Compressed Counting *

Ping Li†

Abstract

We propose *Compressed Counting (CC)* for approximating the αth frequency moments ($0 < \alpha \leq 2$) of data streams under a *relaxed strict-Turnstile* model, using *maximally-skewed stable random projections*. Estimators based on the *geometric mean* and the *harmonic mean* are developed.

When $\alpha = 1$, a simple counter suffices for counting the first moment (i.e., sum). The *geometric mean* estimator of CC has asymptotic variance $\propto \Delta = |\alpha - 1|$, capturing the intuition that the complexity should decrease as $\Delta = |\alpha - 1| \to 0$. However, the previous classical algorithms based on *symmetric stable random projections*[12, 15] required $O\left(1/\epsilon^2\right)$ space, in order to approximate the αth moments within a $1+\epsilon$ factor, for any $0 < \alpha \leq 2$ including $\alpha = 1$.

We show that using the *geometric mean* estimator, CC requires $O\left(\frac{1}{\log(1+\epsilon)} + \frac{2\sqrt{\Delta}}{\log^{3/2}(1+\epsilon)} + o\left(\sqrt{\Delta}\right)\right)$ space, as $\Delta \to 0$. Therefore, in the neighborhood of $\alpha = 1$, the complexity of CC is essentially $O(1/\epsilon)$ instead of $O\left(1/\epsilon^2\right)$.

CC may be useful for estimating Shannon entropy, which can be approximated by certain functions of the αth moments with $\alpha \to 1$. [10, 9] suggested using $\alpha = 1 + \Delta$ with (e.g.,) $\Delta < 0.0001$ and $\epsilon < 10^{-7}$, to rigorously ensure reasonable approximations. Thus, unfortunately, CC is "theoretically impractical" for estimating Shannon entropy, despite its empirical success reported in [16].

1 Introduction

Counting is a fundamental operation. Counting the sum $\sum_{i=1}^{D} A_t[i]$ is the simplest task (where t denotes time). Counting the αth moment $\sum_{i=1}^{D} A_t[i]^\alpha$ is more general. Here, A_t is a time-varying *data stream*[11, 12, 3, 18, 1].

1.1 The Relaxed Strict-Turnstile Model

This study considers a *relaxed strict-Turnstile* model. The input stream $a_t = (i_t, I_t)$, $i_t \in [1,\ D]$ arriving sequentially describes the underlying signal A, meaning

$$A_t[i_t] = A_{t-1}[i_t] + I_t,$$

where I_t can be positive (insertion) or negative (deletion). Restricting $A_t[i] \geq 0$ results in the *strict-Turnstile* model, which suffices for describing most natural phenomena.

For example, $A_{t-1}[i]$ may record the number of items that user i has ordered up to time $t-1$ and I_t denotes the additional orders ($I_t > 0$) or cancels ($I_t < 0$) at t. It is reasonable to assume that it is not possible to cancel orders that do not exist. In general, in a database[18], a record can only be deleted if it was previously inserted.

We consider the *relaxed strict-Turnstile* model, which constrains $A_t[i] \geq 0$ only at the t we care about. At $s \neq t$, we allow $A_s[i]$ to be arbitrary. Under this model, the αth frequency moment of a data stream A_t is defined as

$$F_{(\alpha)} = \sum_{i=1}^{D} A_t[i]^\alpha.$$

When $\alpha = 1$, it is obvious that one can compute $F_{(1)} = \sum_{i=1}^{D} A_t[i] = \sum_{s=1}^{t} I_s$ trivially, using a simple counter.

Counting $F_{(\alpha)}$ for massive data streams is practically important and challenging. Because the elements, $A_t[i]$, are dynamic, a naíve counting mechanism requires D counters to compute $F_{(\alpha)}$ exactly. This is not always realistic when D is large (e.g., $D = 2^{64}$) or when we need to compute $F_{(\alpha)}$ in real-time, e.g., in network measurement/monitoring[23].

Compressed Counting (CC) is the first proposal of using *(maximally) skewed stable random projections* for computing $F_{(\alpha)}$ with $0 < \alpha \leq 2$. The improvement of CC over previous studies is most significant when $\alpha \approx 1$.

1.2 Previous Work

Pioneered by[2], the task of approximating $F_{(\alpha)}$ has been heavily studied. [2] considered $\alpha = 0$, 2, and $\alpha > 2$. [6] provided an algorithm for $\alpha = 1$ and [12] proposed *symmetric stable random projections* for $0 < \alpha \leq 2$. [15] proposed various estimators and tail bounds (with constants explicitly given) for *symmetric stable random projections*, whose required space is $O\left(1/\epsilon^2\right)$ in order to approximate the αth moments within a $1 \pm \epsilon$ factor, for any $0 < \alpha \leq 2$. [7] proposed a different algorithm using space $O\left(1/\epsilon^{2+\alpha}\right)$ to trade for some speedup in the processing time.

[20, 4] proved the space lower bounds for $\alpha > 2$ and [13] provided algorithms for $\alpha > 2$ to achieve the lower bounds. [22] proved the lower bounds for all frequency moments, that any one-pass algorithm for approximating $F_{(\alpha)}$ required space $\Omega\left(1/\epsilon^2\right)$, except for $\alpha = 1$.

1.3 CC Breaks the $O\left(1/\epsilon^2\right)$ Barrier

Compressed Counting (CC) captures the intuition that, when $\alpha = 1$, a simple counter suffices for computing $F_{(1)}$,

*Supported by NSF grant DMS-0808864 and a gift from Google.

†Department of Statistical Science, Faculty of Computing and Information Science, Cornell University. Ithaca NY 14853, USA

and when $\alpha = 1 \pm \Delta$ with small Δ, the complexity should be low and vary continuously as a function of Δ. None of the previous studies, however, captured this intuition.

Roughly speaking, for a fixed (small) ϵ, as $\Delta = |\alpha - 1| \to 0$, the complexity of CC is $O\left(1/\epsilon\right)$, instead of $O\left(1/\epsilon^2\right)$. This result will be stated precisely in Theorem 4.1.

The basic tool for CC is *skewed stable distributions*.

1.4 Skewed Stable Distributions

A random variable Z follows a β-skewed α-stable distribution if the Fourier transform of its density is[24]

$$\mathscr{F}_Z(\theta) = \mathrm{E}\exp\left(\sqrt{-1}Z\theta\right)$$
$$= \exp\left(-F|\theta|^\alpha\left(1 - \sqrt{-1}\beta\mathrm{sign}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right)\right), \quad \alpha \neq 1$$

where $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$ and $F > 0$ is the scale parameter. We denote $Z \sim S(\alpha, \beta, F)$.

Consider two independent variables, $Z_1, Z_2 \sim S(\alpha, \beta, 1)$. For any non-negative constants C_1 and C_2, the "α-stability" follows from properties of Fourier transforms:

$$Z = C_1 Z_1 + C_2 Z_2 \sim S\left(\alpha, \beta, C_1^\alpha + C_2^\alpha\right).$$

However, if C_1 and C_2 do not have the same signs, the above "stability" does not hold (unless $\beta = 0$ or $\alpha = 2, 0+$). To see this, we consider $Z = C_1 Z_1 - C_2 Z_2$, with $C_1 \geq 0$ and $C_2 \geq 0$. Then, because $\mathscr{F}_{-Z_2}(\theta) = \mathscr{F}_{Z_2}(-\theta)$,

$$\mathscr{F}_Z = \exp\left(-|C_1\theta|^\alpha\left(1 - \sqrt{-1}\beta\mathrm{sign}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right)\right) \times$$
$$\exp\left(-|C_2\theta|^\alpha\left(1 + \sqrt{-1}\beta\mathrm{sign}(\theta)\tan\left(\frac{\pi\alpha}{2}\right)\right)\right),$$

which does not represent a stable law, unless $\beta = 0$ or $\alpha = 2, 0+$. This is the fundamental reason why *Compressed Counting* needs the restriction that at the time t of the evaluation, stream elements should have the same signs.

1.5 Skewed Stable Random Projections

First, generate a vector $R \in \mathbb{R}^D$, whose entries are i.i.d. samples of a stable distribution: $r_i \sim S(\alpha, \beta, 1)$. Then

$$R^{\mathrm{T}} A_t = \sum_{i=1}^{D} r_i A_t[i] \sim S\left(\alpha, \beta, F_{(\alpha)} = \sum_{i=1}^{D} A_t[i]^\alpha\right),$$

meaning $R^{\mathrm{T}} A_t$ represents one sample of the stable distribution whose scale parameter $F_{(\alpha)}$ is what we are after.

