input $\mathbf{x}$ satisfies $W^{\infty}(\mathbf{x},\mathbf{z}) \le \beta$. This, by definition of the distortion ∂_{move}, implies $\partial_{\text{move}}(\mathbf{x},\mathbf{z}) \le \beta$. We have also shown in the proof of Theorem 3 that the output $\mathbf{y}$ produced by $\mathcal{M}^{\tau,\epsilon,[0,B)}_{\text{STLap}}$ on input $\mathbf{z}$ satisfies $\partial_{\text{drop}}(\mathbf{z},\mathbf{y}) \le \alpha$. So, we have $\partial_{\text{move}}(\mathbf{x},\mathbf{z}) \le \beta$ and $\partial_{\text{drop}}(\mathbf{z},\mathbf{y}) \le \alpha$. This, together with Lemma 12, implies the existence of a histogram $\mathbf{s}$ such that $\partial_{\text{drop}}(\mathbf{x},\mathbf{s}) \le \alpha$ and $\partial_{\text{move}}(\mathbf{s},\mathbf{y}) \le \beta$. Using these in the definition of $\partial^{\eta}_{\text{drmv}}$ in (39) implies that $\partial^{\eta}_{\text{drmv}}(\mathbf{x},\mathbf{y}) \le \alpha + \eta\beta$. Since we have attributed all the error to the input distortion, we have shown that $\mathcal{M}^{\alpha,\beta,[0,B)}_{\text{BucHist}}$ is $(\alpha+\eta\beta, 0, 0)$-accurate w.r.t. the distortion $\partial^{\eta}_{\text{drmv}}$.

This completes the proof of Theorem 6. □

I.1 Proofs of ∂_{move} Being a Metric and $\partial^{\eta}_{\text{drmv}}$ Being a Quasi-Metric

Showing that $\partial_{\text{move}}(\cdot,\cdot)$ is a metric is trivial; however, showing that $\partial^{\eta}_{\text{drmv}}$ is a quasi-metric is non-trivial, and most of this section is devoted to proving that.

Claim 11. *$\partial_{\text{move}}(\cdot,\cdot)$ is a metric.*

Proof. Since $\partial_{\text{move}}(\cdot,\cdot)$ is defined as the ∞-Wasserstein distance between normalized histograms, it suffices to show that the ∞-Wasserstein distance is a metric. We need to show three things for any triple of distributions P, Q, R over a metric space $(\Omega, \mathfrak{d})$: (i) $W^{\infty}(P,Q) \ge 0$ and equality holds if and only if $P = Q$, (ii) $W^{\infty}(P,Q) = W^{\infty}(Q,P)$, and (iii) $W^{\infty}(P,R) \le W^{\infty}(P,Q) + W^{\infty}(Q,R)$.

By definition, $W^{\infty}(P,R) = \inf_{\phi\in\Phi(P,R)} \sup_{(x,z):\phi(x,z)\neq 0} \mathfrak{d}(x,z)$. Now, the first two conditions follow because $\mathfrak{d}$ is a metric, and the last condition (triangle inequality) we show in Lemma 3 in Appendix A.1.

Note that when $|\mathbf{x}| = |\mathbf{y}| = 0$, the Wasserstein distance is undefined, but we have defined $\partial_{\text{move}}(\mathbf{x},\mathbf{y})$ in this case separately as 0, which is consistent with the properties of a metric. □

We first give an intermediate result (Lemma 12 below) which will be used in proving that $\partial^{\eta}_{\text{drmv}}$ is a quasi-metric. The result of this lemma is also used in the proof of Theorem 6.

Lemma 12. *Let $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ be any three histograms over a ground set $\mathcal{G}$,associated with a metric $\mathfrak{d}$, such that $\partial_{\text{move}}(\mathbf{x},\mathbf{z}) = \alpha_1$ and $\partial_{\text{drop}}(\mathbf{z},\mathbf{y}) = \alpha_2$ with $\alpha_1 \ge 0$ and $\alpha_2 < 1$. Then there exists a histogram $\mathbf{s}$ such that $\partial_{\text{drop}}(\mathbf{x},\mathbf{s}) = \alpha_2$ and $\partial_{\text{move}}(\mathbf{s},\mathbf{y}) \le \alpha_1$.*

Proof. Using the definitions of ∂_{drop} and ∂_{move}, we have the following:

Z.1 $|\mathbf{x}| = |\mathbf{z}|$
Z.2 $W^{\infty}(\frac{\mathbf{x}}{|\mathbf{x}|}, \frac{\mathbf{z}}{|\mathbf{z}|}) \le \alpha_1$. We will use ϕ_z to denote the optimal joint distribution which achieves the infimum in the definition of $W^{\infty}(\frac{\mathbf{x}}{|\mathbf{x}|}, \frac{\mathbf{z}}{|\mathbf{z}|})$.
Z.3 $|\mathbf{y}| = (1-\alpha_2)|\mathbf{z}|$
Z.4 For all $g \in \mathcal{G}$, $0 \le \mathbf{y}(g) \le \mathbf{z}(g)$

