The Podium Mechanism: Improving on the Laplace and Staircase Mechanisms

Vasyl Pihur

Abstract—The Podium mechanism guarantees \((\epsilon, 0)\)-differential privacy by sampling noise from a finite mixture of three uniform distributions. By carefully constructing such a mixture distribution, we trivially guarantee privacy properties, while minimizing the variance of the noise added to our continuous output. Our gains in variance control are due to the “truncated” nature of the Podium mechanism where support for the noise distribution is maintained as close as possible to the sensitivity of our data collection, unlike the infinite support that characterizes both the Laplace and Staircase mechanisms. In a high-privacy regime \((\epsilon < 1)\), the Podium mechanism outperforms the other two by 50-70\% in terms of the noise variance reduction, while in a low privacy regime \((\epsilon \rightarrow \infty)\), it asymptotically approaches the Staircase mechanism.

Index Terms—Data Privacy, randomized algorithm.

I. INTRODUCTION

Since the introduction of differential privacy \([1]\), the Laplace mechanism \([1]\) became de facto a standard way of ensuring that the differential privacy property is satisfied when collecting or releasing continuous outcomes. In fact, to many casual privacy practitioners, the notions of differential privacy and the Laplace mechanism have become two sides of the same coin, one a theoretical and somewhat vague concept, while the second a practical and prescriptive way of achieving it.

The scope of this work is most relevant to the local privacy model where multiple noise additions accumulate through the collection process, so we limit our discussion to this setting. However, the Podium mechanism can be used in place of the Laplace or the Staircase \([2]\) mechanisms in all settings and will result in more precise estimates (smaller variance) in all cases.

It is also important to point out that we are dealing solely with continuous or numeric responses, so randomized algorithms designed for the discrete case are not strictly relevant \([3]\), though one can always discretize a continuous response.

Formally, let \(X\) be a continuous random variable to be collected in a local privacy model \([4]\, \[5]\) given a privacy budget of \(\epsilon\). Let \(M\) be a randomized mechanism \([6]\) that adds zero-mean noise with variance \(\sigma^2\) to each raw data point \(x\). Let \(x' = M(x)\) be the observed, noisy data points satisfying the \(\epsilon\)-differential privacy property.

Definition 1: \((\epsilon\)-differential privacy). A randomized mechanism \(M\) satisfies \(\epsilon\)-differential privacy if for all inputs \(x_i\) and \(x_j\) and all outputs \(x'\),

\[
P(M(x_i) = x') \leq e^\epsilon P(M(x_j) = x').
\]

V. Pihur is with Snap Inc. e-mail: vpihur@snapchat.com.

Definition 2: (Global Sensitivity). Sensitivity of the data collection, \(\Delta\), is defined \(a\ priori\) to the data collection as

\[
\Delta = \max_{x_i, x_j} ||x_i - x_j||_1.
\]

For a one-dimensional case, this is simply a true range \(R(X)\) (or support) of the random variable \(X\), defined as

\[
R(X) = \max(X) - \min(X).
\]

Definition 3: (Relative Efficiency). The relative efficiency \([7]\) of two estimators of an unknown parameter \(\mu\), \(T_1\) and \(T_2\) is defined as

\[
e(T_1, T_2) = \frac{V(T_1)}{V(T_2)},
\]

where \(V(\cdot)\) is the variance of the estimator defined as

\[
V(T_1) = E[(T_1 - \mu)^2],
\]

where \(E(\cdot)\) is the expectation operator.

This is a standard way of comparing two estimators in the statistics literature and, whenever presented with two unbiased estimators for a quantity \(\mu\), one would obviously prefer one with smaller variance.

Definition 4: (Laplace Mechanism). The Laplace Mechanism (LM) ensures \(\epsilon\)-differential privacy by adding noise from a Laplace distribution with mean 0 and scale \(b = \frac{\Delta}{\epsilon}\) such that

\[
x' = x + L \left( 0, \frac{\Delta}{\epsilon} \right) = L \left( x, \frac{\Delta}{\epsilon} \right),
\]

where \(L\) is a random variable with a probability density function

\[
f(z; \mu, b) = \frac{1}{2b} e^{-\frac{|z - \mu|}{b}}, \forall z \in \mathbb{R}.
\]

The Laplace Mechanism is being widely used in practice, mainly due to the ease with which a specific Laplace distribution can be chosen given sensitivity \(\Delta\), as well as actual generation of random variables from that distribution by taking the natural log of a scaled uniform random variable with a random sign.

The same cannot be said about the Staircase mechanism discussed below which is likely the main reason why it has not completely replaced the Laplace mechanism as of the time of this work.

