Infinitely Divisible Noise in the Low Privacy Regime

Rasmus Pagh   PAGH@DI.KU.DK
BARC and University of Copenhagen

Nina Mesing Stausholm   NIMN@ITU.DK
IT University of Copenhagen

Editors: Sanjoy Dasgupta and Nika Haghtalab

Abstract

Federated learning, in which training data is distributed among users and never shared, has emerged as a popular approach to privacy-preserving machine learning. Cryptographic techniques such as secure aggregation are used to aggregate contributions, like a model update, from all users. A robust technique for making such aggregates differentially private is to exploit *infinite divisibility* of the Laplace distribution, namely, that a Laplace distribution can be expressed as a sum of i.i.d. noise shares from a Gamma distribution, one share added by each user.

However, Laplace noise is known to have suboptimal error in the low privacy regime for ε-differential privacy, where $\varepsilon > 1$ is a large constant. In this paper we present the first infinitely divisible noise distribution for real-valued data that achieves ε-differential privacy and has expected error that decreases exponentially with ε.

Keywords: Differential privacy, federated learning.

1. Introduction

Differential privacy, a state-of-the-art privacy definition, is a formal constraint on randomized mechanisms for privately releasing results of computations. It gives a formal framework for quantifying how well the privacy of an individual, whose data is part of the input, is preserved. In recent years, differentially private algorithms for machine learning have been developed and made available, for example, in TensorFlow/privacy and the Opacus library for PyTorch. Using such algorithms ensures that the presence or absence of a single data record in a database does not significantly affect the distribution of the model produced.

Concurrently, the area of *federated learning* (McMahan et al., 2017) has explored how to carry out machine learning in settings where training data is distributed among n users and never shared. Cryptographic techniques such as secure aggregation are used to aggregate contributions, like a model update, from all users (Bonawitz et al., 2017). This setup is used, for example, in the federated learning system run by Google on data from Android phones.[*] A robust technique for making such aggregates differentially private is to exploit *infinite divisibility* of the Laplace distribution, namely, that a Laplace distribution can be expressed as a sum of n i.i.d. noise shares from a Gamma distribution, one share added to the input of each user (Goryczka and Xiong, 2017). That is, if user i holds $x_i \in [0, \Delta]$, the input provided to the secure aggregation is $x_i + \eta_i$, where η_i is sampled from a suitable Gamma distribution. Infinitely divisible noise is resistant to *dropout*, where some users never contribute to the sum, if we set n to be a lower bound on the number of fully participating users.
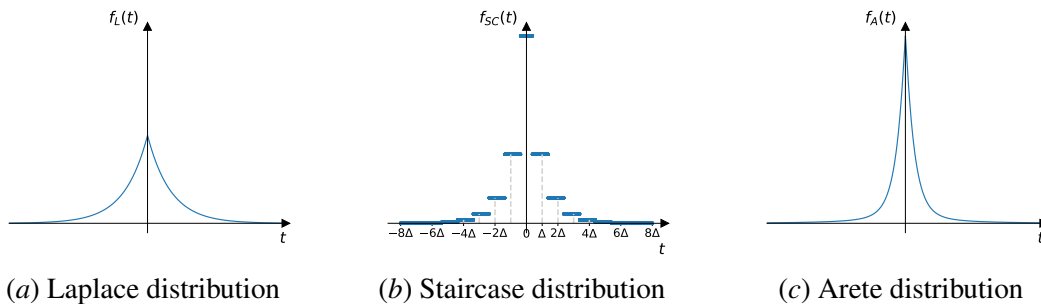
---

[*] https://ai.googleblog.com/2017/04/federated-learning-collaborative.html

(*a*) Laplace distribution   (*b*) Staircase distribution   (*c*) Arete distribution

Figure 1: Illustration of the density functions for the Laplace, Staircase and Arete distributions.

However, the Laplace noise needed for ε-differential privacy yields expected error $\Theta(\Delta/\varepsilon)$, which is not optimal in the "low privacy regime" when $\varepsilon \gg 1$. Geng, Kairouz, Oh, and Viswanath (2015) and Geng and Viswanath (2016a) presented the *Staircase* mechanism (see Lemma 30), which can be parameterized to obtain expected error $\Theta(\Delta e^{-\varepsilon/2})$ or variance $\Theta(\Delta^2 e^{-2\varepsilon/3})$, thus outperforming the Laplace mechanism for ε larger than some constant. However, the noise distribution used by this mechanism is not infinitely divisible, so it cannot replace Laplace noise in federated settings.

In this paper we present the first infinitely divisible noise distribution for real-valued data that achieves ε-differential privacy and has expected error that decreases exponentially with ε. Our new noise distribution, the *Arete** distribution, has expected absolute value and variance exponentially decreasing in ε, and thus comparable to that of the Staircase distribution up to constant factors in ε. Figure 1 illustrates the shape of each of the three distributions. The Arete distribution has a continuous density function implying that the privacy level decreases more smoothly with sensitivity, in contrast to the Staircase distribution (see discussion in Section 2).

The main part of this paper is devoted to the accuracy and privacy analysis of the Arete mechanism. Appendix A discusses applications in distributed private data analysis.

The Arete Distribution

For simplicity, we will limit ourselves to 1-dimensional setting. In order to deal with vectors (with ℓ_∞ sensitivity bounded by Δ), we may simply add independent noise from the Arete distribution to each coordinate. Our goal is to approximate the staircase distribution with an infinitely divisible distribution, so it is instructive to understand the essential properties of the staircase distribution: Only probability mass $\exp(-\Omega(\varepsilon))$ is placed in the tails, which can be seen as a piece-wise uniform version of a scaled Laplace distribution. The majority of the probability mass is placed in a uniform distribution on an interval around zero of length $\exp(-\Omega(\varepsilon))$.

Definition 1 (Arete distribution, informal) *Let independent random variables $X_1, X_2 \sim \Gamma(\alpha, \theta)$ and $Y \sim Laplace(\lambda)$. Then $Z := X_1 - X_2 + Y$ has Arete distribution with parameters α, θ and λ, denoted $Arete(\alpha, \theta, \lambda)$. When the parameters α, θ and λ are understood from the context, we use $f_A(t)$, $t \in \mathbb{R}$, to denote the density function of Z.*

---

*The name *Arete* is inspired by the word arête (pronounced "ah-ray't"), which is a sharp-crested mountain ridge, while also a concept from Greek mythology, Arete (pronounced "ah-reh-'tay") referring to moral virtue and excellence: the notion of the fulfillment of purpose or function and the act of living up to one's full potential (Wikipedia).

Since the Γ and Laplace distributions are continuous and infinitely divisible, and the Laplace distribution is symmetric, it follows that the Arete distribution also has these properties. In Section 3.2 we show:

Lemma 2 *For any choice of parameters $\alpha, \theta, \lambda > 0$, the $Arete(\alpha, \theta, \lambda)$ distribution is infinitely divisible and has density $f_A(t)$ that is continuous, symmetric around 0, and monotonely decreasing for $t > 0$.*

Next, we discuss the intuition behind the noise and privacy properties of the Arete distribution: For privacy parameter $\varepsilon > 0$ and sensitivity $\Delta > 0$ we concern ourselves with distributions $\mathcal{D}$ with support S and density function $f_\mathcal{D}$ satisfying

$$e^{-\varepsilon} \leq \frac{f_\mathcal{D}(t)}{f_\mathcal{D}(t+a)} \leq e^\varepsilon, \qquad \forall t, a \in \mathbb{R}, |a| \leq \Delta \tag{1}$$

as this property is sufficient to ensure differential privacy, which is our main goal. We will refer to the property (1) as the *differential privacy constraint*. In order to minimize the magnitude of the noise, the goal is to find a distribution with minimal expected (absolute) value while satisfying (1).

The difference of two Γ distributed random variables can be parameterized to have similar tails and to "peak" in an interval around zero of the same width as the staircase distribution. But this does not provide differential privacy since the density function has a singularity at zero. To achieve differential privacy we add a small amount of Laplace noise that "smooths out" the singularity. In more detail, the $\Gamma(\alpha, \theta)$-distribution (see Definition 25) with shape $\alpha < 1$, has most of its probability mass on an interval $(0, O(\alpha))$. The difference of two Γ distributions does not satisfy (1) for any choice of $\alpha < 1$, as the density tends to infinity for values going to zero. To fix this we need to "flatten the curve" of the density function in the neighborhood of 0. Consider $Z' := X + Y$ for independent $X \sim \Gamma(\alpha, \theta)$ and $Y \sim Exp(\lambda)$. The Exponential distribution, with a suitable choice of parameter λ, is used to flatten the density function of the Γ distribution close to 0. In order to get a noise distribution that is symmetric around zero, we further consider $Z = X_1 + Y_1 - (X_2 + Y_2)$ for $X_1, X_2 \sim \Gamma(\alpha, \theta)$ and $Y_1, Y_2 \sim Exp(\lambda)$. Our definition of the Arete distribution follows from the fact that if $Y_1, Y_2 \sim Exp(\lambda)$, then $Y = Y_1 - Y_2 \sim Laplace(\lambda)$. We provide an explicit setting for the parameters α, θ, λ in Lemma 3.

We note that the Arete distribution generalizes the Laplace distribution in the sense that we obtain $Laplace(\lambda)$ as the limiting distribution for $\alpha \to 0$. In this sense, the Arete distribution is suitable also for ε close to zero, where it may simply be used to implement Laplace noise.

Main Results

Let the Arete distribution $Arete(\alpha, \theta, \lambda)$ be as in Definition 1 (and formally, Definition 9) with density function f_A. In Section 4.3 we show the following result:

Lemma 3 *For every choice of $\Delta \geq 2/e$ and $\varepsilon \geq 20 + 4\ln(\Delta)$ there exist parameters $\alpha, \beta, \lambda > 0$ such that:*

- *For every choice of $t, a \in \mathbb{R}$ with $|a| \leq \Delta$, $e^{-\varepsilon} \leq \frac{f_A(t)}{f_A(t+a)} \leq e^\varepsilon$:*

- *For $Z \sim Arete(\alpha, \theta, \lambda)$, $\mathrm{E}[|Z|] = O(\Delta e^{-\varepsilon/4})$ and $\mathrm{Var}[Z] = O(\Delta^2 e^{-\varepsilon/4})$.*

Parameters $\alpha = e^{-\varepsilon/4}, \theta = \frac{4\Delta}{\varepsilon}$ and $\lambda = e^{-\varepsilon/4}$ suffice.
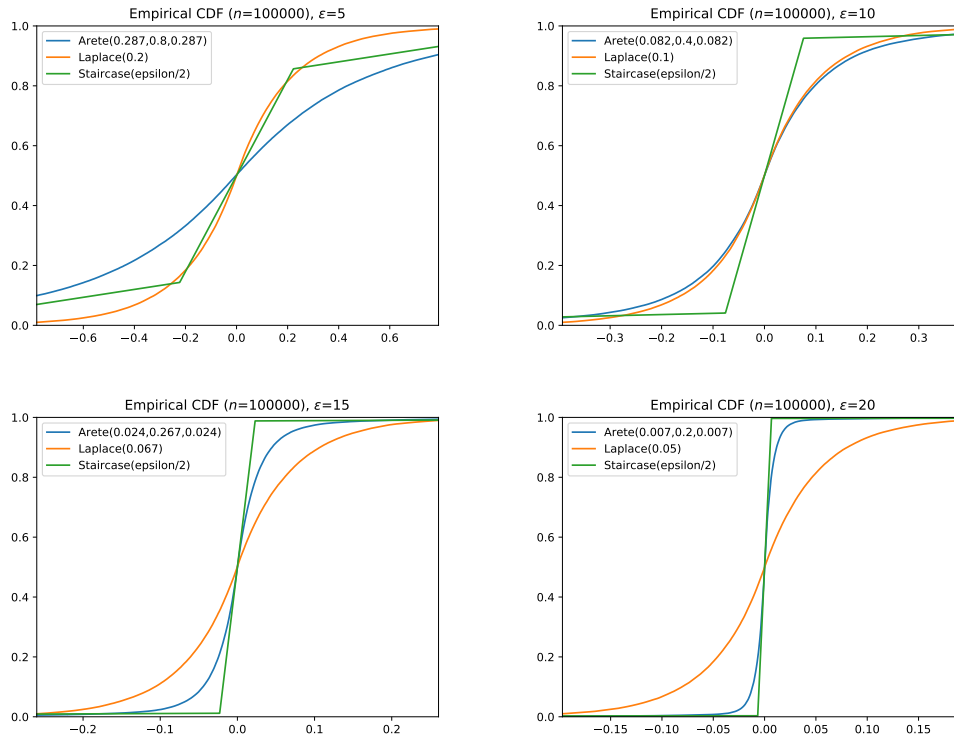
Figure 2: Empirical cumulative distribution functions for Arete distributions parameterized according to Lemma 3 for various values of ε, compared to two known ε-differentially noise distributions: Laplace$(1/\varepsilon)$, which is infinitely divisible, and a Staircase distribution (not infinitely divisible) parameterized to be $(\varepsilon + o(1))$-differentially private at neighbor distance $\Delta = 1 + o(1)$. The values of $\varepsilon < 20$ are for illustration only, since they are too small for Lemma 3 to apply.

The condition $\Delta \geq 2/e$ is not essential in the sense that we may always scale the noise by a factor Δ, which means that it is enough to consider sensitivity 1. This also means that $\varepsilon \geq 20$ suffices if we use a scaled Arete distribution.