If we generate a matrix $\mathbf{R} \in \mathbb{R}^{D \times k}$ with each entry $r_{ij} \sim S(\alpha, \beta, 1)$ i.i.d., the resultant vector $X = \mathbf{R}^{\mathrm{T}} A_t \in \mathbb{R}^k$ contains k i.i.d. samples: $x_j \sim S\left(\alpha, \beta, F_{(\alpha)}\right)$, $j = 1$ to k. We will explain why we recommend $\beta = 1$ (***maximally-skewed***).

Since it is a linear projection, this method is naturally applicable to data streams under the *Turnstile* model (which is also linear), by conducting the matrix-vector multiplication incrementally[12]. That is, for every incoming $a_t = (i_t, I_t)$, we update $x_j \leftarrow x_j + r_{i_t j} I_t$ for $j = 1$ to k, where random numbers $r_{i_t j}$'s are generated *on-demand*.

1.6 Statistical Estimators for Compressed Counting

Compressed Counting (CC) boils down to a statistical estimation problem. This study provides estimators based on the *geometric mean* and the *harmonic mean*.

In terms of the asymptotic variances, Figure 1 compares our two proposed estimators for CC with the *geometric mean* estimator for *symmetric stable distributions[15]*, demonstrating a huge improvement, especially around $\alpha = 1$.

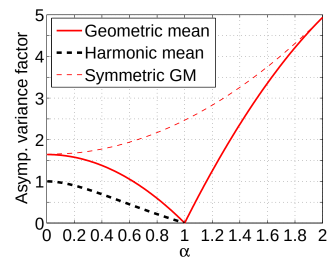


Figure 1: Let $\hat{F}_{(\alpha)}$ be an estimator of $F_{(\alpha)}$ with asymptotic variance $\mathrm{Var}\left(\hat{F}_{(\alpha)}\right) = V\frac{F_{(\alpha)}^2}{k} + O\left(\frac{1}{k^2}\right)$. We plot the V values for the proposed *geometric mean* and the *harmonic mean* estimators, along with the V values for the geometric mean estimator in [15] (symmetric GM).

1.7 Paper Organization

Section 2 describes applications of *Compressed Counting*. Section 3 derives the basic moment formulas for general skewed stable distributions, needed for designing and analyzing the *geometric mean* estimator and the *harmonic mean* estimator. Section 4 focuses on the *geometric mean* estimators. In particular, we show that the complexity bound is essentially $O\left(1/\epsilon\right)$ in the neighborhood of $\alpha = 1$. Section 5 analyzes the *harmonic mean* estimator. Section 6 describes the procedure of sampling from skewed stable distributions.

2 Three Types of Applications of Compressed Counting

2.1 Computing Basic Summary Statistics

The frequency moment $F_{(\alpha)}$ is a very basic summary statistic of the signal A_t. In certain applications, the parameter $\alpha = 1 \pm \Delta$ may bear a physical meaning. For example, the finance department may need to predict future earnings and hence it is useful not only to count the current sum $F_{(1)} = \sum_{i=1}^D A_t$ but also the future sum $F_{(\alpha)} = \sum_{i=1}^D A_t[i]^{1\pm\Delta}$, where Δ may be interpreted as the growth (or interest) rate.

Some useful summary statistics are functions of $F_{(\alpha)}$. For example, Rényi entropy[19], R_α, and Tsallis entropy[21], T_α, are defined as, respectively

$$R_\alpha = \frac{1}{1-\alpha}\log\frac{F_{(\alpha)}}{F_{(1)}^\alpha}, \qquad T_\alpha = \frac{1}{\alpha - 1}\left(1 - \frac{F_{(\alpha)}}{F_{(1)}^\alpha}\right), \tag{2.1}$$

which generalize and approach (as $\alpha \to 1$) Shannon entropy

$$H = -\sum_{i=1}^{D}\frac{A_t[i]}{F_{(1)}}\log\frac{A_t[i]}{F_{(1)}}. \tag{2.2}$$

2.2 Statistical Modeling and Inference

A basic task is to model the distribution of A_t. For example, a three-parameter generalized gamma distribution $GG(\theta_1, \theta_2, \theta_3)$ is flexible for modeling positive data[17]. If $A_t[i] \sim GG(\theta_1, \theta_2, \theta_3)$, then $\mathrm{E}(A_t[i]) = \theta_1\theta_2$, $\mathrm{Var}(A_t[i]) = \theta_1\theta_2^2$, $\mathrm{E}\left(A_t[i] - \mathrm{E}(A_t[i])\right)^3 = (\theta_3+1)\theta_1\theta_2^3$. Thus, one can estimate θ_1, θ_2 and θ_3 by counting the first three moments ($\sum_{i=1} A_t[i]^\alpha$, $\alpha = 1, 2, 3$) and solving equations. However, since some moments may be (much) easier to compute than others, it may be reasonable to estimate the parameters using three fractional moments (e.g., $\alpha = 1.0$, and close to 1.0).

2.3 Basic Building Element for Other Algorithms

One example[23, 10, 9] is to estimate Shannon entropy using Rényi entropy or Tsallis entropy. [10, 9] suggested using $\alpha = 1+\Delta$ with (e.g.,) $\Delta < 0.0001$ and (e.g.,) $\epsilon < 10^{-7}$ for reasonable approximations of Shannon entropy. (Those numbers can be verified in [10, 9]). Thus, CC is "theoretically impractical" for this task:

- CC using the *geometric mean* estimator has complexity $O(1/\epsilon)$ around $\alpha = 1$, which, when $\epsilon < 10^{-7}$, may correspond to $> 10^7$ samples, too large to be practical.
- When estimating Shannon entropy using Rényi or Tsallis entropy, the estimation variance blows up like $\frac{1}{(\alpha-1)^2}\mathrm{Var}\left(\hat{F}_{(\alpha)}\right)$ for an estimator $\hat{F}_{(\alpha)}$. As we will show, the *geometric mean* estimator has variance $\propto |\alpha - 1|$, which clearly does not decrease fast enough.

However, it is evident that CC is "practically practical" for estimating Shannon entropy. A recent work[16] adopted "bias-variance trade-off" by using α not too close to 1, to dramatically reduce the required number of samples (especially for *symmetric stable random projections*). It is, however, merely a statistical trick, not a rigorous theoretical result as [10, 9]. There is opportunity to improve CC, in order to estimate Shannon entropy using the criteria in [10, 9].

3 Moments of Skewed Stable Distributions

Recall, *Compressed Counting (CC)* boils down to estimating the scale parameter $F_{(\alpha)}$, from k i.i.d. samples of a β-skewed α-stable random variable, $x_j \sim S\left(\alpha, \beta, F_{(\alpha)}\right)$, $j = 1$ to k. Recall $\beta = 0$ corresponds to *symmetric stable distributions*.

There is a closed-form moment formula for $\mathrm{E}\left(|x_j|^\lambda\right)$. The proposed *geometric mean* estimator and *harmonic mean* estimator are based on positive moments ($\lambda > 0$) and negative moments ($\lambda < 0$), respectively.

Lemma 3.1 shows that in the strip $-1 < \lambda < \alpha$, the λth moment of $|Z|$ is bounded. The restriction $\lambda < \alpha$ is due to the heavy-tailed nature of $Z \sim S\left(\alpha, \beta, F_{(\alpha)}\right)$. The restriction $\lambda > -1$ is needed for Fubini's Theorem.

LEMMA 3.1. *If $Z \sim S(\alpha, \beta, F_{(\alpha)})$, then for any $-1 < \lambda < \alpha$,*

$$\boldsymbol{E}\left(|Z|^\lambda\right) = F_{(\alpha)}^{\lambda/\alpha}\cos\left(\frac{\lambda}{\alpha}\tan^{-1}\left(\beta\tan\left(\frac{\alpha\pi}{2}\right)\right)\right)$$
$$\times\left(1+\beta^2\tan^2\left(\frac{\alpha\pi}{2}\right)\right)^{\frac{\lambda}{2\alpha}}\left(\frac{2}{\pi}\sin\left(\frac{\pi}{2}\lambda\right)\Gamma\left(1-\frac{\lambda}{\alpha}\right)\Gamma(\lambda)\right),$$

which can be simplified when $\beta = 1$, to be

$$\boldsymbol{E}\left(|Z|^\lambda\right) = F_{(\alpha)}^{\lambda/\alpha}\frac{\cos\left(\frac{\kappa(\alpha)}{\alpha}\frac{\lambda\pi}{2}\right)}{\cos^{\lambda/\alpha}\left(\frac{\kappa(\alpha)\pi}{2}\right)}\left(\frac{2}{\pi}\sin\left(\frac{\pi}{2}\lambda\right)\Gamma\left(1-\frac{\lambda}{\alpha}\right)\Gamma(\lambda)\right),$$

$\kappa(\alpha) = \alpha$ *if* $\alpha < 1$, *and* $\kappa(\alpha) = 2-\alpha$ *if* $\alpha > 1$.