Now we want to prove the existence of a histogram $\mathbf{s}$ with the following property:

S.1 $|\mathbf{s}| = (1-\alpha_2)|\mathbf{x}|$
S.2 For all $g \in \mathcal{G}$, $0 \le \mathbf{s}(g) \le \mathbf{x}(g)$
S.3 $|\mathbf{s}| = |\mathbf{y}|$
S.4 $W^{\infty}(\frac{\mathbf{s}}{|\mathbf{s}|}, \frac{\mathbf{y}}{|\mathbf{y}|}) \le \alpha_1$.

Consider the following joint distribution ϕ_s:

$$\phi_s(g_x, g_y) = \begin{cases} \frac{1}{1-\alpha_2}\phi_z(g_x,g_y)\frac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)} & \text{if } \mathbf{z}(g_y) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{40}$$

We denote the first marginal of ϕ_s by $\frac{\mathbf{s}}{|\mathbf{s}|}$, where $\mathbf{s}$ corresponds to the histogram that we want to show.

By definition, for all $g_x, g_y \in \mathcal{G}$, we have $\phi_s(g_x, g_y) \geq 0$. Also note that, if $\mathbf{z}(g_y) = 0$, then for all $g_x \in \mathcal{G}$, we have $\phi_z(g_x, g_y) = 0$; this is because $\frac{\mathbf{z}}{|\mathbf{z}|}$ is the second marginal of ϕ_z. Now we show that the above-defined ϕ_s satisfies properties S.1-S.4 – we show these in the sequence of S.4, S.3, S.1, S.2.

- **Proof of S.4.** Note that the first marginal of ϕ_s is assumed to be $\frac{\mathbf{s}}{|\mathbf{s}|}$. Now we show that its second marginal is $\frac{\mathbf{y}}{|\mathbf{y}|}$ and that $\max_{(g_x,g_y)\leftarrow\phi_s} \mathfrak{d}(g_x, g_y) \leq \alpha_1$. Note that these together imply that $W^{\infty}(\frac{\mathbf{s}}{|\mathbf{s}|}, \frac{\mathbf{y}}{|\mathbf{y}|}) \leq \alpha_1$.
  - *Second marginal of ϕ_s is $\frac{\mathbf{y}}{|\mathbf{y}|}$:* We show it in two parts, first for $g_y \in \mathcal{G}$ for which $\mathbf{z}(g_y) = 0$ and then for the rest of the $g_y \in \mathcal{G}$. Note that when $\mathbf{z}(g_y) = 0$, we have from Z.4 that $\mathbf{y}(g_y) = 0$. Now we show that $\int_{\mathcal{G}} \phi_s(g_x, g_y)\, \mathrm{d}g_x = 0$. It follows from (40) that for all g_y such that $\mathbf{z}(g_y) = 0$, we have $\phi_s(g_x, g_y) = 0, \forall g_x \in \mathcal{G}$, which implies that $\int_{\mathcal{G}} \phi_s(g_x, g_y)\, \mathrm{d}g_x = 0$. Now we analyze the case when $\mathbf{z}(g_y) > 0$.

$$\begin{aligned}
\int_{\mathcal{G}} \phi_s(g_x, g_y)\, \mathrm{d}g_x &= \int_{\mathcal{G}} \frac{1}{1-\alpha_2} \phi_z(g_x, g_y) \frac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)}\, \mathrm{d}g_x && \text{(using (40))} \\
&= \frac{1}{1-\alpha_2} \frac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)} \int_{\mathcal{G}} \phi_z(g_x, g_y)\, \mathrm{d}g_x \\
&= \frac{1}{1-\alpha_2} \frac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)} \frac{\mathbf{z}(g_y)}{|\mathbf{z}|} && \text{(using Z.2)} \\
&= \frac{\mathbf{y}(g_y)}{(1-\alpha_2)|\mathbf{z}|} \\
&= \frac{\mathbf{y}(g_y)}{|\mathbf{y}|}. && \text{(Using Z.3)}
\end{aligned}$$