Figure 2 shows empirical cumulative distribution functions for Arete distributions derived from Lemma 3. The code for generating these plots, as well as all other plots in this paper, can be found on GitHub.[†]

The following corollary, shown in Section 3.2, says that adding noise from the Arete distribution gives an ε-differentially private mechanism. We refer to Section B.2 for details about differential privacy and the definition of the sensitivity of a query.

Definition 4 (The Arete mechanism) *Let $x \in \mathcal{X}^d$ be an input and $q : \mathcal{X}^d \to \mathbb{R}$ a query with sensitivity bounded by $\Delta \geq 2/e$. Given parameters α, θ, λ, the* Arete *mechanism $\mathcal{M}_{Arete}(x)$ samples $Z \sim Arete(\alpha, \theta, \lambda)$ and returns $q(x) + Z$.*

[†]https://github.com/rasmus-pagh/alt22-code

Corollary 5 *The Arete mechanism $\mathcal{M}_{Arete}$ with parameters as specified in Lemma 3 has expected error $O(\Delta e^{-\varepsilon/4})$ and is ε-differentially private.*

Discussion of Large Values of ε. Values of ε larger than one often appear in practice. Examples of deployments using large values of ε include Google's RAPPOR with ε up to 9, Apple's MacOS with $\varepsilon = 6$, iOS 10 with $\varepsilon = 14$ (Greenberg, 2017), and US Census Bureau with ε up to 19.6.[‡]

These examples are not directly comparable to our setting since they deal with either local differential privacy (data of a single user), or with the release of many statistics (rather than a single aggregate). However, we note that mechanisms in the low-privacy regime can often be "boosted", e.g. using sampling (Balle et al., 2018) or shuffling (Erlingsson et al., 2019), to obtain a mechanism with a better privacy parameter. If sampling is used to improve privacy we will of course get noise due to sampling error, and this will dominate the total error in most cases. One way to use our result (with secure aggregation) is to essentially match the error of estimating a sum from a sample, getting privacy almost for free when the sample is much smaller than the whole data set.

The lower bound on ε in Lemma 3 is high, but we note that *empirically* we achieve differential privacy for significantly lower values of ε (see Section 5).

2. Related Works

A fundamental question is what can be said about the tradeoff between error and privacy. Hardt and Talwar (2010) study this tradeoff for linear queries, showing a lower bound of $\Omega(1/\varepsilon)$ for worst case expected ℓ_2-norm of noise (std. deviation) under the constraint of ε-differential privacy for small ε. Nikolov et al. (2013) extend the work of (Hardt and Talwar, 2010) to the tradeoff between error and (ε, δ)-differential privacy. For error that can be a general function of the added noise, Gupte and Sundararajan (2010) and Ghosh, Roughgarden, and Sundararajan (2012) introduced the Geometric Mechanism for *counting queries* (integer valued) with sensitivity 1, showing that the optimal noise has a (symmetric) Geometric distribution with error (standard deviation) $\Theta(e^{-\varepsilon/2})$. Brenner and Nissim (2010) extend these results, showing that for general queries there is no optimal mechanism for ε-differential privacy. In the high privacy regime, Geng and Viswanath (2016b) present a (near) optimal mechanism for integer-valued vector queries for (ε, δ)-differential privacy, achieving error (for single-dimensional queries) $\Theta(\min\{1/\varepsilon, 1/\delta\})$ for small ε and δ. Though the geometric mechanism yields optimal error in the *discrete* setting, and is infinitely divisible (Goryczka and Xiong, 2017), it does not seem to generalize to a differentially private, infinitely divisible noise distribution in the real-valued setting. Recently, Kairouz et al. (2021) studied a distributed, discrete version of the Gaussian mechanism, which has good composition properties. This mechanism is not aimed at the low privacy regime, and does not have error that decreases exponentially as ε grows.

Generalizing to *real-valued* 1-dimensional queries with arbitrary sensitivity, Geng and Viswanath (2016a) introduced the ε-differentially private Staircase mechanism (see Lemma 30), which adds noise from the Staircase distribution – a geometric mixture of uniform distributions. The density function of the Staircase distribution, f_{SC}, is a piece-wise continuous step (or "staircase-shaped") function, symmetric around zero, with geometrically decaying density as a function of the distance from zero. The staircase mechanism circumvents a lower bound of Koufogiannis et al. (2015) on "Lipschitz" differential privacy, which requires the privacy loss to be bounded by $\varepsilon|q(x) - q(y)|/\Delta$, by only bounding the worst case privacy loss for $|q(x) - q(y)|/\Delta \leq 1$.

[‡] https://www.census.gov/newsroom/press-releases/2021/2020-census-key-parameters.html

[Geng and Viswanath](2016a) prove that the optimal ε-differentially private mechanism for single real-valued queries, measuring error as expected magnitude or variance of noise, is not Laplace but rather Staircase distributed: while the Laplace mechanism is asymptotically optimal as $\varepsilon \to 0$, the Staircase mechanism performs better in the low privacy regime (i.e., for large ε), as the expected magnitude of the noise is exponentially decreasing in ε. Specifically, for sensitivity Δ and for the parameter setting of γ optimizing for expected noise magnitude, the Staircase mechanism achieves error $\Theta(\Delta e^{-\varepsilon/2})$. For the choice of γ optimizing for variance, the Staircase mechanism ensures variance of the noise $\Theta(\Delta^2 e^{-2\varepsilon/3})$. We note that the γ optimizing for noise magnitude is not generally the same a for optimizing for variance. The Laplace distribution has expected noise magnitude $\Theta(\Delta/\varepsilon)$ and variance $\Theta(\Delta^2/\varepsilon^2)$. In comparison, the expected noise magnitude and variance of the Arete distribution is also exponentially decreasing in ε, specifically $O(\Delta e^{-\varepsilon/4})$ and $O(\Delta^2 e^{-\varepsilon/4})$, respectively, for our choice of parameters. The expected error and variance mentioned here are for a single parameter setting for both Laplace and Arete mechanisms.

As we want a noise distribution that is implementable in a distributed setting, we limit our interest to noise distributions that are oblivious of the input data and the query output. A nice property of the Arete distribution is that the density function is continuous and so we get a more graceful decrease in privacy than the Staircase mechanism for inputs that are not quite neighboring. We can measure this using the *worst case privacy loss*, which is the logarithm of the largest ratio between the mechanism's density functions on inputs x and y. If $|q(x) - q(y)| \leq \Delta$, inequality (1) implies that the worst case privacy loss is at most ε. For inputs with $|q(x) - q(y)| > \Delta$ this is no longer guaranteed, but it is still interesting to study how the level of privacy decreases as a function of $|q(x) - q(y)|$. The Staircase mechanism is *exactly* fitted to the sensitivity of the query such that differential privacy is guaranteed for neighboring inputs, but as soon as $|q(x) - q(y)| > \Delta$ the worst case privacy loss is immediately doubled. The privacy loss increases in a smoother fashion when applying the Arete distribution due to the continuity of the density function (See Figure 3).

[Geng et al.](2015) extend the Staircase mechanism from ([Geng and Viswanath](2016a)) to queries in multiple dimensions.

3. The Arete Distribution

This section introduces the Arete distribution and some of its properties. We refer to Appendix B for definitions of probability distributions and differential privacy. The following lemma is well-known from the probability theory literature:

Lemma 6 *If X and Y are independent, continuous random variables with density functions f_X and f_Y, then $Z = X + Y$ is a continuous random variable where the density is the convolution*

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = \int_{-\infty}^{\infty} f_X(z-x) f_Y(x) dx.$$

The following distribution will be useful in defining the Arete distribution:

Definition 7 (The $\Gamma - \Gamma$ Distribution) *Let $X_1, X_2 \sim \Gamma(\alpha, \theta)$ be independent and consider their difference $X := X_1 - X_2$. We say that X has the $\Gamma - \Gamma(\alpha, \theta)$ distribution and the density of X is*

$$f_{\Gamma - \Gamma(\alpha,\theta)}(t) = \int_0^{\infty} f_{\Gamma(\alpha,\theta)}(t+x) f_{\Gamma(\alpha,\theta)}(x) dx = \begin{cases} \int_0^{\infty} f_\Gamma(t+x) f_\Gamma(x) dx, & t \geq 0 \\ \int_{|t|}^{\infty} f_\Gamma(t+x) f_\Gamma(x) dx, & t < 0. \end{cases}$$
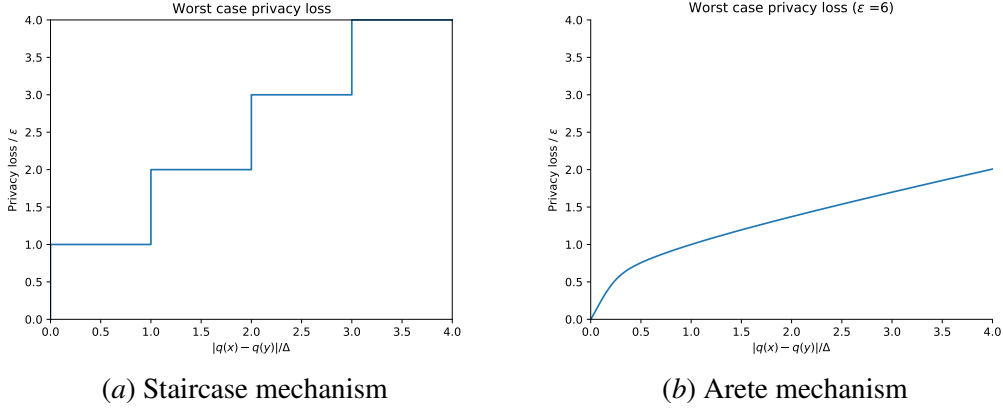
(*a*) Staircase mechanism   (*b*) Arete mechanism

Figure 3: Worst case privacy loss of the Staircase and Arete mechanisms (the latter with $\varepsilon = 6$) as a function of the difference between query outputs. As we have no closed form for the density of the Arete distribution, the plot on the right hand side is a numerical approximation computed by discretizing the involved Γ and Laplace distributions and computing the convolution of the discretizations. The privacy loss of the Arete distribution increases rapidly from distance 0 to about $\Delta/2$ and then increases more slowly from $\Delta/2$ to Δ where it reaches ε. As can be seen the Arete distribution has a smaller privacy loss than the staircase distribution for larger distances $|q(x) - q(y)|$, suggesting that it in fact adds more noise than necessary.

where the integrals are reduced to the intervals where $f_\Gamma(t + x)f_\Gamma(x)$ is non-zero.

Lemma 8   *The $\Gamma - \Gamma$ distribution is infinitely divisible: For $2n$ independent random variables $X_i, Y_i \sim \Gamma(\alpha/n, \theta)$, we have $X = \sum_{i=1}^{n}(X_i - Y_i) \sim \Gamma - \Gamma\left(\alpha, \theta\right)$.*

Proof The result follows immediately from infinite divisibility of the Γ-distribution.   ∎

Definition 9 (The Arete distribution)   *Let $X \sim \Gamma - \Gamma(\alpha, \theta)$ and $Y \sim Laplace(\lambda)$ be independent. Define $Z := X + Y$, then $Z \sim Arete(\alpha, \theta, \lambda)$ for $\alpha, \theta, \lambda > 0$. The density of Z is*

$$f_{A(\alpha, \theta, \lambda)}(t) = \int_{-\infty}^{\infty} f_{\Gamma - \Gamma(\alpha, \theta)}(t - x) f_{L(\lambda)}(x) dx = \int_{-\infty}^{\infty} f_{L(\lambda)}(t - x) f_{\Gamma - \Gamma(\alpha, \theta)}(x) dx, \quad t \in \mathbb{R}.$$

Lemma 10   *The Arete distribution is infinitely divisible: For $4n$ independent random variables $X_{1i}, X_{2i} \sim \Gamma(\alpha/n, \theta)$ and $Y_{1i}, Y_{2i} \sim \Gamma(1/n, \lambda)$, we have $X = \sum_{i=1}^{n}(X_{1i} - X_{2i} + (Y_{1i} - Y_{2i})) \sim Arete\left(\alpha, \theta, \lambda\right)$.*

Proof The result follows immediately from infinite divisibility of the Laplace distribution and Lemma 8.   ∎

Note 1   *We remark that we are only interested in $0 < \alpha < 1$. Furthermore, we do not explicitly state the density of the Arete distribution, as there is no simple closed form for the density of the $\Gamma - \Gamma$ distribution. (It can, however, be expressed in terms of Bessel functions – see (Mathai, 1993).)*

A similar intuitive way of defining our distribution would be to use a symmetric version of the Γ-distribution (two halved Γ-distributions put back-to-back at zero), instead of the $\Gamma - \Gamma$-distribution. An important property of our distribution is infinite divisibility such that we can draw independent noise shares that sum to a random variable following the Arete distribution. As opposed to our $\Gamma - \Gamma$ distribution, it is not clear whether a symmetric Γ-distribution is infinitely divisible.

3.1. Symmetric Density Functions

We observe some simple properties of the Arete distribution.

Lemma 11 *For $f, g : \mathbb{R} \to \mathbb{R}$, that are symmetric around 0, i.e., $f(x) = f(-x)$ and $g(x) = g(-x)$, we have for any $t \in \mathbb{R}$*

$$\int_{-\infty}^{\infty} f(x)g(t - x)dx = \int_{-\infty}^{\infty} f(x)g(|t| - x)dx.$$

In particular, the convolution $f * g$ is symmetric around 0.