Proof: *See Appendix A. Here $\Gamma(.)$ is the gamma function.* □

Lemma 3.2 presents a (seemingly) surprising result that when $\beta = 1$ and $\alpha < 1$, all negative moments are bounded, i.e., estimators based on negative moments (when $\beta = 1$ and $\alpha < 1$) will have bounded moment generating functions.

LEMMA 3.2. *For $\alpha < 1$, $\beta = 1$, and $-\infty < \lambda < \alpha$,*

$$\boldsymbol{E}\left(|Z|^\lambda\right) = F_{(\alpha)}^{\lambda/\alpha}\frac{\Gamma\left(1-\frac{\lambda}{\alpha}\right)}{\cos^{\lambda/\alpha}\left(\frac{\alpha\pi}{2}\right)\Gamma(1-\lambda)}.$$

Proof: *See Appendix B.* □

4 The Geometric Mean Estimator

Although we recommend $\beta = 1$ (**maximally-skewed**), we start with the *geometric mean* estimator for general β and show that $\beta = 1$ achieves the smallest variance. Due to the symmetry, we only have to consider $\beta \in [0, 1]$.

4.1 The Geometric Mean Estimator for General β

Setting $\lambda = \frac{\alpha}{k}$ in Lemma 3.1 yields an unbiased estimator:

$$\hat{F}_{(\alpha),gm,\beta} = \frac{\prod_{j=1}^k |x_j|^{\alpha/k}}{D_{gm,\beta}} \qquad (k \geq 2)$$
$$D_{gm,\beta} = \cos^k\left(\frac{1}{k}\tan^{-1}\left(\beta\tan\left(\frac{\alpha\pi}{2}\right)\right)\right)\times$$
$$\left(1+\beta^2\tan^2\left(\frac{\alpha\pi}{2}\right)\right)^{\frac{1}{2}}\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}\right)\Gamma\left(1-\frac{1}{k}\right)\Gamma\left(\frac{\alpha}{k}\right)\right]^k.$$

Lemma 4.1 illustrates that the variance of $\hat{F}_{(\alpha),gm,\beta}$ decreases with increasing $\beta \in [0, 1]$.

LEMMA 4.1. *The variance of $\hat{F}_{(\alpha),gm,\beta}$ is a decreasing function of $\beta \in [0, 1]$, where*

$$\mathit{Var}\left(\hat{F}_{(\alpha),gm,\beta}\right) = F_{(\alpha)}^2\times$$
$$\left(\frac{\cos^k\left(\frac{2}{k}\tan^{-1}\left(\beta\tan\left(\frac{\alpha\pi}{2}\right)\right)\right)}{\cos^{2k}\left(\frac{1}{k}\tan^{-1}\left(\beta\tan\left(\frac{\alpha\pi}{2}\right)\right)\right)}\frac{\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{k}\right)\Gamma\left(1-\frac{2}{k}\right)\Gamma\left(\frac{2\alpha}{k}\right)\right]^k}{\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}\right)\Gamma\left(1-\frac{1}{k}\right)\Gamma\left(\frac{\alpha}{k}\right)\right]^{2k}}-1\right)$$

Proof: *The result follows from the fact that*

$$\frac{\cos\left(\frac{2}{k}\tan^{-1}\left(\beta\tan\left(\frac{\alpha\pi}{2}\right)\right)\right)}{\cos^2\left(\frac{1}{k}\tan^{-1}\left(\beta\tan\left(\frac{\alpha\pi}{2}\right)\right)\right)} = 2-\sec^2\left(\frac{1}{k}\tan^{-1}\left(\beta\tan\left(\frac{\alpha\pi}{2}\right)\right)\right),$$

is a decreasing function of $\beta \in [0, 1]$. □

Thus, the recommended *geometric mean* estimator is obtained by taking $\beta = 1$:

$$\hat{F}_{(\alpha),gm} = \frac{\cos\left(\frac{\kappa(\alpha)\pi}{2}\right)\prod_{j=1}^k |x_j|^{\alpha/k}}{\cos^k\left(\frac{\kappa(\alpha)\pi}{2k}\right)\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}\right)\Gamma\left(1-\frac{1}{k}\right)\Gamma\left(\frac{\alpha}{k}\right)\right]^k}, \tag{4.3}$$

$\kappa(\alpha) = \alpha$, if $\alpha < 1$, $\kappa(\alpha) = 2-\alpha$ if $\alpha > 1$.

For brevity, we simply use $\hat{F}_{(\alpha),gm}$ instead of $\hat{F}_{(\alpha),gm,1}$. In fact, the rest of the paper will always consider $\beta = 1$ only.

4.2 Moments of the Geometric Mean Estimator $\hat{F}_{(\alpha),gm}$

LEMMA 4.2. *As $k \to \infty$*

$$\left[\cos\left(\frac{\kappa(\alpha)\pi}{2k}\right)\frac{2}{\pi}\Gamma\left(\frac{\alpha}{k}\right)\Gamma\left(1-\frac{1}{k}\right)\sin\left(\frac{\pi}{2}\frac{\alpha}{k}\right)\right]^k \to \exp\left(-\gamma_e\left(\alpha-1\right)\right), \tag{4.4}$$

*decreasing **monotonically** with increasing k, where $\gamma_e = 0.57724...$ is Euler's constant. More precisely,*

$$\left[\cos\left(\frac{\kappa(\alpha)\pi}{2k}\right)\frac{2}{\pi}\Gamma\left(\frac{\alpha}{k}\right)\Gamma\left(1-\frac{1}{k}\right)\sin\left(\frac{\pi}{2}\frac{\alpha}{k}\right)\right]^k = \exp\left(-\gamma_e(\alpha-1)\right)$$
$$\times\exp\left(\frac{1}{k}\left(2+\alpha^2-3\kappa^2(\alpha)\right)\frac{\pi^2}{24}+\frac{1}{k^2}\frac{1-\alpha^3}{3}\zeta_3+...\right)$$

where $\zeta_3 = 1.2020569...$ is Apery's constant.
Proof: *See Appendix C.* $\Box$

LEMMA 4.3. *As $k \to \infty$, for any fixed $t \geq 1$,*

$$E\left(\left(\hat{F}_{(\alpha),gm}\right)^t\right) = F^t_{(\alpha)}\frac{\cos^k\left(\frac{\kappa(\alpha)\pi}{2k}t\right)\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}t\right)\Gamma\left(1-\frac{t}{k}\right)\Gamma\left(\frac{\alpha}{k}t\right)\right]^k}{\cos^{kt}\left(\frac{\kappa(\alpha)\pi}{2k}\right)\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}\right)\Gamma\left(1-\frac{1}{k}\right)\Gamma\left(\frac{\alpha}{k}\right)\right]^{kt}}$$
$$= F^t_{(\alpha)}\exp\left(\frac{1}{k}\frac{\pi^2(t^2-t)}{24}\left(2+\alpha^2-3\kappa^2(\alpha)\right)+\frac{1}{k^2}\frac{t^3-t}{3}(1-\alpha^3)\zeta_3+O\left(\frac{1}{k^3}\right)\right)$$

Proof: *See Appendix D.* $\Box$

LEMMA 4.4. *As $k \to \infty$,*

$$Var\left(\hat{F}_{(\alpha),gm}\right) = \frac{F^2_{(\alpha)}}{k}\frac{\pi^2}{6}\left(1-\alpha^2\right) \qquad (\text{if } \alpha < 1)$$
$$+\frac{F^2_{(\alpha)}}{k^2}\left(\frac{\pi^4}{72}(1-\alpha^2)^2+2(1-\alpha^3)\zeta_3\right)+O\left(\frac{1}{k^3}\right) \tag{4.5}$$

$$Var\left(\hat{F}_{(\alpha),gm}\right) = \frac{F^2_{(\alpha)}}{k}\frac{\pi^2}{6}(\alpha-1)(5-\alpha) \qquad (\text{if } \alpha > 1)$$
$$+\frac{F^2_{(\alpha)}}{k^2}\left(\frac{\pi^4}{72}(\alpha-1)^2(5-\alpha)^2+2(1-\alpha^3)\zeta_3\right)+O\left(\frac{1}{k^3}\right) \tag{4.6}$$

Proof: *A direct consequence of Lemma 4.3.* $\Box$

4.3 Tail Bounds

Tail bounds are crucial for providing a rigorous criterion on choosing k. We will derive tail bounds with all "constants" specified. In fact, as $\alpha \to 1$, those "constants" are so small that they should not be treated as constants any more.