Definition 5: (Staircase Mechanism). To ensure \(\epsilon\)-differential privacy, the Staircase mechanism (SM) \([2]\, \[8]\) samples from a geometric mixture of uniform random
variables with the probability density function $f(z; \gamma)$ defined as

$$f(z; \gamma) = \begin{cases} a(\gamma) & z \in [0, \gamma \Delta) \\ e^{-\gamma}a(\gamma) & z \in [\gamma \Delta, \Delta) \\ e^{-k\epsilon}f_{\gamma}(z-k\Delta) & z \in [k\Delta, (k+1)\Delta) \\ f_{\gamma}(-z) & z < 0 \end{cases}$$

where

$$a(\gamma) = \frac{1 - e^{-\epsilon}}{2\Delta(\gamma + e^{-\epsilon}(1 - \gamma))}$$

and

$$\gamma = -\frac{e^{-\epsilon}}{1 - e^{-\epsilon}} + \frac{(e^{-\epsilon} - 2e^{-2\epsilon} + 2e^{-4\epsilon} - e^{-5\epsilon})^{1/3}}{21^{1/3}(1 - e^{-\epsilon})^2}.$$
us to match the range of input and output values. Ideally, if we are collecting, say age, spread from 13 to 120 (Δ = 107), we would like to output the corresponding noisy values in the same range. Note that both the Laplace and Staircase mechanism generate values on the whole real line, leading to the loss of efficiency due to the possibility of extreme outliers. However, m cannot possibly be 1 because adding noise centered at the extreme values of either −Δ/2 or Δ/2 would require mean to be equal to min or max, which is not possible for any non-degenerate distribution. Therefore, m must be larger than 1, extending the support of the noise distribution in both directions, but should be as small as possible for efficiency reasons. This parameter depends only on ϵ and is determined by minimizing the variance of the Podium distribution.

The second parameter w describes the width of the step. Its value also comes from the variance optimization. It depends on ϵ and Δ and can be pre-computed once.

The third parameter t describes the location of the step under the constraint that the mean (µ) of the Podium distribution is equal the input value x. This parameter ranges between −Δm/2 and Δm/2. Since it changes depending on x, it must be computed every time during the collection process. Because we do not want to perform a constrained optimization, we parameterize t using another unconstrained parameter s as

\[ t = \Delta m \frac{1}{1 + e^{-s}} - \frac{\Delta m}{2}, \]

which translates a real value s into an interval [−Δm/2, Δm/2].

B. Deriving the Podium distribution

To derive the shape of the Podium distribution, we would like to minimize its variance. We are presented with a choice here as the shape and variance of the distribution changes, depending on its mean. It makes sense to perform such minimization under the constraint that its mean is equal to Δ/2 or at its most extreme shape. It is there that we are forced to allocate a margin m to balance the distribution. It is also a shape where the second parameter w becomes a function of t,
as the distribution becomes a mixture of two uniform variables instead of three. This distribution is shown schematically in Figure [3].

We will perform variance optimization calculations at the extreme right shape of the Podium distribution. We have two unknowns \( m \) and \( s \) and two constraints. The first one is that \( \mu = \Delta/2 \) and the second one is that the area under the Podium function should add up to one to be a proper distribution.

Because it is a two-component mixture distribution with mean \( \mu \), its variance can be computed \([11]\) as

\[
V(Z) = E[(Z - \mu)^2] = \sum_{i=1}^{K} p_i (\mu_i^2 + \sigma_i^2) - \mu^2 = p(\mu_1^2 + \sigma_1^2) + (1 - p)(\mu_2^2 + \sigma_2^2) - (\Delta/2)^2,
\]

where \( p \) is the proportion of the first component, \( \mu_1 \) and \( \mu_2 \) are the means of each component and \( \sigma_1^2 \) and \( \sigma_2^2 \) are their corresponding variances.

First, we compute the probability of the first component \( p \) which turns out to be a function of \( \epsilon \) and \( s \) only. Let \( d \) be the density (height) of the first component. To guarantee \( \epsilon \)-differential privacy, the second component must be equal to \( de^\epsilon \) (refer to Figure [3]).

Then,

\[
p = \left( \frac{\Delta m}{1 + e^{-s}} - \frac{\Delta m}{2} + \frac{\Delta m}{2} \right) d = \frac{\Delta md}{1 + e^{-s}}.
\]

Similarly after some algebra,

\[
1 - p = \frac{\Delta mde^\epsilon}{1 + e^{s}}.
\]

Because these must add up to 1 to produce a proper density function, we can solve for \( d \) which is equal to

\[
d = \frac{(1 + e^{-s})(1 + e^{s})}{\Delta m(1 + e^{s} + e^{\epsilon} + e^{\epsilon-s})}.
\]

Plugging \( d \) into the first component probability gives

\[
p = \frac{1 + e^{s}}{1 + e^{s} + e^{\epsilon} + e^{\epsilon-s}},
\]

which does not depend on \( m \) or \( \Delta \).