Proof The statement is immediate for $t \geq 0$, so suppose $t < 0$. Then for any $a, b \in \mathbb{R}$

$$\int_{-a}^{-b} f(x)g(t - x)dx = \int_{-a}^{-b} f(x)g(|t| + x)dx = \int_{b}^{a} f(-x)g(|t| - x)dx = \int_{b}^{a} f(x)g(|t| - x)dx$$

where first step is by symmetry of g, the second step follows from integration by substitution and the last step is by symmetry of f. In particular, we may let a and b be $\pm\infty$. ∎

Lemma 12 $f_{\Gamma - \Gamma}$ *is symmetric around 0.*

Proof We prove that $f_{\Gamma - \Gamma}(t) = f_{\Gamma - \Gamma}(|t|)$ for all $t \in \mathbb{R}$. Clearly, this is the case if $t \geq 0$, so suppose $t < 0$. By Definition 7

$$f_{\Gamma - \Gamma}(t) = \int_{|t|}^{\infty} f_{\Gamma}(t + x) f_{\Gamma}(x)dx = \int_{|t|}^{\infty} f_{\Gamma}(x - |t|) f_{\Gamma}(x)dx = \int_{0}^{\infty} f_{\Gamma}(x) f_{\Gamma}(|t| + x)dx = f_{\Gamma - \Gamma}(|t|).$$

where the penultimate step follows from integration by substitution with $x - |t|$. ∎

Corollary 13 f_A *is symmetric around 0.*

Proof The result follows directly from symmetry of the density of the Laplace distribution, f_L, and Lemmas 11 and 12. ∎

3.2. Properties of the Arete Distribution

We restate the lemma here for convenience:

Lemma 14 *For any choice of parameters $\alpha, \theta, \lambda > 0$, the $Arete(\alpha, \theta, \lambda)$ distribution is infinitely divisible and has density $f_A(t)$ that is continuous, symmetric around 0, and monotonely decreasing for $t > 0$.*

Proof Symmetry of the density function f_A is proven in Corollary 13 and infinite divisibility in Lemma 10. Since f_Γ and f_L are continuous, $f_{\Gamma-\Gamma}$ and f_A are also continuous by Lemma 6. We prove that f_A is monotonely decreasing, i.e., for $|t| \leq |t'|$ we have $f_A(t) \geq f_A(t')$. First, we argue that $f_{\Gamma-\Gamma}$ is monotonely decreasing. Recall Definition 7 and observe

$$f_{\Gamma-\Gamma}(t) = \int_0^\infty f_\Gamma(|t| + x) f_\Gamma(x) dx, \qquad \forall t \in \mathbb{R}$$

which is immediate for $t \geq 0$ while for $t < 0$

$$f_{\Gamma-\Gamma}(t) = \int_{|t|}^\infty f_\Gamma(-|t| + x) f_\Gamma(x) dx = \int_0^\infty f_\Gamma(x') f_\Gamma(x' + |t|) dx'$$

where we substituted $x' := x - |t|$. So assume $|t| \leq |t'|$. Then, since f_Γ is monotonely decreasing

$$f_{\Gamma-\Gamma}(t) = \int_0^\infty f_\Gamma(|t| + x) f_\Gamma(x) dx \geq \int_0^\infty f_\Gamma(|t'| + x) f_\Gamma(x) dx = f_{\Gamma-\Gamma}(t').$$

We prove that f_A is also monotonely decreasing: Assuming that $|t| \leq |t'|$ we prove that $f_A(t) \geq f_A(t')$. Recall Definition 9 and observe

$$f_A(t) = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(|t| - x) f_L(x) dx = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(x) f_L(|t| - x) dx$$

which is obvious for $t \geq 0$ and since for $t < 0$:

$$f_A(t) = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(t - x) f_L(x) dx = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(|t| + x) f_L(x) dx = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(x') f_L(|t| - x') dx'$$

using that $f_{\Gamma-\Gamma}$ and f_L are symmetric and a substitution with $x' := |t| + x$. A similar argument can be made if the convolution is flipped. We conclude that

$$f_A(t) = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(|t| - x) f_L(x) dx \geq \int_{-\infty}^\infty f_{\Gamma-\Gamma}(|t'| - x) f_L(x) dx = f_A(t')$$

using that $f_{\Gamma-\Gamma}$ is monotonely decreasing. ∎

We finally assume Lemma 3 and prove Corollary 5, restated here for convenience. The proof of Lemma 3 is given in Section 4.

Corollary 15 *The Arete mechanism $\mathcal{M}_{Arete}$ with parameters as specified in Lemma 3 has expected error $O(\Delta e^{-\varepsilon/4})$ and is ε-differentially private.*

Proof The expected error bound follows directly from the bound on $\mathrm{E}[|Z|]$ in Lemma 3. For the claim of differential privacy, let $x, x' \in \mathbb{R}$ with $|x - x'| \leq 1$. We show that for any subset $S \subset \mathbb{R}$

$$\Pr[\mathcal{M}_{Arete}(x) \in S] \leq e^{\varepsilon} \Pr[\mathcal{M}_{Arete}(x') \in S]. \tag{2}$$

Let noise $Z \sim Arete(\alpha, \theta, \lambda)$ for parameters α, θ, λ as in Lemma 3.
Define $S' := S - q(x) = \{s - q(x) : s \in S\}$, then:

$$\frac{\Pr[\mathcal{M}_{Arete}(x) \in S]}{\Pr[\mathcal{M}_{Arete}(x') \in S]} = \frac{\int_{S'} f_A(z) dz}{\int_{S'} f_A(z + q(x') - q(x)) dz} \leq \frac{\int_{S'} f_A(|z|) dz}{\int_{S'} f_A(|z| + |q(x') - q(x)|) dz}$$

where we used symmetry of f_A, the triangle inequality, and the fact that $f_A(t)$ is decreasing for $t > 0$. By assumption $|q(x) - q(x')| \leq \Delta$. Lemma 3 says that $f_A(t)/f_A(t + a) \leq e^{\varepsilon}$ for all $t \in \mathbb{R}$ and $a \leq \Delta$, and so we get

$$\frac{\int_{S'} f_A(|z|) dz}{\int_{S'} f_A(|z| + |q(x') - q(x)|) dz} \leq \frac{\int_{S'} e^{\varepsilon} f_A(|z| + \Delta) dz}{\int_{S'} f_A(|z| + |q(x') - q(x)|) dz} \leq e^{\varepsilon}.$$

The first inequality in (2) follows by symmetry. ∎

4. Proof of Main Lemma

In the remaining part of this section we prove a number of theoretical lemmas, that will help prove our main result, Lemma 3. The bulk of the analysis is the proof of the first bullet point of Lemma 3, showing that the given parameters $\alpha, \theta, \lambda > 0$ suffice to bound $f_A(t)/f_A(t + a)$ for all $t, a \in \mathbb{R}$, $|a| \leq \Delta$. We break this part of the analysis down in this section. The intuition behind the structure is as follows: We first remark that (to the best of our knowledge) there is no simple expression for the density of the $\Gamma - \Gamma$ distribution (see Note 1). Hence, we will show upper and lower bounds for $f_{\Gamma - \Gamma}$ and use these to bound the ratio $f_A(t)/f_A(t + a)$. As discussed earlier, we have not optimized for constants, and as our proof includes *several* steps of bounding, our analysis may not be tight, thus leading to the high value of ε required in Lemma 3. A tighter analysis is likely to allow for a better setting of parameters α, θ, λ and a smaller ε. We give the proof of Lemma 3 in Section 4.3.

4.1. Bounds on Density of $\Gamma - \Gamma$ Distribution

We first derive upper and lower bounds on the density function of the $\Gamma - \Gamma$ distribution (see Section B for definitions).

Lemma 16 *For any $t \in \mathbb{R}$ and any $\zeta_{\Gamma} > 0$*

$$f_{\Gamma}(|t| + \zeta_{\Gamma}) c_{\zeta_{\Gamma}} \leq f_{\Gamma - \Gamma}(t) \leq f_{\Gamma}(|t|) \qquad where \qquad c_{\zeta_{\Gamma}} := \int_0^{\zeta_{\Gamma}} f_{\Gamma}(x) dx \ .$$

Proof Recall Definition 7 and Lemma 12 and let $t \in \mathbb{R}$. For the upper bound, we have

$$f_{\Gamma - \Gamma}(t) = \int_0^{\infty} f_{\Gamma}(|t| + x) f_{\Gamma}(x) dx < f_{\Gamma}(|t|) \int_0^{\infty} f_{\Gamma}(x) dx = f_{\Gamma}(|t|).$$

For the lower bound we have for any $\zeta_\Gamma > 0$

$$f_{\Gamma-\Gamma}(t) = \int_0^\infty f_\Gamma(|t| + x)f_\Gamma(x)dx \geq \int_0^{\zeta_\Gamma} f_\Gamma(|t| + x)f_\Gamma(x)dx \geq f_\Gamma(|t| + \zeta_\Gamma) \int_0^{\zeta_\Gamma} f_\Gamma(x)dx \ .$$

■

4.2. Bounds on Density of Arete Distribution

In this section we show that for $\Delta > 0$ and setting of parameters α, θ, λ and for large enough ε:

$$e^{-\varepsilon} \ \leq \ f_A(t)/f_A(t+\Delta) \ \leq \ e^\varepsilon, \qquad \forall t \in \mathbb{R}. \tag{3}$$

We remark that by monotonicity of the density of the Arete distribution, if we show (3) it follows that f_A satisfies (1): Take any $a \in \mathbb{R}$ such that $|a| \leq \Delta$ and suppose without loss of generality that $f(t) \geq f(t+a)$ (if this is not the case, substitute $t' := |t| - a$, such that $f(t') \geq f(t'+a)$). Then $e^{-\varepsilon} \leq \frac{f(t)}{f(t+a)}$. We prove that $f(t+a) \geq f(t+\Delta)$ ensuring $\frac{f(t)}{f(t+a)} \leq \frac{f(t)}{f(t+\Delta)} \leq e^\varepsilon$, which finishes the argument: by assumption $f(t) \geq f(t+a)$ and so $|t| \leq |t+a|$, further implying that $t \geq -a/2$. Hence, as $|a| \leq \Delta$ we have $|t+\Delta| \geq |t+a|$ and so we conclude that $f(t+a) \geq f(t+\Delta)$ as wanted.

Throughout the section we assume that $|t| \leq |t+\Delta|$ (and so $t \geq -\Delta/2$). For such t, $f_A(t) \geq f_A(t+\Delta)$ and so the first inequality in (3) is immediate. Hence, we put our focus toward proving the latter inequality. If $|t+\Delta| \leq |t|$, the result follows by symmetry of f_A (Corollary 13).

We start with the following lower bound on the density f_A:

Lemma 17 *Let ζ_Γ and c_{ζ_Γ} be as in Lemma 16 and assume $\lambda \leq \Delta/\ln(2)$ for $\Delta > 0$. For $-\Delta/2 \leq t \in \mathbb{R}$*

$$f_A(t+\Delta) \geq f_\Gamma(|t+\Delta| + \zeta_\Gamma)c_{\zeta_\Gamma}c_L \qquad where \qquad c_L := 1/4.$$

Proof By Definition 9 and Lemma 16 we have

$$f_A(t+\Delta) = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(t+\Delta-x)f_L(x)dx \geq \int_{-\infty}^\infty f_\Gamma(|t+\Delta-x| + \zeta_\Gamma)c_{\zeta_\Gamma}f_L(x)dx$$

$$\geq c_{\zeta_\Gamma} \int_0^{2(t+\Delta)} f_\Gamma(|t+\Delta-x| + \zeta_\Gamma)f_L(x)dx \geq c_{\zeta_\Gamma}f_\Gamma(t+\Delta+\zeta_\Gamma) \int_0^{2(t+\Delta)} f_L(x)dx,$$

where we used that $f_\Gamma(|t+\Delta-x| + \zeta_\Gamma) \geq f_\Gamma(|t+\Delta| + \zeta_\Gamma)$ for $x \in (0, 2(t+\Delta))$ and that by assumption $t+\Delta \geq \Delta/2$ allowing us to remove the absolute value signs. Again using that $t+\Delta \geq \Delta/2$

$$\int_0^{2(t+\Delta)} f_L(x)dx \geq \int_0^\Delta f_{L(\lambda)}(x)dx = \frac{1}{2}\int_0^\Delta f_{Exp(\lambda)}(x)dx \geq \frac{1}{4}, \qquad \lambda < \Delta/\ln(2)$$

where we noticed that on the positive reals, the density function of the Laplace distribution is $1/2$ times the density function of the Exponential distribution, and used that the median of the latter is $\ln(2)\lambda$, so that the last inequality is true as long as $\ln(2)\lambda \leq \Delta$. Hence,

$$f_A(t+\Delta) \geq c_{\zeta_\Gamma}f_\Gamma(t+\Delta+\zeta_\Gamma)1/4.$$

11

Defining $c_L := 1/4$ finishes the proof. ■

The following three lemmas are technical and give upper bounds for the ratio $f_A(t) \, / \, f_A(t + \Delta)$; first for large and small $|t|$ separately in Lemmas 18 and 19 (i.e., for t close to and far from 0, where "close to/far from" is quantified by a parameter ζ_u, which we will set in Lemma 21). We combine these results to an upper bound for general t in Lemma 20 (still assuming t is s.t. $f_A(t) \geq f_A(t+\Delta)$) and finally choose parameters to ensure an upper bound of e^ε in Lemma 21, thus satisfying the second inequality of (3). Throughout the next three lemmas we make use of the variables $\zeta_\Gamma, c_{\zeta_\Gamma}$ (from Lemma 16) and c_L (from Lemma 17), all of which will be handled in the proof of Lemma 21.