The estimator $\hat{F}_{(\alpha),gm}$ is unbiased, which is nice; but there is a small price to pay. In (4.3), the denominator depends on k for small k, which complicates the analysis of tail bounds (especially the left tail bound). For convenience, we instead consider an asymptotically (as $k \to \infty$) equivalent (but slightly biased at small k) geometric mean estimator:

$$\hat{F}_{(\alpha),gm,b} = \exp\left(\gamma_e(\alpha-1)\right)\cos\left(\frac{\kappa(\alpha)\pi}{2}\right)\prod_{j=1}^k|x_j|^{\alpha/k}. \tag{4.7}$$

LEMMA 4.5. *As $k \to \infty$, when $\alpha < 1$,*

$$E\left(\hat{F}_{(\alpha),gm,b}\right)-F_{(\alpha)} = F_{(\alpha)}\frac{1}{k}\left(1-\alpha^2\right)\frac{\pi^2}{12}$$
$$+F_{(\alpha)}\frac{1}{k^2}\frac{1-\alpha^3}{3}\zeta_3+\frac{F_{(\alpha)}}{k^2}\left(1-\alpha^2\right)^2\frac{\pi^4}{288}+O\left(\frac{1}{k^3}\right)$$

$$Var\left(\hat{F}_{(\alpha),gm,b}\right) = \frac{F^2_{(\alpha)}}{k}\frac{\pi^2}{6}\left(1-\alpha^2\right)$$
$$+\frac{F^2_{(\alpha)}}{k^2}\left(\frac{\pi^4}{24}(1-\alpha^2)^2+2(1-\alpha^3)\zeta_3\right)+O\left(\frac{1}{k^3}\right)$$

And when $\alpha > 1$,

$$E\left(\hat{F}_{(\alpha),gm,b}\right)-F_{(\alpha)} = F_{(\alpha)}\frac{1}{k}\left(5-\alpha\right)\left(\alpha-1\right)\frac{\pi^2}{12}$$
$$+F_{(\alpha)}\frac{1}{k^2}\frac{1-\alpha^3}{3}\zeta_3+\frac{F_{(\alpha)}}{k^2}\left(\alpha-1\right)^2\left(5-\alpha\right)^2\frac{\pi^4}{288}+O\left(\frac{1}{k^3}\right)$$
$$Var\left(\hat{F}_{(\alpha),gm,b}\right) = \frac{F^2_{(\alpha)}}{k}\frac{\pi^2}{6}\left(\alpha-1\right)\left(5-\alpha\right)$$
$$+\frac{F^2_{(\alpha)}}{k^2}\left(\frac{\pi^4}{24}\left(\alpha-1\right)^2\left(5-\alpha\right)^2+2(1-\alpha^3)\zeta_3\right)+O\left(\frac{1}{k^3}\right)$$

Proof: *The proof follows from Lemmas 4.2 and 4.4.* $\Box$

We have carefully analyzed the moments of $\hat{F}_{(\alpha),gm}$ and $\hat{F}_{(\alpha),gm,b}$, to illustrate that two estimators are essentially no different, in case some readers have concerns about it.

LEMMA 4.6. *The right tail bound: for $\epsilon > 0$,*

$$\mathbf{Pr}\left(\hat{F}_{(\alpha),gm,b}-F_{(\alpha)} \geq \epsilon F_{(\alpha)}\right) \leq \exp\left(-k\frac{\epsilon^2}{G_{R,gm}}\right),$$

and left tail bound: for $0 < \epsilon < 1$,

$$\mathbf{Pr}\left(\hat{F}_{(\alpha),gm,b}-F_{(\alpha)} \leq -\epsilon F_{(\alpha)}\right) \leq \exp\left(-k\frac{\epsilon^2}{G_{L,gm}}\right),$$

where

$$\frac{\epsilon^2}{G_{R,gm}} = C_R\log(1+\epsilon)-C_R\gamma_e(\alpha-1)$$
$$-\log\left(\cos\left(\frac{\kappa(\alpha)\pi C_R}{2}\right)\frac{2}{\pi}\Gamma\left(\alpha C_R\right)\Gamma\left(1-C_R\right)\sin\left(\frac{\pi\alpha C_R}{2}\right)\right),$$
$$\frac{\epsilon^2}{G_{L,gm}} = -C_L\log(1-\epsilon)+C_L\gamma_e(\alpha-1)+\log\alpha$$
$$-\log\left(\cos\left(\frac{\kappa(\alpha)\pi}{2}C_L\right)\Gamma\left(C_L\right)\right)+\log\left(\Gamma\left(\alpha C_L\right)\cos\left(\frac{\pi\alpha C_L}{2}\right)\right).$$

C_R and C_L are solutions to

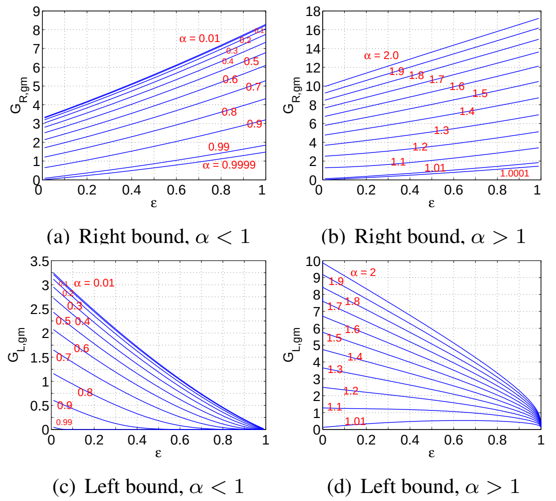
$$-\gamma_e(\alpha-1)+\log(1+\epsilon)+\frac{\kappa(\alpha)\pi}{2}\tan\left(\frac{\kappa(\alpha)\pi}{2}C_R\right)$$
$$-\frac{\alpha\pi/2}{\tan\left(\frac{\alpha\pi}{2}C_R\right)}-\psi\left(\alpha C_R\right)\alpha+\psi\left(1-C_R\right) = 0,$$
$$\log(1-\epsilon)-\gamma_e(\alpha-1)-\frac{\kappa(\alpha)\pi}{2}\tan\left(\frac{\kappa(\alpha)\pi}{2}C_L\right)$$
$$+\frac{\alpha\pi}{2}\tan\left(\frac{\alpha\pi}{2}C_L\right)-\psi\left(\alpha C_L\right)\alpha+\psi\left(C_L\right) = 0.$$

Here $\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ is the Psi function (digamma function).
Proof: *The proof is omitted.* $\Box$.

One can infer the right tail bound of $\hat{F}_{(\alpha),gm}$ from the right tail bound of $\hat{F}_{(\alpha),gm,b}$, because

$$\mathbf{Pr}\left(\hat{F}_{(\alpha),gm}-F_{(\alpha)} \geq \epsilon F_{(\alpha)}\right) \leq \mathbf{Pr}\left(\hat{F}_{(\alpha),gm,b}-F_{(\alpha)} \geq \epsilon F_{(\alpha)}\right)$$

holds due to the monotonicity result (4.4) in Lemma 4.3.



(a) Right bound, $\alpha < 1$ (b) Right bound, $\alpha > 1$

(c) Left bound, $\alpha < 1$ (d) Left bound, $\alpha > 1$

Figure 2: The tail bound constants $G_{R,gm}$ and $G_{L,gm}$ of the *geometric mean* estimator $\hat{F}_{(\alpha),gm,b}$ in Lemma 4.6.

4.4 Behavior of Tail Bounds as $\alpha \to 1$

Lemma 4.7 describes the precise rates of convergence, as $\alpha = 1 \pm \Delta \to 1$, of the constants derived in Lemma 4.6.

LEMMA 4.7. *For fixed ϵ, as $\alpha = 1 \pm \Delta \to 1$ (i.e., $\Delta \to 0$),*

$$G_{R,gm} = \frac{\epsilon^2}{\log(1+\epsilon) - 2\sqrt{\Delta \log(1+\epsilon)} + o\left(\sqrt{\Delta}\right)},$$

$$= \frac{\epsilon^2}{\log(1+\epsilon)} + \frac{2\epsilon^2}{\log^{3/2}(1+\epsilon)}\sqrt{\Delta} + o\left(\sqrt{\Delta}\right).$$

$$G_{L,gm} = \begin{cases} \frac{\epsilon^2}{-\log(1-\epsilon) - 2\sqrt{-2\Delta\log(1-\epsilon)} + o\left(\sqrt{\Delta}\right)}, & \alpha > 1 \\ \frac{\epsilon^2}{\Delta\left(\exp\left(\frac{-\log(1-\epsilon)}{\Delta} - 1 - \gamma_e\right)\right) + o\left(\Delta\exp\left(\frac{1}{\Delta}\right)\right)}, & \alpha < 1 \end{cases}$$

Proof: *See Appendix E. See Figure 3 for verification.* □

We usually consider small ϵ. Thus, roughly speaking, as $\alpha \to 1$, $G_{R,gm} = O(\epsilon)$ and $G_{L,gm} = O(\epsilon)$.