Since each component is simply a uniform random variable on an interval \([a, b]\), its mean is given by \( \frac{a+b}{2} \) and variance by \( \frac{(b-a)^2}{12} \). Thus, the mean of the first component is given by

\[
\mu_1 = -\frac{\Delta m}{2} \left( \frac{1}{1 + e^{s}} \right)
\]

and the mean of the second component by

\[
\mu_2 = \frac{\Delta m}{2} \left( \frac{1}{1 + e^{-s}} \right).
\]

Their variances, of course, are simply

\[
\sigma_1^2 = \frac{\Delta^2 m^2}{12} \left( \frac{1}{1 + e^{-s}} \right)^2
\]

and

\[
\sigma_2^2 = \frac{\Delta^2 m^2}{12} \left( \frac{1}{1 + e^{s}} \right)^2,
\]

respectively.

We now consider our second constraint that the mean of this distribution is equal to \( \Delta/2 \) which implies that

\[
\mu = p\mu_1 + (1 - p)\mu_2
\]

\[
= -\frac{\Delta m}{2} \left( \frac{1}{1 + e^{s}} + \frac{\Delta me^{\epsilon}}{2(1 + e^{s} + e^{\epsilon} + e^{\epsilon-s})} \right)
\]

\[
= \frac{\Delta}{2}.
\]

This allows us to solve for \( m \) which can be expressed as

\[
m = \frac{1 + e^{s} + e^{\epsilon} + e^{\epsilon-s}}{e^{\epsilon} - 1}.
\]

At this point, everything is expressed in terms of \( \epsilon \), \( \Delta \) and \( s \). Plugging individual pieces into the total variance formula above, after combining and rearranging terms, we get

\[
V(Z) = \frac{\Delta^2}{12} \left( 1 + e^{s} + e^{\epsilon} + e^{\epsilon-s}(3 + 6e^{-s} + e^{s}(e^{s} + 3e^{\epsilon})) \right)
\]

\[
- \frac{\Delta^2}{4} (e^{s} - 1)^2(1 + e^{s})
\]

Taking the first derivative of \( V(Z) \) with respect to \( s \) gives

\[
\frac{dV(Z)}{ds} = -2e^{-s} + 2e^{s+\epsilon} - e^{2s-2s} + e^{2s},
\]

which is a quartic function (4\textsuperscript{th}-degree polynomial) in \( s \). Setting \( \frac{dV(Z)}{ds} \) equal to 0 and solving for \( s \), we get

\[
s = \begin{cases} 
\log \left( \frac{-\sqrt{A + e^{s} + e^{\epsilon} + e^{\epsilon}} \pm \sqrt{B + 2e^{s} - A}}{2} \right) & \epsilon \geq \log(\sqrt{2}) \\
\log \left( \frac{-\sqrt{A + e^{s} + e^{\epsilon} + e^{\epsilon}} \pm \sqrt{B + 2e^{s} - A}}{2} \right) & \epsilon < \log(\sqrt{2})
\end{cases}
\]

(1)
where
\[ A = (4(e^{2s} - e^{4s}))^{1/3} \]
and
\[ B = \frac{2(2e^s - e^{3s})}{\sqrt{A + e^{2s}}} \]

The second derivative is given by
\[ \frac{d^2V(Z)}{ds^2} = 4e^s(\cosh(2s - \epsilon) + \cosh(s)) \]
and is always positive as the domain of \( \cosh(x) \) is \( \geq 1 \). Thus, our solution represents a true global minimum.

These expressions for \( s \) look daunting at first, as are all real solutions to quartic equations. We plotted \( \epsilon \) vs \( s \) in Figure 4 and it is apparent that their relationship can be quite closely approximated by a linear function! In fact, \( s = \epsilon/3 \) is a very good approximation for the above equations. It is important to keep in mind that this approximation for \( s \) does not effect the privacy of the Podium mechanism. It only affects its relative efficiency and, as we will demonstrate later, not by much at all.

Our width parameter \( w \) can now be computed as
\[ w = \frac{\Delta m}{2} - t \]
\[ = \frac{\Delta m}{2} - \left( \frac{\Delta m}{1 + e^{x}} - \frac{\Delta m}{2} \right) \]
\[ = \Delta m \left( 1 - \frac{1}{1 + e^{x}} \right) \]
\[ = \frac{\Delta m}{1 + e^{x}}. \]

Both \( m \) and \( s \) are agnostic to the sensitivity \( \Delta \), while \( w \) is linearly proportional to it.