Lemma 18 *Let $\zeta_u > 0$ be given and assume $0 < \alpha \leq 1$. Let $\zeta_\Gamma, c_{\zeta_\Gamma}$ be as in Lemma 16 and c_L as in Lemma 17. Assume $1/\lambda - 1/\theta \geq 1/(\zeta_u + \Delta + \zeta_\Gamma)$ and $\lambda \leq \Delta/\ln(2)$ for $\Delta > 0$. For $-\Delta/2 < t \in \mathbb{R}$ with $|t| \geq \zeta_u$ we have*

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{c_{\zeta_\Gamma}} \left(\left(1 + \frac{\Delta + \zeta_\Gamma}{\zeta_u} \right) + \frac{c_{\zeta_u}\Gamma(\alpha)\theta^\alpha}{2\lambda c_L} e^{\zeta_u/\theta}(\zeta_\Gamma + \Delta + \zeta_u)^{1-\alpha} \right)$$

where $c_{\zeta_u} := 2 \int_0^{\zeta_u} f_\Gamma(x)dx$.

Proof The proof can be found in Appendix C.1. ■

Lemma 19 *Let $\zeta_u > 0$ be given and assume $0 < \alpha \leq 1$. Let $\zeta_\Gamma, c_{\zeta_\Gamma}$ be as in Lemma 16 and c_L as in Lemma 17. Assume $1/\lambda - 1/\theta \geq 1/(\Delta + \zeta_\Gamma)$ and $\lambda \leq \Delta/\ln(2)$ for $\Delta > 0$. For $-\Delta/2 < t \in \mathbb{R}$ such that $|t| \leq \zeta_u$ we have*

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{c_L c_{\zeta_\Gamma}\lambda} \left(\frac{\zeta_u + \zeta_\Gamma + \Delta}{\alpha} + \Gamma(\alpha)\theta^\alpha(\zeta_\Gamma + \Delta)^{1-\alpha} \right)$$

Proof The proof can be found in Appendix C.2. ■

The following lemma combines Lemmas 18 and 19 to give an upper bound for general $t > -\Delta/2$:

Lemma 20 *Let $\zeta_\Gamma, c_{\zeta_\Gamma}$ be as in Lemma 16 and c_L as in Lemma 17. Assume $0 < \alpha \leq 1$, $\theta \leq \zeta_\Gamma + 1$, $\lambda \leq \min\{\theta/2, \Delta/\ln(2)\}$ and $\Gamma(\alpha) \leq 1/\alpha$ for $\Delta > 0$. For $-\Delta/2 < t \in \mathbb{R}$*

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq \frac{2e^{(\zeta_\Gamma + \Delta)/\theta}e^\alpha(\alpha\theta + \zeta_\Gamma + \Delta)}{\alpha c_{\zeta_\Gamma} c_L \lambda}$$

Proof The proof can be found in Appendix C.3. ■

Note 2 *The fact that $\Gamma(\alpha) \leq 1/\alpha$ for $0 < \alpha \leq 1$ follows from Euler's definition of the Gamma function,*

$$\Gamma(\alpha) = \frac{1}{\alpha} \prod_{n=1}^\infty \frac{(1 + \frac{1}{n})^\alpha}{1 + \frac{\alpha}{n}},$$

since $(1 + \frac{1}{n})^\alpha \leq 1 + \frac{\alpha}{n}$ for any $n > 0$ and $0 < \alpha \leq 1$.

We finally choose parameters α, θ, λ ensuring that the ratio $e^{-\varepsilon} \leq f(t)/f(t+\Delta) \leq e^{\varepsilon}$ for ε large enough:

Lemma 21 *Suppose $\varepsilon \geq 20 + 4\ln(\Delta)$ for $\Delta \geq 2/e$. Let $\alpha = e^{-\varepsilon/4}, \theta = \frac{4\Delta}{\varepsilon}$ and $\lambda = e^{-\varepsilon/4}$. Then for $t \in \mathbb{R}$*

$$e^{-\varepsilon} \leq \frac{f_A(t)}{f_A(t+\Delta)} \leq e^{\varepsilon}.$$

Proof The proof can be found in Appendix C.4. ∎

4.3. Putting Things Together

We restate the lemma here for convenience:

Lemma 22 *For every choice of $\Delta \geq 2/e$ and $\varepsilon \geq 20 + 4\ln(\Delta)$ there exist parameters $\alpha, \beta, \lambda > 0$ such that:*

- *For every choice of $t, a \in \mathbb{R}$ with $|a| \leq \Delta$, $e^{-\varepsilon} \leq \frac{f_A(t)}{f_A(t+a)} \leq e^{\varepsilon}$:*

- *For $Z \sim Arete(\alpha, \theta, \lambda)$, $\mathrm{E}[|Z|] = O(\Delta e^{-\varepsilon/4})$ and $\mathrm{Var}[Z] = O(\Delta^2 e^{-\varepsilon/4})$.*

Parameters $\alpha = e^{-\varepsilon/4}, \theta = \frac{4\Delta}{\varepsilon}$ and $\lambda = e^{-\varepsilon/4}$ suffice.

Proof The first bullet with the choice of parameters $\alpha = e^{-\varepsilon/4}, \theta = 4\Delta/\varepsilon$ and $\lambda = e^{-\varepsilon/4}$ follow from Lemma 21 and monotonicity of f_A (Lemma 2), as described at the beginning of Section 4.2. The second bullet also follows from Lemma 21, as the expected error of a random variable $Z = X + Y$, where $X \sim \Gamma - \Gamma(\alpha, \theta)$ and $Y \sim Laplace(\lambda)$, i.e., $Z \sim Arete(\alpha, \theta, \lambda)$, is

$$\mathrm{E}[|Z|] = \mathrm{E}[|X_1 - X_2 + Y|] \leq 2\,\mathrm{E}[X_1] + \mathrm{E}[|Y|] = 2\alpha\theta + \lambda = \frac{8\Delta e^{-\varepsilon/4}}{\varepsilon} + e^{-\varepsilon/4} = O\left(\frac{\Delta}{\varepsilon}e^{-\varepsilon/4}\right)$$

where $X_1, X_2 \sim \Gamma(\alpha, \theta)$, and similarly, by independence

$$\mathrm{Var}[Z] = \mathrm{Var}[X + Y] = \mathrm{Var}[X_1 - X_2 + Y] = 2\,\mathrm{Var}[X_1] + \mathrm{Var}[Y] = 2\alpha\theta^2 + 2\lambda^2$$
$$= 2\frac{16\Delta^2 e^{-\varepsilon/4}}{\varepsilon^2} + 2e^{-\varepsilon/2} = O\left(\frac{\Delta^2}{\varepsilon^2}e^{-\varepsilon/4}\right)$$

with our choice of parameters. Finally, for $\varepsilon \geq 1/\sqrt{2}$ (which is significantly smaller than the values of ε that we are interested in), we may simplify to

$$\mathrm{E}[|Z|] = O\left(\Delta e^{-\varepsilon/4}\right), \qquad \mathrm{Var}[Z] = O\left(\Delta^2 e^{-\varepsilon/4}\right)$$
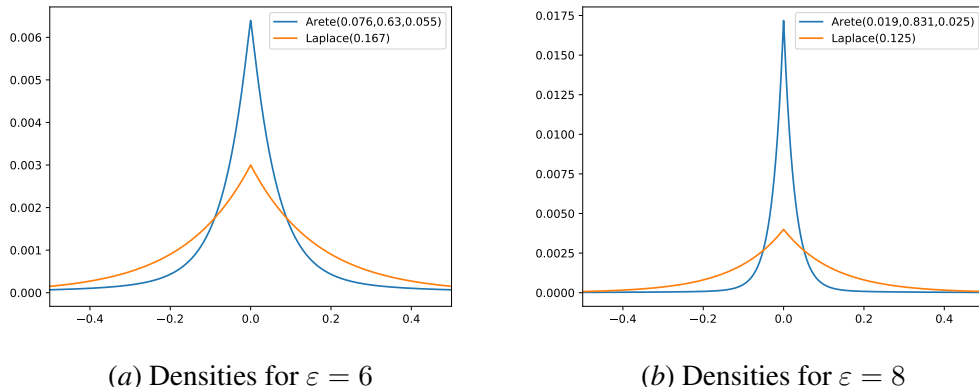
thus finishing the proof. ∎

(a) Densities for $\varepsilon = 6$ (b) Densities for $\varepsilon = 8$

Figure 4: Density functions for Arete distributions that empirically yield ε-DP with $\varepsilon = 6$ and $\varepsilon = 8$, respectively. The density functions were approximated by rounding the constituent Γ and Laplace distribution values to a multiple of 0.001 and computing the discrete convolution. Parameters were found using a local search heuristic. For comparison, Laplace distributions with the same privacy guarantee have been included and are clearly less concentrated around zero.

5. Conclusion

In this work we have seen a new noise distribution, the Arete distribution, which has a continuous density function, is symmetric around zero, monotonely decreasing for $t > 0$, infinitely divisible and has expected absolute value and variance exponentially decreasing in ε. The Arete distribution yields an ε-differentially private mechanism, the Arete mechanism, which is an *infinitely divisible* alternative to the Staircase mechanism (Geng and Viswanath, 2016a) with a *continuous* density function. The Arete mechanism achieves error comparable to the Staircase mechanism and outperforms the Laplace mechanism (Dwork et al., 2006b) in terms of absolute error and variance in the low privacy regime (for large ε).

Simulations suggest that the constant factors of the Arete mechanism, with parameters chosen as we have described, can be improved (see Figure 4). We leave open the question of finding an optimal (up to lower order terms), infinitely divisible error distribution for differential privacy in the low privacy regime.

Acknowledgments

References

Apple. Apple differential privacy technical overview. https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf. Accessed: 2021-07-06.

Borja Balle, Gilles Barthe, and Marco Gaboardi. Privacy amplification by subsampling: Tight analyses via couplings and divergences. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 6280–6290, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/3b5020bb891119b9f5130f1fea9bd773-Abstract.html.

Borja Balle, James Bell, Adrià Gascón, and Kobbi Nissim. The privacy blanket of the shuffle model. In *Advances in Cryptology - CRYPTO*, volume 11693 of *Lecture Notes in Computer Science*, pages 638–667, 2019. doi: 10.1007/978-3-030-26951-7_22. URL https://doi.org/10.1007/978-3-030-26951-7_22.

Borja Balle, James Bell, Adrià Gascón, and Kobbi Nissim. Private summation in the multi-message shuffle model. In *Conference on Computer and Communications Security, CCS*, CCS '20, page 657–676, New York, NY, USA, 2020. ISBN 9781450370899. doi: 10.1145/3372297.3417242. URL https://doi.org/10.1145/3372297.3417242.

Andrea Bittau, Úlfar Erlingsson, Petros Maniatis, Ilya Mironov, Ananth Raghunathan, David Lie, Mitch Rudominer, Ushasree Kode, Julien Tinnés, and Bernhard Seefeld. Prochlo: Strong privacy for analytics in the crowd. In *Symposium on Operating Systems Principles, SOSP*, pages 441–459, 2017. doi: 10.1145/3132747.3132769. URL https://doi.org/10.1145/3132747.3132769.

Keith Bonawitz, Vladimir Ivanov, Ben Kreuter, Antonio Marcedone, H. Brendan McMahan, Sarvar Patel, Daniel Ramage, Aaron Segal, and Karn Seth. Practical secure aggregation for privacy-preserving machine learning. In *Conference on Computer and Communications Security, CCS*, pages 1175–1191, 2017. doi: 10.1145/3133956.3133982. URL https://doi.org/10.1145/3133956.3133982.

Hai Brenner and Kobbi Nissim. Impossibility of differentially private universally optimal mechanisms. In *Symposium on Foundations of Computer Science, FOCS*, pages 71–80, 2010. doi: 10.1109/FOCS.2010.13. URL https://doi.org/10.1109/FOCS.2010.13.

T.-H. Hubert Chan, Elaine Shi, and Dawn Song. Privacy-preserving stream aggregation with fault tolerance. In *Financial Cryptography and Data Security - FC*, volume 7397 of *Lecture Notes in Computer Science*, pages 200–214, 2012. doi: 10.1007/978-3-642-32946-3_15. URL https://doi.org/10.1007/978-3-642-32946-3_15.

Jeesen Chen and Herman Rubin. Bounds for the difference between median and mean of Gamma and Poisson distributions. *Statistics & Probability Letters*, 4(6):281–283, October 1986. URL https://ideas.repec.org/a/eee/stapro/v4y1986i6p281-283.html.

Albert Cheu, Adam D. Smith, Jonathan R. Ullman, David Zeber, and Maxim Zhilyaev. Distributed differential privacy via shuffling. In *Advances in Cryptology, EUROCRYPT*, volume 11476 of *Lecture Notes in Computer Science*, pages 375–403, 2019. doi: 10.1007/978-3-030-17653-2_13. URL https://doi.org/10.1007/978-3-030-17653-2_13.

Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. In *Advances in Neural Information Processing Systems, NIPS*, pages 3571–3580, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/253614bbac999b38b5b60cae531c4969-Abstract.html.

John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. Local privacy and statistical minimax rates. In *54th Symposium on Foundations of Computer Science, FOCS*, pages 429–438, 2013. doi: 10.1109/FOCS.2013.53. URL https://doi.org/10.1109/FOCS.2013.53.