4.5 Sample Complexity Bound

The sample complexity bound follows by letting

$$\mathbf{Pr}\left(\hat{F}_{(\alpha),gm,b} - F_{(\alpha)} \geq \epsilon F_{(\alpha)}\right) \leq \exp\left(-k\frac{\epsilon^2}{G_{R,gm}}\right) \leq \delta.$$

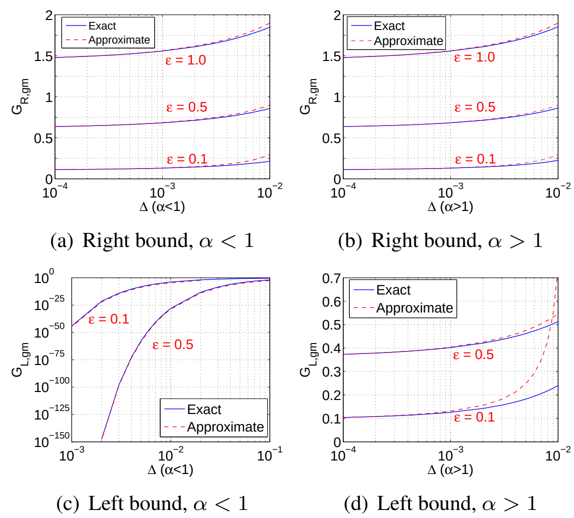
THEOREM 4.1. *Using the geometric mean estimator $\hat{F}_{(\alpha),gm,b}$, as $\Delta = |\alpha - 1| \to 0$, it suffices to let*

$$k = \left(\frac{1}{\log(1+\epsilon)} + \frac{2\sqrt{\Delta}}{\log^{3/2}(1+\epsilon)} + o\left(\sqrt{\Delta}\right)\right)\log\frac{1}{\delta}, \tag{4.8}$$

so that the estimate will be within a $1+\epsilon$ factor of the truth with probability $1-\delta$.

One can similarly write down the sample complexity bound for achieving an accuracy within a $1-\epsilon$ factor.

Using standard arguments, the space complexity in terms of the number of bits can be obtained by multiplying the above sample complexity bound with $\log M$, where M is the size of the "universe," or in this context $M = \sum_{s=1}^{t}|I_s|$.



(a) Right bound, $\alpha < 1$ (b) Right bound, $\alpha > 1$

(c) Left bound, $\alpha < 1$ (d) Left bound, $\alpha > 1$

Figure 3: The tail bound constants proved in Lemma 4.6, together with their approximations in Lemma 4.7, for small $\Delta = |\alpha - 1|$.

5 The Harmonic Mean Estimator

For $\alpha < 1$, the *harmonic mean* estimator can considerably improve $\hat{F}_{(\alpha),gm}$. This estimator takes advantage of the fact in Lemma 3.2 that if $Z \sim S(\alpha < 1, \beta = 1, F_{(\alpha)})$, then $\mathrm{E}\left(|Z|^\lambda\right) = \mathrm{E}\left(Z^\lambda\right)$ exists for $-\infty < \lambda < \alpha$.

LEMMA 5.1. *Assume k i.i.d. samples $x_j \sim S(\alpha < 1, \beta = 1, F_{(\alpha)})$, define the harmonic mean estimator $\hat{F}_{(\alpha),hm}$, and the bias-corrected harmonic mean estimator $\hat{F}_{(\alpha),hm,c}$:*

$$\hat{F}_{(\alpha),hm} = \frac{k\frac{\cos\left(\frac{\alpha\pi}{2}\right)}{\Gamma(1+\alpha)}}{\sum_{j=1}^{k}|x_j|^{-\alpha}},$$

$$\hat{F}_{(\alpha),hm,c} = \frac{k\frac{\cos\left(\frac{\alpha\pi}{2}\right)}{\Gamma(1+\alpha)}}{\sum_{j=1}^{k}|x_j|^{-\alpha}}\left(1 - \frac{1}{k}\left(\frac{2\Gamma^2(1+\alpha)}{\Gamma(1+2\alpha)} - 1\right)\right).$$

The bias and variance of $\hat{F}_{(\alpha),hm,c}$ are

$$E\left(\hat{F}_{(\alpha),hm,c}\right) = F_{(\alpha)} + O\left(\frac{1}{k^2}\right),$$

$$Var\left(\hat{F}_{(\alpha),hm,c}\right) = \frac{F_{(\alpha)}^2}{k}\left(\frac{2\Gamma^2(1+\alpha)}{\Gamma(1+2\alpha)} - 1\right) + O\left(\frac{1}{k^2}\right).$$

The right tail bound of $\hat{F}_{(\alpha),hm}$ is, for $\epsilon > 0$,

$$\mathbf{Pr}\left(\hat{F}_{(\alpha),hm} - F_{(\alpha)} \geq \epsilon F_{(\alpha)}\right) \leq \exp\left(-k\left(\frac{\epsilon^2}{G_{R,hm}}\right)\right),$$

$$\frac{\epsilon^2}{G_{R,hm}} = -\log\left(\sum_{m=0}^{\infty}\frac{\Gamma^m(1+\alpha)}{\Gamma(1+m\alpha)}(-t_1^*)^m\right) - \frac{t_1^*}{1+\epsilon},$$

where t_1^* is the solution to

$$\frac{\sum_{m=1}^{\infty}(-1)^m m(t_1^*)^{m-1}\frac{\Gamma^m(1+\alpha)}{\Gamma(1+m\alpha)}}{\sum_{m=0}^{\infty}(-1)^m(t_1^*)^m\frac{\Gamma^m(1+\alpha)}{\Gamma(1+m\alpha)}} + \frac{1}{1+\epsilon} = 0.$$

The left tail bound of $\hat{F}_{(\alpha),hm}$ is, for $0 < \epsilon < 1$,

$$\mathbf{Pr}\left(\hat{F}_{(\alpha),hm} - F_{(\alpha)} \leq -\epsilon F_{(\alpha)}\right) \leq \exp\left(-k\left(\frac{\epsilon^2}{G_{L,hm}}\right)\right),$$

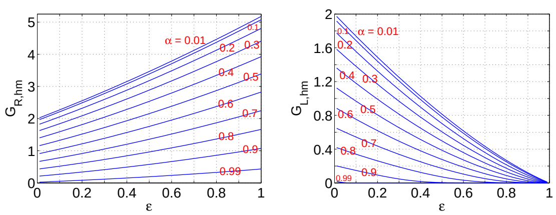
$$\frac{\epsilon^2}{G_{L,hm}} = -\log\left(\sum_{m=0}^{\infty}\frac{\Gamma^m(1+\alpha)}{\Gamma(1+m\alpha)}(t_2^*)^m\right) + \frac{t_2^*}{1-\epsilon}$$

where t_2^* is the solution to

$$-\frac{\sum_{m=1}^{\infty} m(t_2^*)^{m-1}\frac{\Gamma^m(1+\alpha)}{\Gamma(1+m\alpha)}}{\sum_{m=0}^{\infty}(t_2^*)^m\frac{\Gamma^m(1+\alpha)}{\Gamma(1+m\alpha)}} + \frac{1}{1-\epsilon} = 0$$

Proof: *The proof is omitted.* □.

The *harmonic mean* estimator has smaller variance than the *geometric mean* estimator (see Figure 1) and smaller tail bound constants (see Figure 4). However, we have not characterized the behavior of its tail bounds around $\alpha = 1$.



(a) Right tail bound constant (b) Left tail bound constant

Figure 4: The tail bound constants of $\hat{F}_{(\alpha),hm}$ in Lemma 5.1, which are considerably smaller compared to Figure 2(a)(c).

6 Sampling From Skewed Stable Distributions

Sampling from skewed stable distributions is based on the Chambers-Mallows-Stuck method[5]. (Note that [5] adopted a different parameterization.) One first generates an exponential random variable with mean 1, $W \sim \exp(1)$, and a uniform random variable $U \sim uniform\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Let $\rho = \tan^{-1}\left(\beta\tan\left(\frac{\pi\alpha}{2}\right)\right)/\alpha$. Then,

$$Z = \frac{\sin\left(\alpha(U+\rho)\right)}{\left[\cos U\cos\left(\rho\alpha\right)\right]^{\frac{1}{\alpha}}}\left[\frac{\cos\left(U - \alpha(U+\rho)\right)}{W}\right]^{\frac{1-\alpha}{\alpha}} \sim S(\alpha,\beta,1) \tag{6.9}$$

Note that when $\beta = 1$ and $\alpha \to 1$, $\rho \to \frac{\pi}{2}$, i.e., $\cos(\rho\alpha) \to 0$. One might worry about the numerical instability in computing (6.9), or equivalently, the potential problem of using large storage space in order to maintain the desired accuracy. This issue can be completely avoided.