### III. Sampling from the Podium distribution

To generate a random variable from the Podium distribution given the collection parameters \( \epsilon \) and \( \Delta \), one can pre-compute \( m \), \( w \), and \( d \) using Algorithm 1. This can be done once prior to the start of the collection process. One has two choices with respect to this step, depending on how “optimal” one would like to be, either EXACT or APPROXIMATE. If one chooses the EXACT value for \( s \), then one must use the quartic solutions (they are pre-computed for a range of \( \epsilon \) values in Table II along with the rest of the parameters. If one is comfortable with a slight loss in efficiency, then setting \( s = \epsilon/3 \) and computing \( m \), \( w \) and \( d \) as in Algorithm 1 is necessary.

**Algorithm 1** Offline algorithm for computing the extreme right form of the Podium distribution. One has a choice of computing \( s \) either exactly or approximately. This step needs to be performed only once prior to the collection process. If the desired value of \( \epsilon \) can be found in Table I then one can find the output of this algorithm there.

1: **Input:** \( \epsilon \), \( \Delta \)
2: **if** EXACT **then**
3:   Compute \( s \) according to (1) or look it up in Table I
4: **else**
5:   Compute \( s = \epsilon/3 \)
6: **end if**
7: Compute \( m = \frac{1+e^s+e^s+e^{-s}}{1+e^s} \)
8: Compute \( w = \frac{\Delta m}{1+e^s} \)
9: Compute \( d = \frac{\Delta m(1+e^s+e^s+e^{-s})}{(1+e^s)^2} \)
10: **Output:** \( m \), \( w \) and \( d \)

To add the Podium noise during the actual collection, one must perform Algorithm 2 on every noise addition, since the shape of the distribution depends on the input value \( x \). The only shape parameter that changes is \( t \), the location of the step. After computing \( t \), we simply pick at random one of the three mixture components (by generating a standard uniform random variable) and then randomly pick from the selected component with the help of another uniform random variable.

In Table II we pre-computed \( d, w, m \) and \( s \) for a wide variety of \( \epsilon \)’s. This table is meant to be used as a lookup table for the shape of the distribution by practitioners in cases when they do not want to bother with the messy quartic solutions. We range \( \epsilon \) from 0.1 to 10 and show \( d, w/\Delta, m \) with great precision of up to 20 digits.

It takes two uniform random variables to generate one from the Podium distribution with the additional burden of computing the location of the step. In that sense it is similar to the Laplace mechanism which also requires generating two uniform random variables. The Staircase mechanism requires generation of three uniform and one Geometric random variables and, therefore, is a bit more expensive from the computational point of view.

### IV. Privacy of the Podium Mechanism

**Theorem 1:** (Privacy of the Podium mechanism.) The Podium mechanism \( P \) provides \( \epsilon \)-differential privacy, satis-
The truncated nature of the Podium mechanism carries with it a promise of significant reductions in noise variance, as well as a reasonable matching of input and output ranges. In this section, we will study the efficiency implications of
The Podium mechanism has smaller variance in all privacy regimes and outperforms the other two in the most critical high to medium privacy scenarios (small $\epsilon$).

Algorithm 2 Online algorithm for generating a random variable from the Podium distribution with mean $\mu = x$. This algorithm needs to be performed on every noise addition. It requires generation of two uniform random variables.

1: Input: $\epsilon$, $\Delta$, $m$, $w$, $d$ and $x$
2: Compute $t = \frac{2x-w^2d(e^\epsilon-1)}{2wd(e^\epsilon-1)}$
3: Compute probability of first component $p_1 = d(t + \frac{\Delta m}{2})$
4: Compute probability of second component $p_2 = de^\epsilon w$
5: Generate uniform random variable $Y$ in $[0,1]$
6: if $Y < p_1$ then
7: return a uniform random variable $X_1'$ in $[-\frac{\Delta m}{2}, t)$
8: else if $Y < p_1 + p_2$ then
9: return a uniform random variable $X_2'$ in $[t, t + w)$
10: else
11: return a uniform random variable $X_3'$ in $[t + w, \frac{\Delta m}{2}]$
12: end if

The variance of the Staircase mechanism is given by

$$V_{SM} = \frac{\Delta^2(2^{-2/3}e^{-2\epsilon/3}(1 + e^{-\epsilon})^{2/3} + e^{-\epsilon})}{(1 - e^{-\epsilon})^2}.$$ 