Cynthia Dwork. Differential privacy: A survey of results. In Manindra Agrawal, Ding-Zhu Du, Zhenhua Duan, and Angsheng Li, editors, *Theory and Applications of Models of Computation, TAMC*, volume 4978 of *Lecture Notes in Computer Science*, pages 1–19, 2008. doi: 10.1007/978-3-540-79228-4_1. URL https://doi.org/10.1007/978-3-540-79228-4_1.

Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014. doi: 10.1561/0400000042. URL https://doi.org/10.1561/0400000042.

Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology, EUROCRYPT*, volume 4004 of *Lecture Notes in Computer Science*, pages 486–503, 2006a. doi: 10.1007/11761679_29. URL https://doi.org/10.1007/11761679_29.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam D. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography, TCC*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284, 2006b. doi: 10.1007/11681878_14. URL https://doi.org/10.1007/11681878_14.

Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. RAPPOR: randomized aggregatable privacy-preserving ordinal response. In *Conference on Computer and Communications Security*, pages 1054–1067, 2014. doi: 10.1145/2660267.2660348. URL https://doi.org/10.1145/2660267.2660348.

Úlfar Erlingsson, Vitaly Feldman, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Abhradeep Thakurta. Amplification by shuffling: From local to central differential privacy via anonymity. In *Symposium on Discrete Algorithms, SODA*, pages 2468–2479. SIAM, 2019. doi: 10.1137/1.9781611975482.151. URL https://doi.org/10.1137/1.9781611975482.151.

Quan Geng and Pramod Viswanath. The optimal noise-adding mechanism in differential privacy. *IEEE Trans. Inf. Theory*, 62(2):925–951, 2016a. doi: 10.1109/TIT.2015.2504967. URL https://doi.org/10.1109/TIT.2015.2504967.

Quan Geng and Pramod Viswanath. Optimal noise adding mechanisms for approximate differential privacy. *IEEE Trans. Inf. Theory*, 62(2):952–969, 2016b. doi: 10.1109/TIT.2015.2504972. URL https://doi.org/10.1109/TIT.2015.2504972.

Quan Geng, Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The staircase mechanism in differential privacy. *IEEE J. Sel. Top. Signal Process.*, 9(7):1176–1184, 2015. doi: 10.1109/JSTSP.2015.2425831. URL https://doi.org/10.1109/JSTSP.2015.2425831.

Badih Ghazi, Noah Golowich, Ravi Kumar, Pasin Manurangsi, Rasmus Pagh, and Ameya Velingker. Pure differentially private summation from anonymous messages. In *Information-Theoretic Cryptography, ITC*, volume 163 of *LIPIcs*, pages 15:1–15:23, 2020a. doi: 10.4230/LIPIcs.ITC.2020. 15. URL https://doi.org/10.4230/LIPIcs.ITC.2020.15.

Badih Ghazi, Pasin Manurangsi, Rasmus Pagh, and Ameya Velingker. Private aggregation from fewer anonymous messages. In *Advances in Cryptology - EUROCRYPT*, volume 12106 of *Lecture Notes in Computer Science*, pages 798–827, 2020b. doi: 10.1007/978-3-030-45724-2_27. URL https://doi.org/10.1007/978-3-030-45724-2_27.

Badih Ghazi, Ravi Kumar, Pasin Manurangsi, Rasmus Pagh, and Amer Sinha. Differentially private aggregation in the shuffle model: Almost central accuracy in almost a single message. In *International Conference on Machine Learning, ICML*, 2021.

Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. *SIAM J. Comput.*, 41(6):1673–1693, 2012. doi: 10.1137/09076828X. URL https://doi.org/10.1137/09076828X.

Slawomir Goryczka and Li Xiong. A comprehensive comparison of multiparty secure additions with differential privacy. *Trans. Dependable Secur. Comput.*, 14(5):463–477, 2017. doi: 10.1109/TDSC.2015.2484326. URL https://doi.org/10.1109/TDSC.2015.2484326.

Andy Greenberg. How one of apple's key privacy safeguards falls short. *Wired*, 9 2017.

Mangesh Gupte and Mukund Sundararajan. Universally optimal privacy mechanisms for minimax agents. In Jan Paredaens and Dirk Van Gucht, editors, *Symposium on Principles of Database Systems, PODS*, pages 135–146, 2010. doi: 10.1145/1807085.1807105. URL https://doi.org/10.1145/1807085.1807105.

Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Symposium on Theory of Computing, STOC*, pages 705–714. ACM, 2010. doi: 10.1145/1806689.1806786. URL https://doi.org/10.1145/1806689.1806786.

Peter Kairouz, Ziyu Liu, and Thomas Steinke. The distributed discrete gaussian mechanism for federated learning with secure aggregation. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pages 5201–5212. PMLR, 2021. URL http://proceedings.mlr.press/v139/kairouz21a.html.

Shiva Prasad Kasiviswanathan, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam D. Smith. What can we learn privately? *SIAM J. Comput.*, 40(3):793–826, 2011. doi: 10.1137/090756090. URL https://doi.org/10.1137/090756090.

Fragkiskos Koufogiannis, Shuo Han, and George J. Pappas. Optimality of the laplace mechanism in differential privacy. *CoRR*, abs/1504.00065, 2015. URL http://arxiv.org/abs/1504.00065.

AM Mathai. On noncentral generalized laplacianness of quadratic forms in normal variables. *Journal of multivariate analysis*, 45(2):239–246, 1993.

Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. Communication-efficient learning of deep networks from decentralized data. In Aarti Singh and Xiaojin (Jerry) Zhu, editors, *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 54 of *Proceedings of Machine Learning Research*, pages 1273–1282. PMLR, 2017. URL http://proceedings.mlr.press/v54/mcmahan17a.html.

Aleksandar Nikolov, Kunal Talwar, and Li Zhang. The geometry of differential privacy: the sparse and approximate cases. In *Symposium on Theory of Computing Conference, STOC*, 2013. doi: 10.1145/2488608.2488652. URL https://doi.org/10.1145/2488608.2488652.

Elaine Shi, T.-H. Hubert Chan, Eleanor Gilbert Rieffel, Richard Chow, and Dawn Song. Privacy-preserving aggregation of time-series data. In *Distributed System Security Symposium, NDSS*, 2011. URL https://www.ndss-symposium.org/ndss2011/privacy-preserving-aggregation-of-time-series-data.

Salil P. Vadhan. The complexity of differential privacy. In *Tutorials on the Foundations of Cryptography*, pages 347–450. Springer International Publishing, 2017. doi: 10.1007/978-3-319-57048-8_7. URL https://doi.org/10.1007/978-3-319-57048-8_7.

Sameer Wagh, Xi He, Ashwin Machanavajjhala, and Prateek Mittal. DP-cryptography: marrying differential privacy and cryptography in emerging applications. *Commun. ACM*, 64(2):84–93, 2021. doi: 10.1145/3418290. URL https://doi.org/10.1145/3418290.

Stanley L Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

Wikipedia. Arete. https://en.wikipedia.org/wiki/Arete.

Andrew Chi-Chih Yao. Protocols for secure computations (extended abstract). In *Symposium on Foundations of Computer Science, FOCS*, pages 160–164, 1982. doi: 10.1109/SFCS.1982.38. URL https://doi.org/10.1109/SFCS.1982.38.

Appendix A. Applications

In differential privacy, two models are prevalent: the central model and the local model. In the *central* model of differential privacy (Dwork et al., 2006a) all data is held by a single trusted unit who makes the result of a query differentially private before releasing it. This is often done by adding noise to the query result. The central model usually has a very high level of accuracy, but requires a high level of trust. Often, data is split among many players, we refer to these as data owners, and a trusted central unit is not available. This setting is commonly known as the *local* model of differential privacy (Kasiviswanathan et al., 2011; Warner, 1965; Duchi et al., 2013). In this model, each data owner must ensure privacy for their own data, and so applies a differentially private mechanism locally, which is then forwarded to an analyst who combines all reports to compute an approximate answer to the query. For many queries, the overall error in the local model grows rather quickly as a function of the number of players, significantly limiting utility. For example, while we can achieve constant error in the central model (Dwork et al., 2006b), a count query requires $O(\sqrt{n})$ error for

the same level of privacy as in the central model, where n is the number of players (Cheu et al., 2019). The local model is often attractive for data collection as the collecting organizations are not liable for storing sensitive user data in this model – a few examples of deployment include Google's RAPPOR (Erlingsson et al., 2014), Apple (several features such as Lookup Hints, Emoji suggestion etc.) (Apple) and Microsoft Telemetry (Ding et al., 2017).

In order to bridge this trust/utility gap, we may imitate the trusted unit from the centralized setting with cryptographic primitives (Wagh et al., 2021), allowing for differentially private implementations with better utility than in the local model while having lower trust assumptions than in the centralized model. Cryptographic primitives ensure that all parties learn only the output of the computation, while differential privacy further bounds the information leakage from this output, and so the combination gives very strong guarantees. We limit our discussion to the problem of computing the sum of real inputs, which is a basic building block in many other applications. If we can divide the noise among all players we can obtain the same accuracy in a distributed setting as in the central model without the assumption of a trusted aggregator. Luckily, we can divide the noise between the players if the noise distribution $\mathcal{D}$ is infinitely divisible, and so the Arete distribution can be applied in this model.

We discuss differential privacy implementations with two cryptographic primitives: Secure Multiparty Aggregation and Anonymous Communication but note that such implementations come with assumptions about the computational power of the analyst, which are accepted by the security community, but limit the privacy guarantee to computational differential privacy (Wagh et al., 2021).

SECURE MULTIPARTY AGGREGATION

The cryptographic primitive secure multiparty Aggregation, rooted in the work of Yao (Yao, 1982), has often been combined with differential privacy to solve the problem of private real summation, see for example (Shi et al., 2011; Bonawitz et al., 2017; Chan et al., 2012). Goryczka and Xiong (2017) give a comparative study of several protocols for private summation in a distributed setting. These protocols combine common approaches for achieving security (secret sharing, homomorphic encryption and perturbation-based) while each party adds noise shares whose sum follows the Laplace distribution before sharing their data, in order to ensure differential privacy. Continuing their line of work, we may exchange the Laplace noise in (Goryczka and Xiong, 2017) with Arete distributed noise to achieve ε-differentially private protocols with error exponentially small in ε.

ANONYMOUS COMMUNICATION

Another line of work that has received a lot of attention over the past few years is the *shuffle* model of differential privacy (Bittau et al., 2017; Cheu et al., 2019; Erlingsson et al., 2019). Along with Google's Prochlo framework, Bittau et al. (2017) introduced the ESA (Encode Shuffle Analyze) framework where each data owner encodes their data before releasing it to a shuffler. The shuffler randomly permutes the encoded inputs and releases the (private) permuted set of data to an (untrusted) analyst, who then performs statistical analysis on the encoded, shuffled data. For recent work on the problem of summation in the shuffle model and a discussion of error/privacy-tradeoff, we refer to for example (Balle et al., 2019, 2020; Ghazi et al., 2020a, 2021, 2020b). Ghazi et al. (2021) propose an (ε, δ)-differentially private protocol for summation in the shuffle model for summing reals or integers where each user sends expected $1+o(1)$ messages. The protocol adds discrete Laplace noise (also sometimes called Geometrically distributed noise) and achieves error arbitrarily

close to that of the Laplace mechanism (applied in the central model), but do not address the problem of achieving error exponentially decreasing in ε in the shuffle model. The Arete distribution solves this open problem: as it is infinitely divisible, simply exchange the discrete Laplace noise (the "central" noise distribution in the protocol) with the Arete distribution. This yields:

Corollary 23 (Differentially private aggregation in the shuffle model.) *Let n be a positive integer, and let ε, δ be positive real numbers with $\varepsilon = O(\ln n)$. There is an (ε, δ)-differentially private aggregation protocol in the shuffle model for inputs in $[0, 1]$ having absolute error $\frac{1}{e^{\Omega(\varepsilon)} - 1}$ in expectation, using $O\left(1 + \frac{\log(1/\delta)}{\log n}\right)$ messages per party, each consisting of $O(\log n)$ bits.*

Note. By the post-processing property of differential privacy, we still achieve differential privacy if more than n players participate, and so we only need to choose the noise shares based on a lower bound on the number of players in order to ensure differential privacy. Hence, it is not strictly necessary to know the exact number of players in advance.

Appendix B. Basic definitions

B.1. Probability Distributions

In this section we state the definitions and basic facts that we need to analyze the Arete distribution. References to further information can be found in (Goryczka and Xiong, 2017).

Definition 24 (Infinite Divisibility) *A distribution $\mathcal{D}$ is infinitely divisible if, for any random variable X with distribution $\mathcal{D}$, then for every positive integer n there exist n i.i.d. random variables $X_1, ..., X_n$ such that $\sum_{i=1}^{n} X_i$ has the same distribution as X. The random variables X_i need not have distribution $\mathcal{D}$.*

We recall the definitions of the distributions that we use to define the Arete distribution and give a formal definition of the latter. Whenever the parameters are implicit we leave them out and simply write f_Γ, f_L, $f_{\Gamma - \Gamma}$ and f_A for the densities of the Γ, Laplace, $\Gamma - \Gamma$ and Arete distributions, resp.

Definition 25 (The Γ Distribution) *A random variable X has Gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\theta > 0$, denoted $X \sim \Gamma(\alpha, \theta)$, if its density function is*

$$f_{\Gamma(\alpha,\theta)}(t) = \frac{e^{-t/\theta} t^{\alpha - 1}}{\Gamma(\alpha)\theta^\alpha}, \quad t > 0.$$

In the special case $\alpha = 1$, the random variable X has Exponential distribution with parameter θ.