When computing (6.9), we can ignore $\cos^{1/\alpha}(\rho\alpha)$. This is equivalent to sampling $Z' = Z\cos^{1/\alpha}(\rho\alpha) \sim S(\alpha,\beta,\cos(\rho\alpha))$ instead of $Z = S(\alpha,\beta,1)$. We can conduct projections as usual as long as we divide the estimates by $\cos(\rho\alpha)$. Note that the $\cos(\rho\alpha)$ term already exists in the two estimators we have studied. This is nice because we avoid the numerical issue by doing less work.

7 Conclusion

Compressed Counting (CC) is the first proposal of using *skewed stable random projections* for estimating the αth frequency moment $F_{(\alpha)} = \sum_{i=1}^{D} A_t[i]^\alpha$ of a streaming signal A_t, where $0 < \alpha \leq 2$. CC takes advantage of the fact that most data streams encountered in practice are non-negative, although they are subject to deletion and insertion. CC captures the intuition that, when $\alpha = 1$, a simple counter suffices, and when $\alpha = 1 \pm \Delta$ with small Δ, an intelligent counting system should require low space.

Two estimators based on the *geometric mean* and the *harmonic mean* are provided in this study. We show that, as $\Delta = |\alpha - 1| \to 0$, the complexity of CC (using the *geometric mean estimator*) is essentially $O(1/\epsilon)$, instead of the previously believed $O\left(1/\epsilon^2\right)$ bound.

At least three lines of research will benefit from CC. (1): $F_{(\alpha)}$ itself is a useful summary statistic and some important summary statistics (such as Rényi entropy and Tsallis entropy) are functions of $F_{(\alpha)}$. (2): CC will be useful for statistical modeling and parameter inference of data streams using the *method of moments*. (3): CC can be a basic building element for designing other algorithms, for example, estimating Shannon entropy of data streams using Rényi or Tsallis entropy with $\alpha \to 1$. Unfortunately, following the rigorous criteria in[10, 9], CC is still "theoretically impractical" for approximating Shannon entropy, despite its empirical success reported in [16].

Acknowledgement

The author highly appreciates three anonymous referees for their detailed reviews. Jon Kleinberg read the very first draft of this work and suggested several important references. Gennady Samorodnitsky provided the technical report[14], which is otherwise difficult to obtain. Tong Zhang suggested replacing the notation $a_t = (i, I_t)$ by $a_t = (i_t, I_t)$. The author thanks Jelani Nelson, Ravi Kumar, S. Muthukrishnan, Graham Cormode, Phil Long, Johannes Gehrke, Anand Vidyashankar, and Martin Wells for helpful communications.

A Proof of Lemma 3.1

Assume $Z \sim S(\alpha,\beta,F_{(\alpha)})$. The goal is to compute $\mathbf{E}\left(|Z|^\lambda\right)$, $-1 < \lambda < \alpha$. [24, Theorem 2.6.3] provided a partial answer:

$$\int_0^\infty z^\lambda f_Z(z;\alpha,\beta_B,F_{(\alpha)})dz = F_{(\alpha)}^{\lambda/\alpha}\frac{\sin(\pi\rho\lambda)}{\sin(\pi\lambda)}\frac{\Gamma\left(1-\frac{\lambda}{\alpha}\right)}{\Gamma(1-\lambda)}\cos^{-\lambda/\alpha}\left(\pi\beta_B\kappa(\alpha)/2\right)$$

where $\kappa(\alpha) = \alpha$ if $\alpha < 1$, and $\kappa(\alpha) = 2 - \alpha$ if $\alpha > 1$, and according to the parametrization in [24, I.19, I.28]:

$$\beta_B = \frac{2}{\pi\kappa(\alpha)}\tan^{-1}\left(\beta\tan\left(\frac{\pi\alpha}{2}\right)\right), \quad \rho = \frac{1-\beta_B\kappa(a)/\alpha}{2}.$$

Note that

$$\cos^{-\lambda/\alpha}\left(\pi\beta_B\kappa(\alpha)/2\right) = \left(1+\tan^2\left(\pi\beta_B\kappa(\alpha)/2\right)\right)^{\frac{\lambda}{2\alpha}}$$
$$= \left(1+\tan^2\left(\tan^{-1}\left(\beta\tan\left(\frac{\pi\alpha}{2}\right)\right)\right)\right)^{\frac{\lambda}{2\alpha}} = \left(1+\beta^2\tan^2\left(\frac{\pi\alpha}{2}\right)\right)^{\frac{\lambda}{2\alpha}}.$$

Therefore, for $-1 < \lambda < \alpha$,

$$\int_0^\infty z^\lambda f_Z(z;\alpha,\beta_B,F_{(\alpha)})dz = F_{(\alpha)}^{\lambda/\alpha}\frac{\sin(\pi\rho\lambda)}{\sin(\pi\lambda)}\frac{\Gamma\left(1-\frac{\lambda}{\alpha}\right)}{\Gamma(1-\lambda)}\left(1+\beta^2\tan^2\left(\frac{\pi\alpha}{2}\right)\right)^{\frac{\lambda}{2\alpha}}$$

To compute $\mathrm{E}\left(|Z|^{\lambda}\right)$, we use the property[24, page 65]: $f_Z(-z;\alpha,\beta_B,F_{(\alpha)}) = f_Z(z;\alpha,-\beta_B,F_{(\alpha)})$ to obtain:

$$
\begin{aligned}
&\mathrm{E}\left(|Z|^{\lambda}\right) = \int_{-\infty}^{0}(-z)^{\lambda} f_Z(z;\alpha,\beta_B,F_{(\alpha)})dz + \int_0^{\infty} z^{\lambda} f_Z(z;\alpha,\beta_B,F_{(\alpha)})dz \\
&= \int_0^{\infty} z^{\lambda} f_Z(z;\alpha,-\beta_B,F_{(\alpha)})dz + \int_0^{\infty} z^{\lambda} f_Z(z;\alpha,\beta_B,F_{(\alpha)})dz \\
&= \frac{F_{(\alpha)}^{\lambda/\alpha}}{\sin(\pi\lambda)} \frac{\Gamma\left(1-\frac{\lambda}{\alpha}\right)}{\Gamma(1-\lambda)} \left(1+\beta^2\tan^2\left(\frac{\pi\alpha}{2}\right)\right)^{\frac{\lambda}{2\alpha}} \\
&\quad\times\left(\sin\left(\pi\lambda\frac{1-\beta_B\kappa(\alpha)/\alpha}{2}\right) + \sin\left(\pi\lambda\frac{1+\beta_B\kappa(\alpha)/\alpha}{2}\right)\right) \\
&= \frac{F_{(\alpha)}^{\lambda/\alpha}}{\sin(\pi\lambda)} \frac{\Gamma\left(1-\frac{\lambda}{\alpha}\right)}{\Gamma(1-\lambda)} \left(1+\beta^2\tan^2\left(\frac{\pi\alpha}{2}\right)\right)^{\frac{\lambda}{2\alpha}} \left(2\sin\left(\frac{\pi\lambda}{2}\right)\cos\left(\frac{\pi\lambda}{2}\beta_B\kappa(\alpha)/\alpha\right)\right) \\
&= \frac{F_{(\alpha)}^{\lambda/\alpha}}{\cos(\pi\lambda/2)} \frac{\Gamma\left(1-\frac{\lambda}{\alpha}\right)}{\Gamma(1-\lambda)} \left(1+\beta^2\tan^2\left(\frac{\pi\alpha}{2}\right)\right)^{\frac{\lambda}{2\alpha}} \cos\left(\frac{\lambda}{\alpha}\tan^{-1}\left(\beta\tan\left(\frac{\pi\alpha}{2}\right)\right)\right) \\
&= F_{(\alpha)}^{\lambda/\alpha}\left(1+\beta^2\tan^2\left(\frac{\pi\alpha}{2}\right)\right)^{\frac{\lambda}{2\alpha}} \cos\left(\frac{\lambda}{\alpha}\tan^{-1}\left(\beta\tan\left(\frac{\pi\alpha}{2}\right)\right)\right) \\
&\quad\times\left(\frac{2}{\pi}\sin\left(\frac{\pi}{2}\lambda\right)\Gamma\left(1-\frac{\lambda}{\alpha}\right)\Gamma(\lambda)\right),
\end{aligned}
$$

which can be simplified when $\beta = 1$, to be

$$
\mathrm{E}\left(|Z|^{\lambda}\right) = F_{(\alpha)}^{\lambda/\alpha} \frac{\cos\left(\frac{\kappa(\alpha)}{\alpha}\frac{\lambda\pi}{2}\right)}{\cos^{\lambda/\alpha}\left(\frac{\kappa(\alpha)\pi}{2}\right)} \left(\frac{2}{\pi}\sin\left(\frac{\pi}{2}\lambda\right)\Gamma\left(1-\frac{\lambda}{\alpha}\right)\Gamma(\lambda)\right).
$$

For $0<\lambda<\alpha$, in an unpublished work[14], a partial result for $\mathrm{E}\left(|Z|^{\lambda}\right)$ was proved in an integral form, using a different method (via local properties of characteristic functions).