  - *$W^{\infty}(\frac{\mathbf{s}}{|\mathbf{s}|}, \frac{\mathbf{y}}{|\mathbf{y}|}) \leq \alpha_1$:* We have shown that the first and the second marginals of ϕ_s are $\frac{\mathbf{s}}{|\mathbf{s}|}$ and $\frac{\mathbf{y}}{|\mathbf{y}|}$, respectively. So, it suffices to show that $\max_{(g_x,g_y)\leftarrow\phi_s} \mathfrak{d}(g_x, g_y) \leq \alpha_1$. Consider any pair $(g_x, g_y) \in \mathcal{G}^2$ s.t. $\phi_s(g_x, g_y) > 0$. This is possible only if $\phi_z(g_x, g_y) > 0$ (see (40)), which, when combined with Z.2, gives $\mathfrak{d}(g_x, g_y) \leq \alpha_1$. Hence, for any pair $(g_x, g_y) \in \mathcal{G}^2$ s.t. $\phi_s(g_x, g_y) > 0$, we have $\mathfrak{d}(g_x, g_y) \leq \alpha_1$.
- **Proof of S.3.** Note that (40) gives the normalized $\mathbf{s}$, but we still have the freedom to choose $|\mathbf{s}|$. To satisfy S.3, we set $|\mathbf{s}| = |\mathbf{y}|$.
- **Proof of S.1.** Note that S.1 is already satisfied using Z.1, Z.3, and S.3.
- **Proof of S.2.** Let us denote $\{g \in \mathcal{G} \mid \mathbf{z}(g) > 0\}$ by $\mathcal{G}_z$. We will show that for any $g \in \mathcal{G}$, we have $\mathbf{x}(g) - \mathbf{s}(g) \geq 0$:

$$\begin{aligned}
\mathbf{x}(g) - \mathbf{s}(g) &= |\mathbf{x}| \int_{\mathcal{G}} \phi_z(g, g_y)\, \mathrm{d}g_y - |\mathbf{s}| \int_{\mathcal{G}} \phi_s(g, g_y)\, \mathrm{d}g_y && \text{(using Z.2 and S.4)} \\
&= |\mathbf{x}| \int_{\mathcal{G}_z} \phi_z(g, g_y)\, \mathrm{d}g_y - |\mathbf{s}| \int_{\mathcal{G}} \phi_s(g, g_y)\, \mathrm{d}g_y && \text{(Since } \mathbf{z}(g_y) = 0 \Rightarrow \phi_z(g, g_y) = 0, \forall g \in \mathcal{G}\text{; Z.2)} \\
&= |\mathbf{x}| \int_{\mathcal{G}_z} \phi_z(g, g_y)\, \mathrm{d}g_y - |\mathbf{s}| \int_{\mathcal{G}_z} \frac{1}{1-\alpha_2} \phi_z(g, g_y) \frac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)}\, \mathrm{d}g_y && \text{(using (40))} \\
&= |\mathbf{x}| \int_{\mathcal{G}_z} \phi_z(g, g_y)\, \mathrm{d}g_y - \frac{(1-\alpha_2)|\mathbf{x}|}{1-\alpha_2} \int_{\mathcal{G}_z} \phi_z(g, g_y) \frac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)}\, \mathrm{d}g_y && \text{(using S.1)} \\
&= |\mathbf{x}| \int_{\mathcal{G}_z} \phi_z(g, g_y)\, \mathrm{d}g_y - |\mathbf{x}| \int_{\mathcal{G}_z} \phi_z(g, g_y) \frac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)}\, \mathrm{d}g_y \\
&= |\mathbf{x}| \int_{\mathcal{G}_z} \phi_z(g, g_y) \left(1 - \frac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)}\right) \mathrm{d}g_y \\
&\geq 0 && \text{(using Z.4, } \tfrac{\mathbf{y}(g_y)}{\mathbf{z}(g_y)} \leq 1\text{)}
\end{aligned}$$

Thus, we have shown that the joint distribution ϕ_s defined in (40) satisfies all four properties S.1-S.4. This completes the proof of Lemma 12. □

Claim 12. *For all $\eta \in \mathbb{R}_{\geq 0}$, $\partial^{\eta}_{\mathrm{drmv}}(\cdot, \cdot)$ is a quasi metric.*

Proof. Note that both ∂_{drop} and ∂_{move} are quasi-metrics. Hence, for any $\mathbf{x}, \mathbf{y}$, $\partial_{\text{drop}}(\mathbf{x}, \mathbf{y}) \geq 0$ and $\partial_{\text{move}}(\mathbf{x}, \mathbf{y}) \geq 0$. This implies that for every $\mathbf{x}, \mathbf{y}$, $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{y}) \geq 0$. Now we one by one prove that $\partial^{\eta}_{\text{drmv}}$ satisfies the properties of quasi-metric:

Property #1: For all $\mathbf{x}$ *and* $\mathbf{y}$, $\mathbf{x} = \mathbf{y} \Leftrightarrow \partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{y}) = 0$.