The variance considerations of the Podium mechanism are a little more involved, as its variance varies depending on its mean $\mu$. It has the smallest variance when $\mu = 0$ (the Podium distribution is symmetric) and this variance is given by

$$V_{PM}^{\mu=0} = \frac{d}{12} \left( \Delta^2 m^2 + w^3 (e^\epsilon - 1) \right)$$

$$\Delta^2 m^2 \left( 1 + e^{-s} \right) \left( 1 + e^s \right)$$

$$\Delta^2 \left( 1 + e^s + e^s + e^{-s} \right) \left( 1 + e^{-s} \right) \left( 1 + e^s \right)$$

$$\left( e^\epsilon - 1 \right)^2.$$ 

The Podium mechanism takes on the largest variance in case of the most off-center location of the step, i.e., when a large portion of the mass is in one of the tails (its mean is equal to $-\frac{\Delta}{2}$ or $\frac{\Delta}{2}$). Here, the variance is equal to

$$V_{PM}^{\mu=\Delta/2} = \frac{\Delta^2 m^2}{12} \times \frac{1}{1 + e^s + e^s + e^{-s}}$$

$$\times 3 + e^{-s} + e^s (e^s + 3e^s) - \frac{\Delta^2}{4}$$

$$\Delta^2 \cosh(2s - \epsilon) + 4 \cosh(s) + 3$$

$$\cosh(\epsilon) - 1.$$
Its variance is somewhere in between these two values for any intermediate shape of the distribution. In Figure 5 we compare variances of the three mechanisms for four different values of $\epsilon$. We also consider the exact and approximate versions of the Podium mechanism. The following observations are worth noting:

- The Podium mechanism is more efficient than the other two in all privacy regimes. It is especially pronounced for the commonly used $\epsilon = \log(3)$ where its variance is essentially halved.
- The two versions of the Podium mechanism are virtually indistinguishable in terms of efficiency.
- The variance of the Podium mechanism is smallest at 0 and is monotonically increasing towards the extremes.
- The Laplace and Staircase mechanisms are asymptotically equivalent in the high-privacy regime ($\epsilon \to 0$), the Podium and Staircase mechanisms are asymptotically equivalent in the low-privacy regime ($\epsilon \to \infty$).

To study the optimality implications of the Podium mechanism, we will make use of the fact that $s = \epsilon/3$ is a good approximation for the optimal $s$ to make this a more tractable exercise.

In the high privacy regime ($\epsilon \to 0$), the step width $w$ is equal to $\Delta m$, i.e., the Podium distribution becomes equivalent to the uniform distribution on the interval $[-\Delta m, \Delta m]$. Thus, its variance (it is also apparent by plugging $\epsilon = 0$ into $V_{PM}^{\mu = 0}$) is $V_{PM}^{\epsilon = 0} = \frac{\Delta^2 m^2}{12}$, which is exactly what we would want for perfect privacy.

**Theorem 2:** (High Privacy Regime). In the high privacy regime ($\epsilon \to 0$), the variance of the Podium mechanism is equal to

$$V_{PM}^{\epsilon = 0} = \Theta \left( \frac{4 \Delta^2}{3 \epsilon^2} \right).$$

**Proof.** Since

$$m_{\epsilon \to 0} = \frac{1 + e^x + e^t + e^{-x} - s}{e^\epsilon - 1} = \Theta \left( \frac{4}{\epsilon} \right),$$

the result immediately follows.

The Podium mechanism is $33\%$ more efficient at the extreme high privacy regime relative to either the Laplace or Staircase mechanisms

$$\epsilon_{\epsilon \to 0}(PM, LM) = \frac{V_{PM}}{V_{LM}} = \Theta \left( \frac{4 \Delta^2}{2 \Delta x^2} \right) = \Theta \left( \frac{2}{3} \right).$$

In the low privacy regime ($\epsilon \to \infty$), the Podium mechanism is asymptotically equivalent to the Staircase mechanism and exponentially outperforms the Laplace mechanism.

**Theorem 3:** (Low Privacy Regime). In the low privacy regime, $\epsilon \to \infty$, the variance of the Podium mechanism in the extreme right shape is equal to

$$V_{PM}^{\epsilon \to \infty} = \Theta \left( \Delta^2 e^{-\frac{2\epsilon}{3}} \right).$$

**Proof.** In case when $s = \epsilon/3$,

$$V_{PM}^{\epsilon \to \infty} = \frac{\Delta^2 \cosh(-\frac{\epsilon}{3}) + 4 \cosh(\frac{\epsilon}{3}) + 3}{12 \cosh(\epsilon) - 1} = \frac{\Delta^2 5 \cosh(\frac{\epsilon}{3}) + 3}{12 \cosh(\epsilon) - 1} = \Theta \left( \frac{\Delta^2 \cosh(\frac{\epsilon}{3})}{\cosh(\epsilon)} \right)$$

$$= \Theta \left( \frac{\Delta^2 e^{3 \epsilon/8} e^{\epsilon}}{e^{3 \epsilon/2} e^{\epsilon}} \right) = \Theta \left( \Delta^2 e^{-\frac{2\epsilon}{3}} \right).$$

The Podium mechanism is exponentially more efficient than the Laplace mechanism in the low privacy regime.

$$e_{\epsilon \to \infty}(PM, LM) = \frac{V_{PM}}{V_{LM}} = \Theta \left( \frac{\Delta^2 e^{-\frac{2\epsilon}{3}}}{2 \Delta x^2} \right) = \Theta \left( e^{-\frac{2\epsilon}{3}} \right)$$

and is asymptotically equivalent to the Staircase mechanism.