B Proof of Lemma 3.2

Note that when $\alpha<1$ and $\beta=1$, Z is always non-negative. As shown in the proof of [24, Theorem 2.6.3],

$$
\begin{aligned}
&\mathrm{E}\left(|Z|^{\lambda}\right) = F_{(\alpha)}^{\lambda/\alpha}\cos^{-\lambda/\alpha}\left(\frac{\pi\alpha}{2}\right)\frac{1}{\pi}\mathrm{Im}\int_0^{\infty} z^{\lambda}\int_0^{\infty} \\
&\exp\left(-zu\exp(\sqrt{-1}\pi/2) - u^{\alpha}\exp(-\sqrt{-1}\pi\alpha/2) + \frac{\sqrt{-1}\pi}{2}\right)dudz \\
&= F_{(\alpha)}^{\lambda/\alpha}\cos^{-\lambda/\alpha}\left(\frac{\pi\alpha}{2}\right)\frac{1}{\pi}\mathrm{Im}\int_0^{\infty}\int_0^{\infty} \\
&z^{\lambda}\exp\left(-zu\sqrt{-1} - u^{\alpha}\exp(-\sqrt{-1}\pi\alpha/2)\right)\sqrt{-1}dudz.
\end{aligned}
$$

We can show, in the proof of [24, Theorem 2.6.3], Fubini's condition still holds when $\alpha<1$, $\beta=1$, and $\lambda<-1$:

$$
\begin{aligned}
&\int_0^{\infty}\int_0^{\infty}\left|z^{\lambda}\exp\left(-zu\sqrt{-1} - u^{\alpha}\exp(-\sqrt{-1}\pi\alpha/2)\right)\sqrt{-1}\right|dudz \\
&= \int_0^{\infty}\int_0^{\infty} z^{\lambda}\left|\exp\left(-u^{\alpha}\cos(\pi\alpha/2) + \sqrt{-1}u^{\alpha}\sin(\pi\alpha/2)\right)\right|dudz \\
&= \int_0^{\infty}\int_0^{\infty} z^{\lambda}\exp\left(-u^{\alpha}\cos(\pi\alpha/2)\right)dudz < \infty,
\end{aligned}
$$

provided $\lambda \neq -1,-2,\ldots$ (which do not affect the results due to continuity) and $\cos(\pi\alpha/2)>0$, i.e., $\alpha<1$. Once Fubini's condition has been shown to hold, we can exchange the order of integration and the rest follows from [24, Theorem 2.6.3].

C Proof of Lemma 4.2

This section concerns the convergence and monotonicity of

$$
\left[\cos\left(\frac{\kappa(\alpha)\pi}{2k}\right)\frac{2}{\pi}\Gamma\left(\frac{\alpha}{k}\right)\Gamma\left(1-\frac{1}{k}\right)\sin\left(\frac{\pi}{2}\frac{\alpha}{k}\right)\right]^k. \tag{3.10}
$$

First, we can show

$$
\begin{aligned}
&\cos^k\left(\frac{\kappa(\alpha)\pi}{2k}\right) = \exp\left(k\log\cos\left(\frac{\kappa(\alpha)\pi}{2k}\right)\right) \\
&= \exp\left(k\left(-\frac{1}{2}\left(\frac{\kappa(\alpha)\pi}{2k}\right)^2 - \frac{1}{12}\left(\frac{\kappa(\alpha)\pi}{2k}\right)^4 + \ldots\right)\right) = \exp\left(-\frac{\kappa^2\pi^2}{8k} - \frac{\kappa^4\pi^4}{192k^3} + \ldots\right).
\end{aligned}
$$

Next, by Euler's reflection formula, $\Gamma(1-z)\Gamma(z) = \frac{\pi}{\sin(\pi z)}$,

$$
\left[\frac{2}{\pi}\Gamma\left(\frac{\alpha}{k}\right)\Gamma\left(1-\frac{1}{k}\right)\sin\left(\frac{\pi}{2}\frac{\alpha}{k}\right)\right]^k = \left[\frac{2\Gamma\left(\frac{\alpha}{k}\right)\sin\left(\frac{\pi\alpha}{2k}\right)}{\Gamma\left(\frac{1}{k}\right)\sin\left(\frac{\pi}{k}\right)}\right]^k
$$

We take advantage of the infinite-product representations of the Gamma and sin functions[8, 8.322,1.431.1]:

$$
\Gamma(z) = \frac{\exp(-\gamma_e z)}{z}\prod_{s=1}^{\infty}\left(1+\frac{z}{s}\right)^{-1}\exp\left(\frac{z}{s}\right), \quad \sin(z) = z\prod_{s=1}^{\infty}\left(1-\frac{z^2}{s^2\pi^2}\right),
$$

where $\gamma_e = 0.577215665\ldots$, is Euler's constant, to obtain

$$
\left[\frac{2\Gamma\left(\frac{\alpha}{k}\right)\sin\left(\frac{\pi\alpha}{2k}\right)}{\Gamma\left(\frac{1}{k}\right)\sin\left(\frac{\pi}{k}\right)}\right]^k = \exp\left(-\gamma_e(\alpha-1)\right)\times\left(\prod_{s=1}^{\infty}W_s\right)^k,
$$

$$
\begin{aligned}
&W_s = \exp\left(\frac{\alpha-1}{sk}\right)\left(1+\frac{\alpha}{ks}\right)^{-1}\left(1+\frac{1}{ks}\right)\left(1-\frac{\alpha^2}{4k^2s^2}\right)\left(1-\frac{1}{s^2k^2}\right)^{-1} \\
&= \left(1+\frac{\alpha-1}{sk}+\frac{(\alpha-1)^2}{2s^2k^2}+\frac{(\alpha-1)^3}{6s^3k^3}+\ldots\right)\left(1-\frac{\alpha}{sk}+\frac{\alpha^2}{s^2k^2}-\frac{\alpha^3}{s^3k^3}+\ldots\right) \\
&\left(1+\frac{1}{ks}-\frac{\alpha^2}{4k^2s^2}-\frac{\alpha^2}{4k^3s^3}\right)\left(1+\frac{1}{s^2k^2}-\frac{1}{s^4k^4}+\ldots\right) \\
&= 1+\frac{1}{s^2k^2}\left(\frac{1}{2}+\frac{\alpha^2}{4}\right)+\frac{1}{s^3k^3}(1-\alpha)\left(\alpha+\frac{(1-\alpha)^2}{3}\right)+\ldots
\end{aligned}
$$

$$
\begin{aligned}
&\left(\prod_{s=1}^{\infty}W_s\right)^k = \exp\left(k\sum_{s=1}^{\infty}\log W_s\right) \\
&= \exp\left(k\sum_{s=1}^{\infty}\log\left(1+\frac{1}{s^2k^2}\left(\frac{1}{2}+\frac{\alpha^2}{4}\right)+\frac{1-\alpha}{s^3k^3}\left(\alpha+\frac{(1-\alpha)^2}{3}\right)+\ldots\right)\right) \\
&= \exp\left(\frac{1}{k}\left(\frac{1}{2}+\frac{\alpha^2}{4}\right)\sum_{i=1}^{\infty}\frac{1}{s^2}+\frac{1-\alpha}{k^2}\left(\alpha+\frac{(1-\alpha)^2}{3}\right)\sum_{i=1}^{\infty}\frac{1}{s^3}+\ldots\right) \\
&= \exp\left(\frac{1}{k}\left(\frac{1}{2}+\frac{\alpha^2}{4}\right)\frac{\pi^2}{6}+\frac{1-\alpha}{k^2}\left(\alpha+\frac{(1-\alpha)^2}{3}\right)\zeta_3+\ldots\right) \\
&= \exp\left(\frac{1}{k}\left(2+\alpha^2\right)\frac{\pi^2}{24}+\frac{1}{k^2}\frac{1-\alpha^3}{3}\zeta_3+\ldots\right)
\end{aligned}
$$

where $\sum_{s=1}^{\infty}\frac{1}{s^2} = \frac{\pi^2}{6}$ and $\zeta_3 = \sum_{s=1}^{\infty}\frac{1}{s^3} = 1.2020569\ldots$ Thus,

$$
\begin{aligned}
&\left[\cos\left(\frac{\kappa(\alpha)\pi}{2k}\right)\frac{2}{\pi}\Gamma\left(\frac{\alpha}{k}\right)\Gamma\left(1-\frac{1}{k}\right)\sin\left(\frac{\pi}{2}\frac{\alpha}{k}\right)\right]^k = \exp\left(-\gamma_e(\alpha-1)\right) \\
&\times\exp\left(\frac{1}{k}\left(2+\alpha^2-3\kappa^2(\alpha)\right)\frac{\pi^2}{24}+\frac{1}{k^2}\frac{1-\alpha^3}{3}\zeta_3+\ldots\right) \to \exp\left(-\gamma_e(\alpha-1)\right)
\end{aligned}
$$

It remains to show (3.10) is monotonically decreasing.