1. For all $\mathbf{x}$, $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{x}) = 0$:

$$\begin{aligned}\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{x}) &= \inf_{\mathbf{z}} \left(\partial_{\text{drop}}(\mathbf{x}, \mathbf{z}) + \eta \cdot \partial_{\text{move}}(\mathbf{z}, \mathbf{x})\right) \\ &\leq \partial_{\text{drop}}(\mathbf{x}, \mathbf{x}) + \eta \cdot \partial_{\text{move}}(\mathbf{x}, \mathbf{x}) \qquad \text{(infimum over a set is } \leq \text{ the value at any fixed point in set)} \\ &= 0\end{aligned}$$

   Since $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{x}) \geq 0$ as well as ≤ 0, $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{x}) = 0$.
2. For all $\mathbf{x}, \mathbf{y}$, $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{y}) = 0 \Rightarrow \mathbf{x} = \mathbf{y}$:
   $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{y}) = 0$ implies that $\inf_{\mathbf{z}} \left(\partial_{\text{drop}}(\mathbf{x}, \mathbf{z}) + \eta \cdot \partial_{\text{move}}(\mathbf{z}, \mathbf{y})\right) = 0$. As both $\partial_{\text{drop}}(\mathbf{x}, \mathbf{z})$ and $\partial_{\text{move}}(\mathbf{z}, \mathbf{y})$ are ≥ 0 for any value of $\mathbf{x}, \mathbf{y}, \mathbf{z}$, this is possible only if $\partial_{\text{drop}}(\mathbf{x}, \mathbf{z}) = \partial_{\text{move}}(\mathbf{z}, \mathbf{y}) = 0$ which means that $\mathbf{x} = \mathbf{z} = \mathbf{y}$. Hence $\mathbf{x} = \mathbf{y}$.

Property #2: For all $\mathbf{x}$, $\mathbf{y}$ *and* $\mathbf{z}$, $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{z}) \leq \partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{y}) + \partial^{\eta}_{\text{drmv}}(\mathbf{y}, \mathbf{z})$.

We assume that the infimum in both $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{y})$ and $\partial^{\eta}_{\text{drmv}}(\mathbf{y}, \mathbf{z})$ is achieved by $\mathbf{s}_1$ and $\mathbf{s}_2$, respectively (the proof can be easily extended to the case when the infimum is not achieved). This means that there exists $a, b, c, d \geq 0$, such that

$$\partial_{\text{drop}}(\mathbf{x}, \mathbf{s}_1) = a;\ \partial_{\text{move}}(\mathbf{s}_1, \mathbf{y}) = b;\ \partial_{\text{drop}}(\mathbf{y}, \mathbf{s}_2) = c;\ \partial_{\text{move}}(\mathbf{s}_2, \mathbf{z}) = d,$$

which implies $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{y}) = a + \eta b$ and $\partial^{\eta}_{\text{drmv}}(\mathbf{y}, \mathbf{z}) = c + \eta d$. We need to show that $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{z}) \leq (a+c) + \eta(b+d)$.

Using Lemma 12 with $\partial_{\text{move}}(\mathbf{s}_1, \mathbf{y}) = b$ and $\partial_{\text{drop}}(\mathbf{y}, \mathbf{s}_2) = c$, we get that there is a $\mathbf{y}'$ such that $\partial_{\text{drop}}(\mathbf{s}_1, \mathbf{y}') = c$ and $\partial_{\text{move}}(\mathbf{y}', s_2) \leq b$. This gives the following:

$$\partial_{\text{drop}}(\mathbf{x}, \mathbf{s}_1) = a;\ \partial_{\text{drop}}(\mathbf{s}_1, \mathbf{y}') = c;\ \partial_{\text{move}}(\mathbf{y}', \mathbf{s}_2) \leq b;\ \partial_{\text{move}}(\mathbf{s}_2, \mathbf{z}) = d.$$

Now we prove that $\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{z}) \leq (a + c) + \eta(b + d)$:

$$\begin{aligned}\partial^{\eta}_{\text{drmv}}(\mathbf{x}, \mathbf{z}) &= \inf_{\mathbf{z}} \left(\partial_{\text{drop}}(\mathbf{x}, \mathbf{y}) + \eta \cdot \partial_{\text{move}}(\mathbf{y}, \mathbf{z})\right) \\ &\leq \partial_{\text{drop}}(\mathbf{x}, \mathbf{y}') + \eta \cdot \partial_{\text{move}}(\mathbf{y}', \mathbf{z}) \\ &\leq \partial_{\text{drop}}(\mathbf{x}, \mathbf{s}_1) + \partial_{\text{drop}}(\mathbf{s}_1, \mathbf{y}') + \eta \cdot \partial_{\text{move}}(\mathbf{y}', \mathbf{z}) && (\partial_{\text{drop}} \text{ is a quasi-metric}) \\ &\leq \partial_{\text{drop}}(\mathbf{x}, \mathbf{s}_1) + \partial_{\text{drop}}(\mathbf{s}_1, \mathbf{y}') + \eta \cdot (\partial_{\text{move}}(\mathbf{y}', \mathbf{s}_2) + \partial_{\text{move}}(\mathbf{s}_2, \mathbf{z})) && (\partial_{\text{move}} \text{ is a metric}) \\ &\leq (a + c) + \eta(b + d).\end{aligned}$$