The Γ-distribution is infinitely divisible: For n independent random variables $X_i \sim \Gamma(\alpha_i, \theta)$, we have $X = \sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} \alpha_i, \theta\right)$. Furthermore, for $X \sim \Gamma(\alpha, \theta)$ we have $\mathrm{E}[X] = \alpha\theta$ and $\mathrm{Var}[X] = \alpha\theta^2$.

Definition 26 (The Laplace Distribution) *A random variable X has Laplace distribution with location parameter μ and scale parameter $\lambda > 0$, denoted $X \sim Laplace(\mu, \lambda)$, if its density function is*

$$f_{L(\mu,\lambda)}(t) = \frac{e^{-|t-\mu|/\lambda}}{2\lambda}, \quad t \in \mathbb{R}.$$

If $\mu = 0$ we just write $Laplace(\lambda)$.

If $X \sim Laplace(\lambda)$, then $|X| \sim Exp(\lambda)$ and $\mathrm{E}[X] = 0$ while $\mathrm{E}[|X|] = \lambda$. Similarly, $\mathrm{Var}[X] = 2\lambda^2$ while $\mathrm{Var}[|X|] = \lambda^2$.

The Laplace distribution is infinitely divisible: For $2n$ independent random variables $X_i, Y_i \sim \Gamma(1/n, \lambda)$, we have $X = \sum_{i=1}^{n}(\mu/n + X_i - Y_i) \sim Laplace\,(\mu, \lambda)$.

B.2. Differential Privacy

Informally, differential privacy promises that an analyst cannot, given a query answer, decide whether the underlying data contains a specific data record or not, and so differential privacy relies on the notion of *neighboring inputs*: datasets $x, y \in \mathcal{X}^d$ are neighbors, denoted $x \doteq y$, if they differ by one data record. The *sensitivity* of a query quantifies how much the output of the query can differ for neighboring inputs, and so describes how much difference the added noise needs to hide.

Definition 27 (Sensitivity (Dwork et al., 2006b)) *For a real-valued query $q : \mathcal{X}^d \to \mathbb{R}$, the sensitivity of q is defined as*
$\max_{x,y\in\mathcal{X}^d:\ x\doteq y} |q(x) - q(y)|.$

Definition 28 (Differential Privacy (Dwork et al., 2006b,a)) *Let $\mathcal{M}$ be a randomized mechanism. For privacy parameters $\varepsilon, \delta \geq 0$, we say that $\mathcal{M}$ is (ε, δ)-differentially private if, for any neighboring inputs $x, y \in \mathcal{X}^d$ and all $S \in Range(\mathcal{M})$ we have*

$$\Pr[\mathcal{M}(x) \in S] \leq e^\varepsilon \Pr[\mathcal{M}(y) \in S] + \delta.$$

If $\delta = 0$ we say that $\mathcal{M}$ is ε-differentially private.

For more details about differential privacy, we refer the reader to (Dwork, 2008; Dwork and Roth, 2014; Vadhan, 2017).

Lemma 29 (The Laplace Mechanism (Dwork et al., 2006b)) *For real-valued query $q : \mathcal{X}^d \to \mathbb{R}$ and input $x \in \mathcal{X}^d$, the Laplace mechanism outputs $q(x) + X$ where $X \sim Lap(\lambda)$. If Δ is the sensitivity of q, the Laplace mechanism with parameter $\lambda = \Delta/\varepsilon$ is ε-differentially private.*

Lemma 30 (The Staircase Mechanism (Geng and Viswanath, 2016a)) *Let $q : \mathbb{R} \to \mathbb{R}$ be a real-valued query with sensitivity Δ. Let random variable $X \sim SC(\gamma, \Delta)$ have Staircase distribution with parameters $\gamma \in [0, 1]$ and $\Delta > 0$ such that the density of X is*

$$f_{SC}(t) = \begin{cases} a(\gamma), & t \in [0, \gamma\Delta) \\ e^{-\varepsilon}a(\gamma), & t \in [\gamma\Delta, \Delta) \\ e^{-k\varepsilon}f_{SC}(t - k\Delta), & t \in [k\Delta, (k+1)\Delta),\ k \in \mathbb{N} \\ f_{SC}(-t), & t < 0 \end{cases}$$

where $a(\gamma) = \frac{1-e^{-\varepsilon}}{2\Delta(\gamma+e^{-\varepsilon}(1-\gamma))}$ is a normalization factor. Then for input $x \in \mathbb{R}$, the Staircase mechanism which outputs $q(x) + X$ where $X \sim SC(\gamma, \Delta)$ is ε-differentially private.

For optimal parameter γ, the Staircase mechanism achieves expected absolute error $\Theta(\Delta e^{-\varepsilon/2})$ and variance $\Theta(\Delta^2 e^{-2\varepsilon/3})$. We remark that the γ optimizing for expected magnitude of the noise is not the same as the γ optimizing for variance.

Appendix C. Omitted Proofs for Technical Results

Supporting lemmas

Lemma 31 *Let ζ_Γ be as in Lemma 16 and $\Delta > 0$. Assume $0 < \alpha < 1$, $1/\lambda - 1/\theta \geq \frac{1}{\kappa + \Delta + \zeta_\Gamma}$. Then $\forall t \geq \kappa \geq 0$*

$$\frac{(\zeta_\Gamma + \Delta + t)^{1-\alpha}}{e^{t(1/\lambda - 1/\theta)}} \leq \frac{(\zeta_\Gamma + \Delta + \kappa)^{1-\alpha}}{e^{\kappa(1/\lambda - 1/\theta)}}.$$

Proof The function

$$g(t) = \frac{(\zeta_\Gamma + \Delta + t)^{1-\alpha}}{e^{t(1/\lambda - 1/\theta)}}$$

maximized for

$$t^* = \frac{1 - \alpha - (1/\lambda - 1/\theta)(\zeta_\Gamma + \Delta)}{1/\lambda - 1/\theta} = \frac{1-\alpha}{1/\lambda - 1/\theta} - (\zeta_\Gamma + \Delta), \qquad 1/\lambda - 1/\theta > 0, \quad 0 < \alpha < 1,$$

and monotonely decreasing for $t \geq t^*$. By assumption

$$\kappa \geq \frac{1}{1/\lambda - 1/\theta} - (\Delta + \zeta_\Gamma) \geq t^*.$$

and so $g(\kappa) \leq g(t^*)$. Furthermore, for all $t \geq \kappa$

$$g(t) = \frac{(\zeta_\Gamma + \Delta + t)^{1-\alpha}}{e^{t(1/\lambda - 1/\theta)}} \leq \frac{(\zeta_\Gamma + \Delta + \kappa)^{1-\alpha}}{e^{\kappa(1/\lambda - 1/\theta)}} = g(\kappa).$$

∎

C.1. Proof of Lemma 18

Lemma 32 *Let $\zeta_u > 0$ be given and assume $0 < \alpha \leq 1$. Let $\zeta_\Gamma, c_{\zeta_\Gamma}$ be as in Lemma 16 and c_L as in Lemma 17. Assume $1/\lambda - 1/\theta \geq 1/(\zeta_u + \Delta + \zeta_\Gamma)$ and $\lambda \leq \Delta/\ln(2)$ for $\Delta > 0$. For $-\Delta/2 < t \in \mathbb{R}$ with $|t| \geq \zeta_u$ we have*

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{c_{\zeta_\Gamma}} \left(\left(1 + \frac{\Delta + \zeta_\Gamma}{\zeta_u}\right) + \frac{c_{\zeta_u}\Gamma(\alpha)\theta^\alpha}{2\lambda c_L} e^{\zeta_u/\theta}(\zeta_\Gamma + \Delta + \zeta_u)^{1-\alpha} \right)$$

where $c_{\zeta_u} := 2\int_0^{\zeta_u} f_\Gamma(x)dx$.

Proof [Proof of Lemma 18] Suppose $\zeta_u \leq |t|$. By Lemma 16

$$\frac{f_A(t)}{f_A(t + \Delta)} = \frac{\int_{-\infty}^\infty f_{\Gamma-\Gamma}(t-x)f_L(x)dx}{\int_{-\infty}^\infty f_{\Gamma-\Gamma}(t+\Delta-x)f_L(x)dx} \leq \frac{\int_{-\infty}^\infty f_\Gamma(|t-x|)f_L(x)dx}{\int_{-\infty}^\infty f_\Gamma(|t+\Delta-x| + \zeta_\Gamma)c_{\zeta_\Gamma}f_L(x)dx}. \quad (4)$$

Note that $|t + \Delta - x| \leq |t - x| + \Delta$ and

$$|t-x|^{\alpha-1} = (|t-x| + \zeta_\Gamma + \Delta)^{\alpha-1}\left(1 + \frac{\Delta + \zeta_\Gamma}{|t-x|}\right)^{1-\alpha}. \quad (5)$$

22

So filling in the density f_Γ and applying (5), we can write (4) as

$$\frac{f_A(t)}{f_A(t+\Delta)} \leq \frac{\int_{-\infty}^{\infty} \frac{1}{\Gamma(\alpha)\theta^\alpha} e^{-|t-x|/\theta}(|t-x|+\zeta_\Gamma+\Delta)^{\alpha-1}\left(1+\frac{\Delta+\zeta_\Gamma}{|t-x|}\right)^{1-\alpha} f_L(x)dx}{c_{\zeta_\Gamma}\int_{-\infty}^{\infty} f_\Gamma(|t-x|+\Delta+\zeta_\Gamma)f_L(x)dx}.$$

Since $\left(1+\frac{\Delta+\zeta_\Gamma}{|t-x|}\right)^{1-\alpha}$ is maximized for $x \to t$, we can bound this term as long as x is not too close to t. Hence, rewind to equation (4) and treat the cases where x is far from t and x is close to t separately by splitting the numerator from (4) at the intervals $x \in (-\infty, t-\zeta_u) \cup (t+\zeta_u, \infty)$ and $x \in (t-\zeta_u, t+\zeta_u)$ (these intervals are well-defined since $\zeta_u > 0$):

$$\frac{f_A(t)}{f_A(t+\Delta)} \leq \frac{\int_{-\infty}^{t-\zeta_u} f_\Gamma(|t-x|)f_L(x)dx + \int_{t+\zeta_u}^{\infty} f_\Gamma(|t-x|)f_L(x)dx}{c_{\zeta_\Gamma}\int_{-\infty}^{\infty} f_\Gamma(|t-x|+\Delta+\zeta_\Gamma)f_L(x)dx}$$

$$+ \frac{\int_{t-\zeta_u}^{t+\zeta_u} f_\Gamma(|t-x|)f_L(x)dx}{c_{\zeta_\Gamma}\int_{-\infty}^{\infty} f_\Gamma(|t+\Delta-x|+\zeta_\Gamma)f_L(x)dx}$$

$$= \frac{\int_{-\infty}^{t-\zeta_u} e^{-|t-x|/\theta}(|t-x|+\zeta_\Gamma+\Delta)^{\alpha-1}\left(1+\frac{\Delta+\zeta_\Gamma}{|t-x|}\right)^{1-\alpha} f_L(x)dx}{c_{\zeta_\Gamma}\int_{-\infty}^{\infty} e^{-(|t-x|+\Delta+\zeta_\Gamma)/\theta}(|t-x|+\Delta+\zeta_\Gamma)^{\alpha-1}f_L(x)dx}$$

$$+ \frac{\int_{t+\zeta_u}^{\infty} e^{-|t-x|/\theta}(|t-x|+\zeta_\Gamma+\Delta)^{\alpha-1}\left(1+\frac{\Delta+\zeta_\Gamma}{|t-x|}\right)^{1-\alpha} f_L(x)dx}{c_{\zeta_\Gamma}\int_{-\infty}^{\infty} e^{-(|t-x|+\Delta+\zeta_\Gamma)/\theta}(|t-x|+\Delta+\zeta_\Gamma)^{\alpha-1}f_L(x)dx}$$

$$+ \frac{\int_{t-\zeta_u}^{t+\zeta_u} f_\Gamma(|t-x|)f_L(x)dx}{c_{\zeta_\Gamma}\int_{-\infty}^{\infty} f_\Gamma(|t+\Delta-x|+\zeta_\Gamma)f_L(x)dx}$$

where we in the last step again filled in the density function f_Γ and applied (5) in the first two terms and left the last term as it was. Note that the constant $\Gamma(\alpha)\theta^\alpha$ from the density f_Γ cancels out in the fraction.