Relative efficiencies for the three mechanisms are directly compared in Table [V]. The Podium mechanism outperforms the other two mechanisms and its approximation of $s = \epsilon/3$ is shown to be a very good one for all levels of $\epsilon$.

**Table II: Relative Efficiencies of the Three Mechanisms for Different Levels of $\epsilon$.**

| $\epsilon$ | $V_{PM}$ | $V_{LM}$ | $V_{PM,LM}$ | $V_{PM}$ | $V_{LM}$ | $V_{PM,LM}$ | $V_{PM}$ | $V_{LM}$ | $V_{PM,LM}$ |
|------------|----------|----------|-------------|----------|----------|-------------|----------|----------|-------------|
| 0.10       | 0.9639   | 0.6663   | 0.6666      | 0.6425   | 0.6666   |
| 0.20       | 0.9004   | 0.6653   | 0.6993      | 0.6200   | 0.6664   |
| 0.30       | 0.8933   | 0.6635   | 0.9963      | 0.5990   | 0.6661   |
| 0.40       | 0.8705   | 0.6611   | 0.9933      | 0.5794   | 0.6656   |
| 0.50       | 0.8438   | 0.6581   | 0.9896      | 0.5611   | 0.6650   |
| 0.60       | 0.8191   | 0.6543   | 0.9851      | 0.5441   | 0.6642   |
| 0.70       | 0.7962   | 0.6500   | 0.9798      | 0.5282   | 0.6634   |
| 0.80       | 0.7749   | 0.6450   | 0.9736      | 0.5133   | 0.6624   |
| 0.90       | 0.7553   | 0.6394   | 0.9667      | 0.4995   | 0.6614   |
| 1.00       | 0.7370   | 0.6332   | 0.9590      | 0.4866   | 0.6603   |
| log(3)     | 0.7204   | 0.6266   | 0.9508      | 0.4748   | 0.6590   |
| log(16)    | 0.6662   | 0.6403   | 0.7251      | 0.3594   | 0.6348   |
| log(32)    | 0.5409   | 0.3813   | 0.6082      | 0.3391   | 0.6270   |
| 5          | 0.5143   | 0.2296   | 0.3714      | 0.3180   | 0.6183   |
| 10         | 0.5005   | 0.0264   | 0.0424      | 0.3123   | 0.6239   |
| 20         | 0.5000   | 0.0000   | 0.0000      | 0.3149   | 0.6297   |
| 30         | 0.5000   | 0.0000   | 0.0000      | 0.3150   | 0.6300   |
| 40         | 0.5000   | 0.0000   | 0.0000      | 0.3150   | 0.6299   |
| 50         | 0.4977   | 0.0000   | 0.0000      | 0.3153   | 0.6335   |

**VI. Empirical Results and Comparisons**

We simulated our raw input values from the Beta(2, 2) distribution [II]. This distribution is symmetric around its mean 0.5 with the support in $[0, 1]$. Therefore, $\Delta$ in this case is equal to 1. We shift our distribution to the left by 0.5 to center it around 0. The range of input values after this transformation is $[-0.5, 0.5]$.

We simulated 10,000 random variables from the Beta distribution (raw input $x$) and added noise using the three
randomized mechanisms discussed in this work: the Laplace, Staircase and Podium. Results of our simulations are shown in Figure [6], which has six panels. The two rows represent different levels of privacy. The first row can be considered a high-privacy regime ($\epsilon = 1$) and the second row is a relatively low privacy regime ($\epsilon = 5$). The three columns represent the three different mechanisms.

In each panel, we plot the raw input $x$ on the x-axis versus the noisy privatized versions of $x$, $x'$, on the y-axis. In addition, we compute the variance of the noisy $x'$s in each panel for comparison. In a case one was interested in estimating the mean of $\mu = E(X)$ using the sample mean $\bar{x}'$, then the estimate of its variance would be equal to

$$V(\bar{x}') = \frac{V(x')}{10000}.$$

Therefore, the scaled variances shown in each panel also indicate how variable our estimate of $\mu$ would be in each case. Note that these are not noise variances, but the sum of variance of $X$ plus the noise variance.