This concludes the proof of Claim 12 □

J d-Dimensional Analogues of our Mechanisms/Results

Algorithm 4 Bucketing Mechanism over $[0, B)^d$, $\mathcal{M}^{w,[0,B)^d}_{\text{buc}}$

Parameter: Bucket (which is d-dimensional cube) side length w, ground set $[0, B)^d$.
Input: A histogram $\mathbf{x}$ over $[0, B)$.
Output: A histogram $\mathbf{y}$ over $S = T^d$ where $T = \{w(i - \frac{1}{2}) : i \in [t], t = \lceil \frac{B}{w} \rceil\}$, and $|\mathbf{y}| = |\mathbf{x}|$.

1: **for all** $s \in S$ **do**
2: $\quad \mathbf{y}(s) := \sum_{g : g - s \in [\frac{-w}{2}, \frac{w}{2})^d} \mathbf{x}(g)$
3: **end for**
4: Return $\mathbf{y}$

In our d-dimensional bucketing mechanism for $\mathcal{G} = [0, B)^d$, we divide $[0, B)^d$ into $t = \lceil \frac{B}{w} \rceil^d$ d-dimensional cubes (buckets), each of side length w, and map each input point to the center of the nearest cube (bucket). Note that

Algorithm 5 BucketHist Mechanism over $[0,B)^d$, $\mathcal{M}^{\alpha,\beta,[0,B)^d}_{\text{BucHist}}$

Parameter: Accuracy parameters α, β; ground set $[0,B)^d$.
Input: A histogram $\mathbf{x}$ over $[0,B)^d$.
Output: A histogram $\mathbf{y}$ over $[0,B)^d$.

1: $w := 2\beta$, $t := \lceil \frac{B}{w} \rceil^d$, $\tau := \alpha/t$
2: Return $\mathcal{M}^{\tau,\epsilon,[0,B)^d}_{\text{STLap}} \circ \mathcal{M}^{w,[0,B)^d}_{\text{buc}}(\mathbf{x})$

the distance between any point in $[0,B)^d$ to the center of the nearest bucket is $\frac{w}{2}\sqrt{d}$. In the following, we will ignore the ceil/floor for simplicity.

Our d-dimensional bucketing mechanism $\mathcal{M}^{w,[0,B)^d}_{\text{buc}}$ and the final d-dimensional bucketed-histogram mechanism $\mathcal{M}^{\alpha,\beta,[0,B)^d}_{\text{BucHist}}$ are presented in Algorithm 4 and Algorithm 5, respectively.

As mentioned in Remark 5 in Appendix I, with these modified mechanisms, all our results in Theorem 4, Theorem 5, and Theorem 6 will hold verbatim, except for the value of τ, which will be replaced by $\tau = \alpha(\frac{2\beta}{B\sqrt{d}})^d$. Note that for the one dimensional case, we have $\tau = \frac{\alpha}{t} = \alpha(\frac{w}{B})$, where $w = 2\beta$, which comes from the $(0, \frac{w}{2}, 0)$-accuracy of the bucketing mechanism $\mathcal{M}^{w,[0,B)}_{\text{buc}}$ (see Claim 6 in Appendix G). The d-dimensional analogue of that result is stated in the following claim which can be proven along the lines of Claim 6.

Claim 13. *$\mathcal{M}^{w,[0,B)^d}_{\text{buc}}$ is $\left(0, \frac{w}{2}\sqrt{d}, 0\right)$-accurate for the identity function f_{id} over $\mathcal{H}_{[0,B)^d}$ w.r.t the metric $\mathfrak{d}_{\text{hist}}$.*

It follows from Claim 13 that the output error of $\mathcal{M}^{w,[0,B)^d}_{\text{buc}}$ is $\beta = \frac{w}{2}\sqrt{d}$. This implies $\tau = \frac{\alpha}{t} = \alpha(\frac{w}{B})^d = \alpha(\frac{2\beta}{B\sqrt{d}})^d$.

K Details Omitted from Section 3

In this section, we prove that $\widehat{\partial}$ is a quasi-metric (assuming that ∂ is a quasi-metric).