Now (still leaving the last term alone), for the first two terms upper bound the factor

$$\left(1+\frac{\Delta+\zeta_\Gamma}{|t-x|}\right)^{1-\alpha} \leq \left(1+\frac{\Delta+\zeta_\Gamma}{\zeta_u}\right)^{1-\alpha}$$

and pull out $e^{(\Delta+\zeta_\Gamma)/\theta}$ from the denominator, to see that

$$\frac{f_A(t)}{f_A(t+\Delta)}$$

$$\leq e^{(\Delta+\zeta_\Gamma)/\theta}\left(1+\frac{\Delta+\zeta_\Gamma}{\zeta_u}\right)^{1-\alpha} \frac{\left(\int_{-\infty}^{t-\zeta_u} \frac{(|t-x|+\zeta_\Gamma+\Delta)^{\alpha-1}}{e^{|t-x|/\theta}} f_L(x)dx + \int_{t+\zeta_u}^{\infty} \frac{(|t-x|+\zeta_\Gamma+\Delta)^{\alpha-1}}{e^{|t-x|/\theta}} f_L(x)dx\right)}{c_{\zeta_\Gamma}\int_{-\infty}^{\infty} \frac{(|t-x|+\zeta_\Gamma+\Delta)^{\alpha-1}}{e^{|t-x|/\theta}} f_L(x)dx}$$

$$+ \frac{\int_{t-\zeta_u}^{t+\zeta_u} f_\Gamma(|t-x|)f_L(x)dx}{c_{\zeta_\Gamma}\int_{-\infty}^{\infty} f_\Gamma(|t+\Delta-x|+\zeta_\Gamma)f_L(x)dx}$$

$$\leq \frac{1}{c_{\zeta_\Gamma}}\left(e^{(\Delta+\zeta_\Gamma)/\theta}\left(1+\frac{\Delta+\zeta_\Gamma}{\zeta_u}\right) + \frac{\int_{t-\zeta_u}^{t+\zeta_u} f_\Gamma(|t-x|)f_L(x)dx}{\int_{-\infty}^{\infty} f_\Gamma(|t+\Delta-x|+\zeta_\Gamma)f_L(x)dx}\right)$$

23

where in the last step we upper bounded the fraction in the first term by $1/c_{\zeta_\Gamma}$ and removed the $(1-\alpha)$-exponent for simpler notation. The following claim handles the last term and finishes the proof:

Claim 1 *Let $\zeta_\mathrm{u} > 0$ be given. Let $\zeta_\Gamma, c_{\zeta_\Gamma}$ be as in Lemma 16 and c_L as in Lemma 17. Assume $|t| \geq \zeta_\mathrm{u}$, $|t| \leq |t + \Delta|$, $\lambda \leq \Delta/\ln(2)$ and $1/\lambda - 1/\theta \geq 1/(\zeta_\Gamma + \Delta + \zeta_\mathrm{u})$. Then*

$$\frac{\int_{t-\zeta_\mathrm{u}}^{t+\zeta_\mathrm{u}} f_\Gamma(|t-x|)f_L(x)dx}{\int_{-\infty}^{\infty} f_\Gamma(\zeta_\Gamma + |t+1-x|)f_L(x)dx} \leq \frac{c_{\zeta_\mathrm{u}}\Gamma(\alpha)\theta^\alpha}{2\lambda c_L} e^{(\zeta_\Gamma + \zeta_\mathrm{u} + \Delta)/\theta}(\zeta_\Gamma + \Delta + \zeta_\mathrm{u})^{1-\alpha}$$

where

$$c_{\zeta_\mathrm{u}} := 2\int_0^{\zeta_\mathrm{u}} f_\Gamma(x)dx.$$

∎

Proof [Proof of Claim] As $|t| > \zeta_\mathrm{u}$, $\{0\} \notin (t - \zeta_\mathrm{u}, t + \zeta_\mathrm{u})$ and so $f_L(x)$ is maximal at

$$x = \min\{|t - \zeta_\mathrm{u}|, |t + \zeta_\mathrm{u}|\} = \min\{|t| + \zeta_\mathrm{u}, |t| - \zeta_\mathrm{u}\} = |t| - \zeta_\mathrm{u}.$$

Hence

$$\int_{t-\zeta_\mathrm{u}}^{t+\zeta_\mathrm{u}} f_\Gamma(|t-x|)f_L(x)dx \leq f_L(|t| - \zeta_\mathrm{u})\int_{t-\zeta_\mathrm{u}}^{t+\zeta_\mathrm{u}} f_\Gamma(|t-x|)dx = f_L(|t| - \zeta_\mathrm{u})c_{\zeta_\mathrm{u}} \qquad (6)$$

where we defined

$$c_{\zeta_\mathrm{u}} := \int_{t-\zeta_\mathrm{u}}^{t+\zeta_\mathrm{u}} f_\Gamma(|t-x|)dx = \int_{-\zeta_\mathrm{u}}^{\zeta_\mathrm{u}} f_\Gamma(|x|)dx = 2\int_0^{\zeta_\mathrm{u}} f_\Gamma(x)dx.$$

Now, consider

$$\frac{\int_{t-\zeta_\mathrm{u}}^{t+\zeta_\mathrm{u}} f_\Gamma(|t-x|)f_L(x)dx}{\int_{-\infty}^{\infty} f_\Gamma(\zeta_\Gamma + |t+\Delta-x|)f_L(x)dx}.$$

Recalling the assumptions $|t| \leq |t + \Delta|$ and $\lambda \leq \Delta/\ln(2)$, apply Lemma 17 in the denominator and (6) in the numerator, we get

$$\frac{\int_{t-\zeta_\mathrm{u}}^{t+\zeta_\mathrm{u}} f_\Gamma(|t-x|)f_L(x)dx}{\int_{-\infty}^{\infty} f_\Gamma(\zeta_\Gamma + |t+\Delta-x|)f_L(x)dx} \leq \frac{f_L(|t| - \zeta_\mathrm{u})c_{\zeta_\mathrm{u}}}{f_\Gamma(\zeta_\Gamma + |t+\Delta|)c_L} \overset{(*)}{\leq} \frac{c_{\zeta_\mathrm{u}} e^{-||t|-\zeta_\mathrm{u}|/\lambda}\Gamma(\alpha)\theta^\alpha}{2\lambda c_L e^{-(\zeta_\Gamma + \Delta + |t|)/\theta}(\zeta_\Gamma + \Delta + |t|)^{\alpha-1}}$$

$$\leq \frac{c_{\zeta_\mathrm{u}}\Gamma(\alpha)\theta^\alpha}{2\lambda c_L} e^{\zeta_\mathrm{u}/\lambda} e^{(\zeta_\Gamma + \Delta)/\theta}\frac{(\zeta_\Gamma + \Delta + |t|)^{1-\alpha}}{e^{|t|(1/\lambda - 1/\theta)}}$$

where we at $(*)$ filled in the density functions f_Γ and f_L and used that $|t + \Delta| \leq |t| + \Delta$. In the last step, recall $|t| > \zeta_\mathrm{u}$, so $|t| - \zeta_\mathrm{u} > 0$. Applying Lemma 31 (recall $0 < \alpha < 1$ and the assumption $1/\lambda - 1/\theta \geq 1/(\Delta + \zeta_\Gamma + \zeta_\mathrm{u})$) with $\kappa = \zeta_\mathrm{u}$, we get

$$\frac{\int_{t-\zeta_\mathrm{u}}^{t+\zeta_\mathrm{u}} f_\Gamma(|t-x|)f_L(x)dx}{\int_{-\infty}^{\infty} f_\Gamma(\zeta_\Gamma + |t+\Delta-x|)f_L(x)dx} \leq \frac{c_{\zeta_\mathrm{u}}\Gamma(\alpha)\theta^\alpha}{2\lambda c_L} e^{\zeta_\mathrm{u}/\lambda} e^{(\zeta_\Gamma + \Delta)/\theta}\frac{(\zeta_\Gamma + \Delta + \zeta_\mathrm{u})^{1-\alpha}}{e^{\zeta_\mathrm{u}(1/\lambda - 1/\theta)}}$$

$$= \frac{c_{\zeta_\mathrm{u}}\Gamma(\alpha)\theta^\alpha}{2\lambda c_L} e^{(\zeta_\Gamma + \zeta_\mathrm{u} + \Delta)/\theta}(\zeta_\Gamma + \Delta + \zeta_\mathrm{u})^{1-\alpha}.$$

∎

C.2. Proof of Lemma 19

Lemma 33 *Let $\zeta_u > 0$ be given and assume $0 < \alpha \le 1$. Let $\zeta_\Gamma, c_{\zeta_\Gamma}$ be as in Lemma 16 and c_L as in Lemma 17. Assume $1/\lambda - 1/\theta \ge 1/(\Delta + \zeta_\Gamma)$ and $\lambda \le \Delta/\ln(2)$ for $\Delta > 0$. For $-\Delta/2 < t \in \mathbb{R}$ such that $|t| \le \zeta_u$ we have*

$$\frac{f_A(t)}{f_A(t+\Delta)} \le \frac{e^{(\zeta_\Gamma+\Delta)/\theta}}{c_L c_{\zeta_\Gamma}\lambda}\left(\frac{\zeta_u+\zeta_\Gamma+\Delta}{\alpha} + \Gamma(\alpha)\theta^\alpha(\zeta_\Gamma+\Delta)^{1-\alpha}\right)$$

Proof [Proof of Lemma 19] Suppose $|t| < \zeta_u$. By Lemmas 11 and 16 we have

$$f_A(t) = f_A(|t|) = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(x)f_L(|t|-x)dx \le \int_{-\infty}^\infty f_\Gamma(|x|)f_L(|t|-x)dx.$$

Note that $f_L(|t|-x) \le f_L(t)$ when $||t|-x| \ge |t|$, which is satisfied whenever $x \notin (0, 2|t|)$.

$$f_A(t) = f_A(|t|) = \int_{-\infty}^\infty f_{\Gamma-\Gamma}(x)f_L(|t|-x)dx \le \int_{-\infty}^\infty f_\Gamma(|x|)f_L(|t|-x)dx$$

$$= \int_{-\infty}^0 f_\Gamma(|x|)f_L(|t|-x)dx + \int_0^{2|t|} f_\Gamma(|x|)f_L(|t|-x)dx + \int_{2|t|}^\infty f_\Gamma(|x|)f_L(|t|-x)dx$$

$$\le f_L(t)\left(\int_{-\infty}^0 f_\Gamma(|x|)dx + \int_{2|t|}^\infty f_\Gamma(|x|)dx\right) + \int_0^{2|t|} f_\Gamma(|x|)f_L(|t|-x)dx$$

$$\overset{(*)}{\le} 2f_L(t) + 2\int_0^{|t|} f_\Gamma(|x|)f_L(|t|-x)dx$$

$$= 2f_L(t) + \frac{2}{\Gamma(\alpha)\theta^\alpha 2\lambda}\int_0^{|t|} e^{-x/\theta}x^{\alpha-1}e^{-||t|-x|/\lambda}dx$$

$$= 2f_L(t) + \frac{1}{\Gamma(\alpha)\theta^\alpha\lambda}e^{-|t|/\lambda}\int_0^{|t|} e^{x(1/\lambda-1/\theta)}x^{\alpha-1}dx$$

$$\le 2f_L(t) + \frac{1}{\Gamma(\alpha)\theta^\alpha\lambda}e^{-|t|/\theta}\int_0^{|t|} x^{\alpha-1}dx$$

$$= 2f_L(t) + \frac{1}{\Gamma(\alpha)\theta^\alpha\lambda}e^{-|t|/\theta}\frac{|t|^\alpha}{\alpha}$$

At $(*)$ we use that $f_\Gamma(|x|)$ is smaller on $(t, 2t)$ than on $(0, t)$. In the last step we used that

$$\int_0^\kappa x^n dx = \frac{\kappa^{n+1}}{n+1}, \qquad n \neq -1.$$

Recalling the assumptions $|t| \le |t + \Delta|$ and $\lambda \le \Delta/\ln(2)$, apply Lemma 17 to get

$$
\begin{aligned}
\frac{f_A(t)}{f_A(t+\Delta)} &\le \frac{2f_L(t) + \frac{1}{\Gamma(\alpha)\theta^\alpha \lambda}e^{-|t|/\theta}\frac{|t|^\alpha}{\alpha}}{c_L c_{\zeta_\Gamma} f_\Gamma(\zeta_\Gamma + |t+\Delta|)} = \frac{1}{c_L c_{\zeta_\Gamma}}\left(\frac{2f_L(t)}{f_\Gamma(\zeta_\Gamma + |t+\Delta|)} + \frac{1}{\Gamma(\alpha)\theta^\alpha \lambda}\frac{e^{-|t|/\theta}\frac{|t|^\alpha}{\alpha}}{f_\Gamma(\zeta_\Gamma + |t+\Delta|)} \right) \\
&\overset{(*)}{=} \frac{1}{c_L c_{\zeta_\Gamma}}\left(\frac{2e^{-|t|/\lambda}\Gamma(\alpha)\theta^\alpha}{2\lambda e^{-(\zeta_\Gamma + |t+\Delta|)/\theta}(\zeta_\Gamma + |t+\Delta|)^{\alpha-1}} + \frac{\Gamma(\alpha)\theta^\alpha}{\Gamma(\alpha)\theta^\alpha \lambda}\frac{e^{-|t|/\theta}\frac{|t|^\alpha}{\alpha}}{e^{-(\zeta_\Gamma + |t+\Delta|)/\theta}(\zeta_\Gamma + |t+\Delta|)^{\alpha-1}} \right) \\
&\overset{(**)}{\le} \frac{1}{\lambda c_L c_{\zeta_\Gamma}}\left(\frac{e^{-|t|/\lambda}\Gamma(\alpha)\theta^\alpha}{e^{-(\zeta_\Gamma + |t|+\Delta)/\theta}(\zeta_\Gamma + |t| + \Delta)^{\alpha-1}} + \frac{e^{-|t|/\theta}\frac{|t|^\alpha}{\alpha}}{e^{-(\zeta_\Gamma + |t|+\Delta)/\theta}(\zeta_\Gamma + |t| + \Delta)^{\alpha-1}} \right) \\
&= \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{\lambda c_L c_{\zeta_\Gamma}}\left(\frac{e^{-|t|/\lambda}\Gamma(\alpha)\theta^\alpha}{e^{-|t|/\theta}(\zeta_\Gamma + |t| + \Delta)^{\alpha-1}} + \frac{e^{-|t|/\theta}\frac{|t|^\alpha}{\alpha}}{e^{-|t|/\theta}(\zeta_\Gamma + |t| + \Delta)^{\alpha-1}} \right) \\
&= \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{\lambda c_L c_{\zeta_\Gamma}}\left(\Gamma(\alpha)\theta^\alpha \frac{(\zeta_\Gamma + |t| + \Delta)^{1-\alpha}}{e^{|t|(1/\lambda - 1/\theta)}} + (\zeta_\Gamma + |t| + \Delta)^{1-\alpha}\frac{|t|^\alpha}{\alpha} \right).
\end{aligned}
$$

where we at $(*)$ filled in the density functions and at $(**)$ used that $|t + \Delta| \le |t| + \Delta$.