There are three important takeaways from this visualization:

- The Podium mechanism outperforms the other two mechanisms in terms of efficiency. In fact, for $\epsilon = 1$, it reduces the variance of the privatized values by approximately half, even relative to the Staircase mechanism.

- For the Podium mechanism, it is easy to notice how truncation is contributing to the variance reduction. Its noise is distributed differently for different levels of $x$. At the extremes, the noise appears to be one-sided, which of course is not true because of the margin $m$.

- The Laplace mechanism is quite inefficient both in the low and medium privacy regimes.

The Podium mechanism results in smaller variance for the privatized distribution, and, therefore, smaller variance when estimating the mean or other measures of central tendency.

It is important to make the following observation from Figure 6. In the previous section, we have shown that the Podium and Staircase mechanisms are exponentially better than the Laplace mechanism in terms of efficiency as $\epsilon \to \infty$. Yet already at $\epsilon = 5$, variances of the samples collected are much closer together than variances of samples collect when $\epsilon = 1$, which seems counter-intuitive at first. This, of course, is very easily explainable upon further examination. The variance of samples collected is the sum of variance of the original $X$ and the variance added by the randomized mechanism $M$. For $\epsilon = 1$, these variances are comparable in magnitude and, therefore, it matters a lot which mechanism one chooses. But for $\epsilon = 5$, the variance of the randomized algorithm is so much smaller than the variance of $X$, that despite significant differences in their efficiency levels, the resulting variances of $X'$ under different mechanisms are quite close. As $\epsilon$ gets
even bigger, these variances essentially converge to the value of \( V(X) \), regardless of which mechanism one chooses. So, in practice, the exponential efficiency benefits of the Podium and Staircase mechanisms in low-privacy regimes matter little beyond theoretical curiosities.

VII. DISCUSSION

We presented a novel randomized algorithm, called the Podium mechanism, for achieving \( \varepsilon \)-differential privacy. It is characterized by the changing shape of its distribution depending on the input value \( x \) and its truncated nature. This is the first time that such truncated distribution was proposed for achieving \( \varepsilon \)-differential privacy.

The Podium mechanism is strictly better than either the Laplace or Staircase mechanisms and can be used in all places where the Laplace mechanism is currently being used, despite its optimality claims in a more narrow sense [12], [13]. Just like the Laplace mechanism, it requires generation of two uniform random variables, but has the additional burden of computing \( t \), the location of the step, at each noise addition. At this time, there is a plethora of literature on differential privacy in more complex settings [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24] where this work could be applicable or extended.

The benefits of the Podium mechanism really come through at the medium privacy regime (\( \varepsilon \in [1,3] \)). It has the smallest noise variance for values close to middle of the input range, making it necessary for practitioners to consider the shape of their input distribution. The more symmetric and centered it is, the more efficient the collection will be.

We hope that a slight additional complexity of generating random variables from the Podium distribution will not deter its practical adoption. After all, we all strive for more utility out of our data.

REFERENCES

[1] C. Dwork, F. McSherry, K. Nissim, and A. Smith, “Calibrating noise to sensitivity in private data analysis,” in Theory of Cryptography Conference (TCC), 2006, pp. 265–284.

[2] Q. Geng and P. Viswanath, “Optimal noise-adding mechanism in differential privacy,” CoRR, vol. abs/1212.1186, 2012. [Online]. Available: http://arxiv.org/abs/1212.1186

[3] S. L. Warner, “Randomized response: A survey technique for eliminating evasive answer bias,” Journal of the American Statistical Association, vol. 60, pp. 63–69, 1965.

[4] C. Dwork and A. Roth, “The algorithmic foundations of differential privacy,” Found. Trends Theor. Comput. Sci., vol. 9, no. 3, pp. 211–407, Aug. 2014. [Online]. Available: http://dx.doi.org/10.1561/0400000042

[5] U. Erlingsson, V. Pihur, and A. Korolova, “Rappor: Randomized aggregatable privacy-preserving ordinal response,” in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, ser. CCS ’14. New York, NY, USA: ACM, 2014, pp. 1054–1067. [Online]. Available: http://dx.doi.org/10.1145/2660267.2660348

[6] F. McSherry and K. Talwar, “Mechanism design via differential privacy,” in Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science, ser. FOCS ’07. Washington, DC, USA: IEEE Computer Society, 2007, pp. 94–103. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2007.41

[7] E. Lehmann, Elements of Large-Sample Theory, ser. Springer Texts in Statistics. Springer New York, 2004. [Online]. Available: https://books.google.com/books?id=peoxyvTXiEC

[8] Q. Geng, P. Kairouz, S. Oh, and P. Viswanath, “The staircase mechanism in differential privacy,” IEEE Journal of Selected Topics in Signal Processing, vol. 9, pp. 1176–1184, 2015.