Lemma 13. *If ∂ is a quasi-metric, then $\widehat{\partial}$ is a quasi-metric.*

Proof. We need to show that for any three distributions P, Q, and R over the same space A, we have (i) $\widehat{\partial}(P,Q) \geq 0$, where the equality holds if and only if $P = Q$, and (ii) $\widehat{\partial}$ satisfies the triangle inequality: $\widehat{\partial}(P,Q) \leq \widehat{\partial}(P,R) + \widehat{\partial}(R,Q)$. We show them one by one below:

1. The first property follows from the definition of $\widehat{\partial}$ (see Definition 2): If $\widehat{\partial}(P,Q) = 0$, then the optimal $\phi \in \Phi(P,Q)$ is a diagonal distribution, which means that $P = Q$. On the other hand, if $P = Q$, then there exists a coupling ϕ in $\Phi(P,Q)$, which is a diagonal distribution and hence $\widehat{\partial}(P,Q) = 0$.
2. Since the definition of $\widehat{\partial}$ is the same as that of W^∞, except for that the former is defined w.r.t. a quasi-metric, whereas, the latter is defined w.r.t. a metric, we can show the triangle inequality for $\widehat{\partial}$ along the lines of the proof of Lemma 3. Note that we did not use the symmetric property of W^∞ while proving Lemma 3; we only used that the underlying metric $\mathfrak{d}$ satisfies the triangle inequality, which also holds for $\widehat{\partial}$ which is a quasi-metric.

This completes the proof of Lemma 13. □

L Details Omitted from Section 6 and Additional Experimental Evaluations

We empirically compare our basic mechanism $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ (Algorithm 1) on a ground set $\mathcal{G} = \{1, \cdots, B\}$, against various competing mechanisms, for accuracy on a few histogram-based statistics computed on it. We plot average errors (actual and flexible), on different histogram distributions[8] for functions $\max_k(\mathbf{x}) := \max\{i \mid \mathbf{x}(i) \geq k\}$,

[8] For each data distribution, the plots were averaged over 100 data sets, with 100 runs each for each mechanism.

$\max := \max_1$, and $\text{mode}(\mathbf{x}) := \arg\max_i \mathbf{x}(i)$; note that $\text{mode}(\mathbf{x})$ is equal to the most frequently occurring data item in $\mathbf{x}$. The parameters for $\mathcal{M}^{\tau,\epsilon,\mathcal{G}}_{\text{STLap}}$ that we will use in the section are given in Corollary 1.

We emphasize that the plots are only indicative of the performance of our algorithm on specific histograms, and do not suggest *worst-case accuracy guarantees*. On the other hand, our theorems do provide worst-case accuracy guarantees.

We will empirically compare our results against the Exponential Mechanism [McSherry and Talwar, 2007], Propose-Test-Release Mechanism [Dwork and Lei, 2009], Smooth-sensitivity Mechanism [Nissim et al., 2007], Stability-Based Sanitized Histogram [Bun et al., 2019], and Choosing-Based Histogram Mechanism [Beimel et al., 2016]. We describe these mechanisms briefly in Appendix L.2. We point out one notable omission from our plots: the Encode-Shuffle-Analyze histogram mechanism [Erlingsson et al., 2020], which appeared independently and concurrently to our mechanism,[9] also uses a shifted (but not truncated) Laplace mechanism, and in all the examples plotted, yields a behavior that is virtually identical to our mechanism's. However, we emphasize that Erlingsson et al. [2020] claim accuracy only for the histogram itself, and indeed, for the functions that we consider, it does not enjoy the *worst-case accuracy guarantees* that we provide.

L.1 Evaluations Carried Out

In each of the following empirical evaluations, a histogram distribution and one of the following functions were fixed: $\max_k(\mathbf{x}) := \max\{i \mid \mathbf{x}(i) \geq k\}$, $\max := \max_1$, and $\text{mode}(\mathbf{x}) := \arg\max_i \mathbf{x}(i)$.

(1) Function max. Histogram of about 10,000 items drawn i.i.d. from a Cauchy distribution with median 45 and scale 4, restricted to 100 bars, with the last 10 set to empty bars.
(2) Function max. Step histogram with two steps (height $\times$ width) : $[1000 \times 50,\ 1 \times 50]$.
(3) Function $\max_{500}$. Same histogram distribution as in (1) above, but without zeroing out the right-most bars.
(4) Function $\max_{500}$. Step histogram with 100 bars, with two steps (height $\times$ width) : $[540 \times 50,\ 490 \times 50]$.
(5) Function mode. Histogram of 30 bars, each bar has height drawn from i.i.d. Poisson with mean 250.
(6) Function mode. Noisy step histogram, with steps $[130 \times 120,\ 200 \times 5,\ 185 \times 85,\ 190 \times 10,\ 130 \times 80]$.

The results are shown in Figure 4. In each experiment, a range of values for ϵ are chosen, while we fixed $\delta = 2^{-20}$. Errors are shown in the y-axis as a percentage of the full range $[0, B)$. For each experiment and mechanism, we also compute flexible accuracy allowing a distortion of $\partial_{\text{drop}} = 0.005$.