Using the identity $|t|^\alpha = |t|/|t|^{1-\alpha}$ we see that

$$
\begin{aligned}
\frac{f_A(t)}{f_A(t+\Delta)} &\le \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{\lambda c_L c_{\zeta_\Gamma}}\left(\Gamma(\alpha)\theta^\alpha \frac{(\zeta_\Gamma + |t| + \Delta)^{1-\alpha}}{e^{|t|(1/\lambda - 1/\theta)}} + \frac{|t|}{\alpha}\left(\frac{\zeta_\Gamma + |t| + \Delta}{|t|} \right)^{1-\alpha} \right) \\
&\le \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{\lambda c_L c_{\zeta_\Gamma}}\left(\Gamma(\alpha)\theta^\alpha \frac{(\zeta_\Gamma + |t| + \Delta)^{1-\alpha}}{e^{|t|(1/\lambda - 1/\theta)}} + \frac{\zeta_\Gamma + |t| + \Delta}{\alpha} \right) \\
&\le \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{\lambda c_L c_{\zeta_\Gamma}}\left(\Gamma(\alpha)\theta^\alpha \frac{(\zeta_\Gamma + |t| + \Delta)^{1-\alpha}}{e^{|t|(1/\lambda - 1/\theta)}} + \frac{\zeta_\Gamma + \zeta_u + \Delta}{\alpha} \right),
\end{aligned}
$$

where in the last two steps, we removed the $(1-\alpha)$-exponent on the last term and used that $|t| \le \zeta_u$. Finally applying Lemma 31 (recall $0 < \alpha < 1$ and the assumption $1/\lambda - 1/\theta \ge 1/(\Delta + \zeta_\Gamma)$) with $\kappa = 0$ finishes the proof:

$$
\frac{f_A(t)}{f_A(t+\Delta)} \le \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{\lambda c_L c_{\zeta_\Gamma}}\left(\Gamma(\alpha)\theta^\alpha(\zeta_\Gamma + \Delta)^{1-\alpha} + \frac{\zeta_\Gamma + \zeta_u + \Delta}{\alpha} \right).
$$

∎

C.3. Proof of Lemma 20

Lemma 34 *Let $\zeta_\Gamma, c_{\zeta_\Gamma}$ be as in Lemma 16 and c_L as in Lemma 17. Assume $0 < \alpha \le 1$, $\theta \le \zeta_\Gamma + 1$, $\lambda \le \min\{\theta/2, \Delta/\ln(2)\}$ and $\Gamma(\alpha) \le 1/\alpha$ for $\Delta > 0$. For $-\Delta/2 < t \in \mathbb{R}$*

$$
\frac{f_A(t)}{f_A(t+\Delta)} \le \frac{2e^{(\zeta_\Gamma + \Delta)/\theta}e^\alpha(\alpha\theta + \zeta_\Gamma + \Delta)}{\alpha c_{\zeta_\Gamma} c_L \lambda}
$$

Proof We first give the intuition behind the proof: Lemmas 18 and 19 give upper bounds on the ratio for certain values of t, assuming $1/\lambda - 1/\theta \ge \max\{1/(\zeta_u + \zeta_\Gamma + \Delta), 1/(\zeta_\Gamma + \Delta)\} = 1/(\zeta_\Gamma + \Delta)$.

An upper bound on both of these bounds simultaneously gives us a bound on the ratio, which holds for general $t > -\Delta/2$. We note that

$$1/\lambda - 1/\theta \geq 1/(\zeta_\Gamma + \Delta) \quad \Leftrightarrow \quad \lambda \leq \frac{\theta}{\frac{\theta}{\zeta_\Gamma + \Delta} + 1},$$

so if $\theta \leq \zeta_\Gamma + \Delta$, then $\lambda \leq \theta/2$ suffices. Hence, our assumptions $\lambda \leq \theta/2$, $\lambda \leq \Delta/\ln(2)$, $\theta \leq \zeta_\Gamma + \Delta$ and $t > -\Delta/2$ ensure that we can use Lemmas 18 and 19.

So, by Lemmas 18 and 19 we have for $-\Delta/2 < t \in \mathbb{R}$

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq \frac{e^{(\zeta_\Gamma + \Delta)/\theta}(\zeta_u + \zeta_\Gamma + \Delta)}{c_{\zeta_\Gamma}} \max\left\{\frac{1}{\alpha c_L \lambda}, \frac{1}{\zeta_u}\right\}$$

$$+ \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{c_{\zeta_\Gamma}} \frac{\Gamma(\alpha)\theta^\alpha}{\lambda c_L} \max\left\{(\zeta_\Gamma + \Delta)^{1-\alpha}, \frac{c_{\zeta_u} e^{\zeta_u/\theta}}{2}(\zeta_\Gamma + \Delta + \zeta_u)^{1-\alpha}\right\}.$$

As, by assumption, $\theta \leq \zeta_\Gamma + \Delta$, we see

$$\theta^\alpha(\zeta_\Gamma + \Delta)^{1-\alpha} < \theta^\alpha(\zeta_u + \zeta_\Gamma + \Delta)^{1-\alpha} < \zeta_u + \zeta_\Gamma + \Delta$$

and so we may simplify to

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq \frac{e^{(\zeta_\Gamma + \Delta)/\theta}}{c_{\zeta_\Gamma}} \left(\max\left\{\frac{1}{\alpha c_L \lambda}, \frac{1}{\zeta_u}\right\}(\zeta_u + \zeta_\Gamma + \Delta) + \frac{\Gamma(\alpha)(\zeta_\Gamma + \Delta + \zeta_u)}{\lambda c_L} \max\left\{1, \frac{c_{\zeta_u} e^{\zeta_u/\theta}}{2}\right\}\right)$$

$$= \frac{e^{(\zeta_\Gamma + \Delta)/\theta}(\zeta_u + \zeta_\Gamma + \Delta)}{c_{\zeta_\Gamma}} \left(\max\left\{\frac{1}{\alpha c_L \lambda}, \frac{1}{\zeta_u}\right\} + \frac{\Gamma(\alpha)}{\lambda c_L} \max\left\{1, \frac{c_{\zeta_u} e^{\zeta_u/\theta}}{2}\right\}\right).$$

Let $\zeta_u = \alpha\theta$ (i.e., the mean of the Γ-distribution). Recalling that by definition $c_L = 1/4$ and by assumption $\lambda \leq \theta/2$, so $\alpha c_L \lambda \leq \alpha\theta/8 < \alpha\theta = \zeta_u$:

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq \frac{e^{(\zeta_\Gamma + \Delta)/\theta}(\alpha\theta + \zeta_\Gamma + \Delta)}{c_{\zeta_\Gamma}} \left(\frac{1}{\alpha c_L \lambda} + \frac{\Gamma(\alpha)}{c_L \lambda} \max\left\{1, \frac{c_{\zeta_u} e^\alpha}{2}\right\}\right)$$

$$\leq \frac{e^{(\zeta_\Gamma + \Delta)/\theta}(\alpha\theta + \zeta_\Gamma + \Delta)}{c_{\zeta_\Gamma} c_L \lambda} \left(\frac{1}{\alpha} + \Gamma(\alpha)e^\alpha\right),$$

where the last step follows from the observation that $1 \leq c_{\zeta_u} \leq 2$ (recall c_{ζ_u} was defined in Lemma 18) and $e^\alpha > 1$ for $\alpha > 0$.

By assumption $\Gamma(\alpha) < 1/\alpha$, then

$$1/\alpha + \Gamma(\alpha)e^\alpha \leq \frac{2e^\alpha}{\alpha}$$

and so we conclude

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq \frac{2e^{(\zeta_\Gamma + \Delta)/\theta} e^\alpha(\alpha\theta + \zeta_\Gamma + \Delta)}{\alpha c_{\zeta_\Gamma} c_L \lambda}.$$

∎

C.4. Proof of Lemma 21

Lemma 35 *Suppose $\varepsilon \geq 20 + 4\ln(\Delta)$ for $\Delta \geq 2/e$. Let $\alpha = e^{-\varepsilon/4}, \theta = \frac{4\Delta}{\varepsilon}$ and $\lambda = e^{-\varepsilon/4}$. Then for $t \in \mathbb{R}$*

$$e^{-\varepsilon} \leq \frac{f_A(t)}{f_A(t + \Delta)} \leq e^{\varepsilon}.$$

Proof Suppose $|t| \leq |t + \Delta|$. The first inequality is satisfied as $f_A(t) \geq f_A(t + \Delta)$. Let ζ_Γ be as in Lemma 16. We turn to prove the latter inequality: In order to apply Lemma 20 we make the following assumptions:

$$\theta \leq \zeta_\Gamma + \Delta, \quad \lambda \leq \theta/2, \quad \lambda \leq \Delta/\ln(2) \quad \text{and} \quad \Gamma(\alpha) \leq 1/\alpha. \tag{7}$$

We choose parameters satisfying these assumptions towards the end of the proof.

If ζ_Γ is at least the median of the Γ-distribution then $c_{\zeta_\Gamma} \geq 1/2$. So let $\zeta_\Gamma = \alpha\theta$ be the mean of the Γ-distribution (the mean is an upper bound on the median of the Γ-distribution (Chen and Rubin, 1986)), to see

$$\frac{f_A(t)}{f_A(t + \Delta)} \overset{\text{(Lemma 20)}}{\leq} \frac{2e^{(\Delta + \zeta_\Gamma)/\theta}e^\alpha(\alpha\theta + \zeta_\Gamma + \Delta)}{\alpha c_{\zeta_\Gamma} c_L \lambda} \leq \frac{2 \cdot (2\alpha\theta + \Delta)}{1/2 \cdot 1/4} \frac{e^{\Delta/\theta}e^{2\alpha}}{\alpha\lambda}. \tag{8}$$

Suppose $\alpha \leq 1/2$ and recall by assumption $\theta \leq \zeta_\Gamma + \Delta = \alpha\theta + \Delta$, so $\theta \leq \Delta/(1 - \alpha)$. Then $\theta \leq \Delta/(1 - \alpha) \leq \Delta/\alpha$ and so $\alpha\theta \leq \Delta$. We revise our set of assumptions, to also ensure that $\alpha\theta \leq \Delta$, and so our set of assumptions is:

$$\theta \leq \frac{\Delta}{1 - \alpha}, \quad \lambda \leq \min\{\theta/2, \Delta/\ln(2)\}, \quad \alpha \leq 1/2 \quad \text{and} \quad \Gamma(\alpha) \leq 1/\alpha. \tag{9}$$

Under these assumptions we have $2\alpha\theta + \Delta \leq 3\Delta$ and inserting into (8), we conclude

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq 48\Delta\frac{e^{\Delta/\theta}e^{2\alpha}}{\alpha\lambda}.$$

Now define

$$\alpha = 1/e^{\varepsilon/k_\alpha}, \quad \theta = \frac{k_\theta\Delta}{\varepsilon} \quad \text{and} \quad \lambda = 1/e^{\varepsilon/k_\lambda}$$

Observing $2\alpha \leq 1$ and $\ln(48e) < 4.9$

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq 48e^{2\alpha}\Delta e^{\varepsilon(1/k_\theta + 1/k_\alpha + 1/k_\lambda)} < e^{\varepsilon(1/k_\theta + 1/k_\alpha + 1/k_\lambda) + 4.9 + \ln(\Delta)}.$$

Hence, we ensure that

$$\frac{f_A(t)}{f_A(t + \Delta)} \leq e^{\varepsilon}$$

when the assumptions in (9) are satisfied and

$$\varepsilon(1/k_\theta + 1/k_\alpha + 1/k_\lambda) + 5 + \ln(\Delta) \leq \varepsilon \quad \Leftrightarrow \quad 1/k_\theta + 1/k_\alpha + 1/k_\lambda \leq 1 - \frac{5 + \ln(\Delta)}{\varepsilon}. \tag{10}$$

Let $k_\alpha = k_\lambda = k_\theta = 4$. It is easy to check that the assumptions on θ and α in (9) are satisfied simultaneously for, $\varepsilon \geq 4\ln(2)$ (and we can check that $\Gamma(\alpha) \leq 1/\alpha$ numerically). Furthermore, for $\varepsilon \geq 4\ln(2)$, we require $\lambda \leq \Delta \min\{2/\varepsilon, 1/\ln(2)\} = 2\Delta/\varepsilon$ and so the assumption on λ is satisfied when $\Delta \geq \varepsilon\lambda/2 = \varepsilon e^{-\varepsilon/4}/2$. Observing that $\Delta \geq 2/e \geq \varepsilon e^{-\varepsilon/4}/2$, we conclude that the assumptions in (9) are satisfied for $\varepsilon \geq 4\ln(2)$ and $\Delta \geq 2/e$. The inequality in (10) is satisfied for

$$3/4 \leq 1 - \frac{5 + \ln(\Delta)}{\varepsilon} \quad \Leftrightarrow \quad 20 + 4\ln(\Delta) \leq \varepsilon.$$

Finally, observe that if $|t| \geq |t + \Delta|$, the result follows by symmetry of f_A (Corollary 13). ∎