[9] P. Kairouz, S. Oh, and P. Viswanath, “Extremal mechanisms for local differential privacy,” J. Mach. Learn. Res., vol. 17, no. 1, pp. 492–542, Jan. 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2946645.2946662

[10] Q. Geng, W. Ding, R. Guo, and S. Kumar, “Truncated laplacian mechanism for approximate differential privacy,” CoRR, vol. abs/1810.00877, 2018. [Online]. Available: http://arxiv.org/abs/1810.00877

[11] G. Casella and R. Berger, Statistical Inference. Duxbury Resource Center, June 2001.

[12] F. Koufogiannis, S. Han, and G. J. Pappas, “Optimality of the laplace mechanism in differential privacy,” CoRR, vol. abs/1504.00065, 2015. [Online]. Available: http://arxiv.org/abs/1504.00065

[13] R. Sarathy and K. Muralendhar, “Evaluating laplace noise addition to satisfy differential privacy for numeric data,” Trans. Data Privacy, vol. 4, no. 1, pp. 1–17, Apr. 2011. [Online]. Available: http://dl.acm.org/citation.cfm?id=2019312.2019313

[14] M. Hardt, K. Ligett, and F. McSherry, “A simple and practical algorithm for differentially private data release,” in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 2339–2347. [Online]. Available: http://papers.nips.cc/paper/4548-a-simple-and-practical-algorithm-for-differentially-private-data-release.pdf

[15] F. McSherry, “Privacy integrated queries: An extensible platform for privacy-preserving data analysis,” in Communications of The ACM - CACM, vol. 53, 01 2009, pp. 19–30.

[16] A. Roth and T. Roughgarden, “Interactive privacy via the median mechanism,” in Proceedings of the Forty-second ACM Symposium on Theory of Computing, ser. STOC ’10. New York, NY, USA: ACM, 2010, pp. 765–774. [Online]. Available: http://doi.acm.org/10.1145/1806689.1806794

[17] K. Chaudhuri and C. Monteleoni, “Privacy-preserving logistic regression,” in Proceedings of the 21st International Conference on Neural Information Processing Systems, ser. NIPS’08. USA: Curran Associates Inc., 2008, pp. 289–296. [Online]. Available: http://dl.acm.org/citation.cfm?id=2991780.2981817

[18] M. Hardt and N. Rothblum, “A multiplicative weights mechanism for privacy-preserving data analysis,” in Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science, ser. FOCS ’10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 61–70. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2010.85

[19] J. Zhang, Z. Zhang, X. Xiao, Y. Yang, and M. Winslett, “Functional mechanism: Regression analysis under differential privacy,” Proc. VLDB Endow., vol. 5, no. 11, pp. 1364–1375, Jul. 2012. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2147782.2150255

[20] M. Hardt, G. N. Rothblum, and R. A. Servedio, “Private data release via learning thresholds,” in Proceedings of the Twenty-third Annual ACM-SIAM Symposium on Discrete Algorithms, ser. SODA’12. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2012, pp. 168–187. [Online]. Available: http://dl.acm.org/citation.cfm?id=2095116.2095131

[21] V. Karwa, S. Raskhodnikova, A. D. Smith, and G. Yaroslavtsev, “Private analysis of graph structure,” ACM Trans. Database Syst., vol. 39, no. 3, pp. 22:1–22:33, 2014.

[22] J. Hsu, S. Khanna, and A. Roth, “Distributed private heavy hitters,” in Proceedings of the 39th International Colloquium Conference on Automata, Languages, and Programming - Volume Part I, ser. ICALP’12. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 461–472.

[23] A. Ghosh, T. Roughgarden, and M. Sundararajan, “Universally utility-maximizing privacy mechanisms,” in Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing, ser. STOC ’09. New York, NY, USA: ACM, 2009, pp. 351–360.

[24] C. Dwork, M. Naor, T. Pitassi, and G. N. Rothblum, “Differential privacy under continual observation,” in Proceedings of the Forty-second ACM Symposium on Theory of Computing, ser. STOC ’10. New York, NY, USA: ACM, 2010, pp. 715–724. [Online]. Available: http://doi.acm.org/10.1145/1806689.1806787

[25] L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 61–70. [Online]. Available: http://doi.acm.org/10.1145/1806689.1806794

[26] Duxbury Resource Center, June 2001.

[27] G. Casella and R. Berger, Statistical Inference. Duxbury Resource Center, June 2001.

[28] C. Dwork, F. McSherry, K. Nissim, and A. Smith, “Calibrating noise to sensitivity in private data analysis,” in Theory of Cryptography Conference (TCC), 2006, pp. 265–284.