L.2 Description of the Compared Mechanisms

Exponential Mechanism. The Exponential Mechanism [McSherry and Talwar, 2007] can be tailored for an abstract utility function. We consider the negative of the error as the utility of a response y on input histogram $\mathbf{x}$, i.e., $q(\mathbf{x}, y) = -\text{err}(\mathbf{x}, y) = -|\max(\mathbf{x}) - y|$. However, for both $\max_k$ and mode, error has high sensitivity – changing a single element in the histogram can change the error by as much as the number of bars in the histogram. Since the mechanism produces an output r with probability proportional to $e^{\frac{\epsilon q(\mathbf{x},r)}{2\Delta_{\text{err}}}}$, where Δ_{err} is the sensitivity of $\text{err}(\cdot,\cdot)$, having a large sensitivity has the effect of moving the output distribution close to a uniform distribution. This is reflected in the performance of this mechanism in all our plots.

Propose-Test-Release Mechanism (PTR). We consider the commonly used form of the PTR mechanism of Dwork and Lei [Dwork and Lei, 2009], namely, "releasing stable values" (see [Vadhan, 2017, Section 3.3]). On input $\mathbf{x}$, the mechanism either releases the correct result $f(\mathbf{x})$ or refuses to do so (replacing it with a random output value), depending on whether the radius of the neighborhood of $\mathbf{x}$ where it remains constant is sufficiently large (after adding some noise). For computing a function f and a setting of parameter $\beta = 0$ and privacy parameters ϵ, δ, the mechanism calculates this radius for an input $\mathbf{x}$ as, $r = d(\mathbf{x}, \{\mathbf{x}' : \text{LS}_f(\mathbf{x}') > 0\}) + \text{Lap}(1/\epsilon)$, where $d(\mathbf{x}, \mathcal{S})$ is the minimum Hamming distance between $\mathbf{x}$ and any point in the set $\mathcal{S}$ and $\text{LS}_f(\mathbf{y}) := \max\{|f(\mathbf{y}) - f(\mathbf{z})| : \mathbf{y} \sim \mathbf{z}\}$ is local sensitivity of function f at $\mathbf{y}$. If this radius r is greater than $\ln(1/\delta)/\epsilon$, the mechanism will output the exact answer $f(\mathbf{x})$, otherwise it outputs a random value from the domain. For the functions we consider, this radius

[9] A preliminary version of the current work was presented as an invited talk at a pre-pandemic conference, in December 2019 (citation omitted for anonymity).

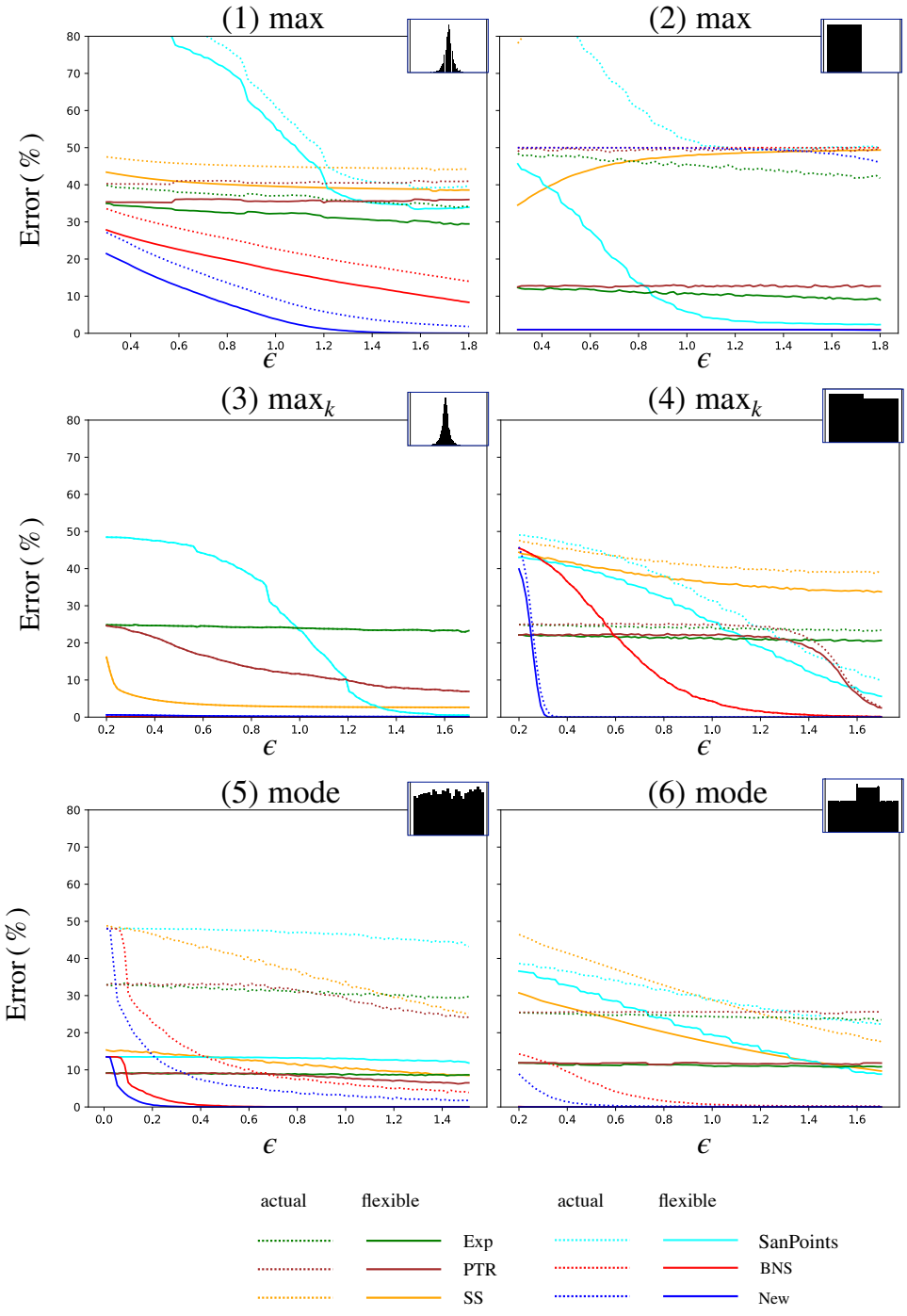


Figure 4: For each evaluation, a typical histogram used is shown in inset. The different data distributions elicit a variety of behaviors of the different mechanisms. Experiment (2) shows an instance which is hard for all the mechanisms without considering flexible accuracy; on the other hand, in Experiment (3), flexible accuracy makes no difference (the plots overlap). In these two experiments, BNS and our new mechanism match each other. In all the other experiments, our new mechanism dominates the others, with or without considering flexible accuracy.

of stable region can be computed efficiently and is typically small or even zero for input distributions considered which is reflected in our plots.

Smooth-sensitivity Mechanism (SS). This mechanism, due to Nissim et al. [Nissim et al., 2007], uses the smooth sensitivity of query f as $SS^{\epsilon}_{f}(\mathbf{x}) = \max\{LS_f(\mathbf{x}')e^{-\epsilon d(\mathbf{x},\mathbf{x}')}|\mathbf{x}' \in \mathcal{H}_{\mathcal{G}}\}$, where $LS_f(\mathbf{x}')$ denotes the *local sensitivity* of f at $\mathbf{x}'$, and $d(\cdot,\cdot)$ is the Hamming distance. Given an input histogram $\mathbf{x}$, the mechanism adds noise roughly $O(SS^{\epsilon}_{f}(\mathbf{x})/\epsilon)$ to $f(x)$ for (ϵ,δ)-privacy. For functions like $\max_k$ and mode, again, the local sensitivity tends to be large on many histograms, and this affects the performance of this mechanism on such inputs.

Stability-Based Sanitized Histogram Mechanism. This mechanism was proposed by Bun et al. [Bun et al., 2019] (also see [Vadhan, 2017, Theorem 3.5]) for releasing histograms with provable worst-case guarantees. However, these guarantees are in terms of the errors in the individual bar heights of the histogram, and does not necessarily translate to the histogram based functions, as we consider. Nevertheless, this mechanism provides a potential candidate for a mechanism for any histogram based statistic.

For each bar of the histogram, the mechanism adds Laplace noise to the bar height, and the resulting value is reported only if it is more than a threshold, and otherwise a 0 is reported. By treating empty bars differently, this mechanism achieves comparable flexible accuracy as our mechanism in the case of max. However, this does not generalize to $\max_k$. In particular, in the example in (4) in Figure 4, by adding (possibly positive) noise to histogram bars of height lower than k, the mechanism is very likely to find a bar which is much further to the right than the point where the bar heights cross k.

Choosing-Based Histogram Mechanism. Beimel et al. [Beimel et al., 2016] presented a mechanism SanPoints for producing a sanitized histogram, with formal (α,β) PAC-guarantees for approximating the height of each bar of the histogram. For a given α,β and privacy paramters ϵ,δ, on an input $\mathbf{x}$ with m elements, the mechanism adds little noise, $\text{Lap}(1/m\epsilon)$ to heights of bars which are iteratively chosen without repetition as per a choosing mechanism. The choosing mechanism privately picks the index with maximum bar height. It either chooses the max height index if its bar height is sufficiently large after adding some Laplace noise or otherwise chooses as per an exponential mechanism with bar height as an index's score.

For a given input $\mathbf{x}$ with m elements and parameters β,ϵ,δ, the mechanism guarantees $\alpha = O\big(\big(\frac{\sqrt{\ln(1/\delta)}\ln(\ln(1/\delta))}{\epsilon m}\big)^{2/3}\big)$. Again though the mechanism doesn't give formal guarantees for the functions we consider, such a histogram-release mechanism can be heuristically used for any histogram based statistic. In our evaluations, SanPoints yields mixed results, but is dominated by BNS and our new mechanism.