Suppose $\alpha>1$, i.e., $\kappa(\alpha) = 2-\alpha<1$. For simplicity, we take the logarithm of (3.10) and replace $1/k$ by t, where $0\le t\le 1/2$ (recall $k\ge 2$). It suffices to show that $g(t) = \frac{1}{t}W(t)$ increases with increasing $t\in[0,1/2]$, where

$$
\begin{aligned}
W(t) = &\log\left(\cos\left(\frac{\kappa(\alpha)\pi}{2}t\right)\right)+\log\left(\Gamma(\alpha t)\right)+\log\left(\sin\left(\frac{\pi\alpha}{2}t\right)\right) \\
&-\log\left(\Gamma(t)\right)-\log\left(\sin(\pi t)\right)+\log(2).
\end{aligned}
$$

Because $g'(t) = \frac{1}{t}W'(t) - \frac{1}{t^2}W(t)$, to show $g'(t)\ge 0$ in $t\in[0,1/2]$, it suffices to show $tW'(t)-W(t)\ge 0$.

One can check $tW'(t)\to 0$ and $W(t)\to 0$, as $t\to 0+$.

$$
W'(t) = -\tan\left(\frac{\kappa(\alpha)\pi}{2}t\right)\frac{\kappa\pi}{2}+\psi(\alpha t)\alpha+\frac{1}{\tan\left(\frac{\pi\alpha}{2}t\right)}\left(\frac{\alpha\pi}{2}\right)-\psi(t)-\frac{\pi}{\tan(\pi t)}.
$$

Here $\psi(x) = \frac{\partial\log(\Gamma(x))}{\partial x}$ is the "Psi" function. Therefore, to show $tW'(t)-W(t)\ge 0$, it suffices to show that $tW'(t)-W(t)$ is an increasing function of $t\in[0,1/2]$, i.e.,

$$
\begin{aligned}
&\left(tW'(t)-W(t)\right)' = W''(t)\ge 0, \quad \text{i.e.,} \\
&W''(t) = -\sec^2\left(\frac{\kappa(\alpha)\pi}{2}t\right)\left(\frac{\kappa(\alpha)\pi}{2}\right)^2+\psi'(\alpha t)\alpha^2 \\
&\qquad -\csc^2\left(\frac{\pi\alpha}{2}t\right)\left(\frac{\pi\alpha}{2}\right)^2-\psi'(t)+\csc^2(\pi t)\pi^2\ge 0.
\end{aligned}
$$

Using series representation of $\psi(x)$ [8, 8.363.8] yields

$$\psi'(\alpha t)\,\alpha^2 - \psi'(t) = \sum_{s=0}^{\infty} \frac{\alpha^2}{(\alpha t+s)^2} - \sum_{s=0}^{\infty} \frac{1}{(t+s)^2} \geq 0,$$

because we consider $\alpha > 1$. Thus, it suffices to show that

$$Q(t;\alpha) = -\sec^2\left(\frac{\kappa\pi}{2}t\right)\left(\frac{\kappa\pi}{2}\right)^2 - \csc^2\left(\frac{\pi\alpha}{2}t\right)\left(\frac{\pi\alpha}{2}\right)^2 + \csc^2(\pi t)\pi^2 \geq 0.$$

To show $Q(t;\alpha) \geq 0$,we notice both $\frac{1}{\sin(x)}$ and $\frac{1}{\cos(x)}$ are convex functions of $x \in [0,\pi/2]$, and hence $Q(t;\alpha)$ is a concave function of α (for fixed t). Because $\lim_{\alpha\to 1+} Q(t;\alpha) = 0$ and $\lim_{\alpha\to 2-} Q(t;\alpha) = 0$, and $Q(t;\alpha)$ is concave in $\alpha \in [1,2]$, we must have $Q(t;\alpha) \geq 0$; and consequently, $W''(t) \geq 0$ and $g'(t) \geq 0$. Therefore, we have proved that (3.10) decreases monotonically with increasing k, when $1 < \alpha \leq 2$.

To prove the monotonicity for $\alpha < 1$, we re-write (3.10):

$$\left[\frac{\Gamma\left(\frac{\alpha}{k}\right)\sin\left(\frac{\pi\alpha}{k}\right)}{\Gamma\left(\frac{1}{k}\right)\sin\left(\frac{\pi}{k}\right)}\right]^k = \exp\left(-\gamma_e(\alpha-1)\right) \times \left(\prod_{s=1}^{\infty} D_s\right)^k$$

$$D_s = \exp\left(\frac{\alpha-1}{sk}\right)\left(1+\frac{\alpha}{ks}\right)^{-1}\left(1+\frac{1}{ks}\right)\left(1-\frac{\alpha^2}{k^2s^2}\right)\left(1-\frac{1}{s^2k^2}\right)^{-1}.$$

It suffices to show for any $s \geq 1$

$$\left(\left(1+\frac{\alpha}{ks}\right)^{-1}\left(1+\frac{1}{ks}\right)\left(1-\frac{\alpha^2}{k^2s^2}\right)\left(1-\frac{1}{s^2k^2}\right)^{-1}\right)^k$$

decreases monotonically, which is equivalent to show the monotonicity of $g(t)$ with increasing t, for $t \geq 2$, where

$$g(t) = t\log\left(\left(1+\frac{\alpha}{t}\right)^{-1}\left(1+\frac{1}{t}\right)\left(1-\frac{\alpha^2}{t^2}\right)\left(1-\frac{1}{t^2}\right)^{-1}\right) = t\log\left(\frac{t-\alpha}{t-1}\right),$$

which is monotonically decreasing with increasing t ($t \geq 2$).

D Proof of Lemma 4.3

Applying the moment formula in Lemma 3.1 yields

$$\mathrm{E}\left(\left(\hat{F}_{(\alpha),gm}\right)^t\right) = F_{(\alpha)}^t \frac{\cos^k\left(\frac{\kappa(\alpha)\pi}{2k}t\right)\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}t\right)\Gamma\left(1-\frac{t}{k}\right)\Gamma\left(\frac{\alpha}{k}t\right)\right]^k}{\cos^{kt}\left(\frac{\kappa(\alpha)\pi}{2k}\right)\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}\right)\Gamma\left(1-\frac{1}{k}\right)\Gamma\left(\frac{\alpha}{k}\right)\right]^{kt}}$$

In [15], it was proved that, as $k \to \infty$,

$$\frac{\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}t\right)\Gamma\left(1-\frac{t}{k}\right)\Gamma\left(\frac{\alpha}{k}t\right)\right]^k}{\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}\right)\Gamma\left(1-\frac{1}{k}\right)\Gamma\left(\frac{\alpha}{k}\right)\right]^{kt}}$$
$$= \exp\left(\frac{1}{k}\frac{\pi^2(t^2-t)}{24}\left(\alpha^2+2\right) + \frac{1}{k^2}\frac{t^3-t}{3}(1-\alpha^3)\zeta_3 + O\left(\frac{1}{k^3}\right)\right).$$

Using the infinite product representation of cosine[8, 1.43.3] $\cos(z) = \prod_{s=0}^{\infty}\left(1-\frac{4z^2}{(2s+1)^2\pi^2}\right)$, we can re-write

$$\frac{\cos^k\left(\frac{\kappa(\alpha)\pi}{2k}t\right)}{\cos^{kt}\left(\frac{\kappa(\alpha)\pi}{2k}\right)} = \left(\prod_{s=0}^{\infty}\left(1-\frac{\kappa^2(\alpha)t^2}{(2s+1)^2k^2}\right)\left(1-\frac{\kappa^2(\alpha)}{(2s+1)^2k^2}\right)^{-t}\right)^k$$
$$= \left(\prod_{s=0}^{\infty}\left(1-\frac{\kappa^2(\alpha)t^2}{(2s+1)^2k^2}\right)\left(1+\frac{t\kappa^2(\alpha)}{(2s+1)^2k^2}+\frac{t(t-1)\kappa^4(\alpha)}{2(2s+1)^4k^4}+O\left(\frac{1}{k^5}\right)\right)\right)^k$$
$$= \left(\prod_{s=0}^{\infty} 1-\frac{\kappa^2(\alpha)(t^2-t)}{(2s+1)^2k^2}+O\left(\frac{1}{k^4}\right)\right)^k$$
$$= \exp\left(k\sum_{s=0}^{\infty}\log\left(1-\frac{\kappa^2(\alpha)(t^2-t)}{(2s+1)^2k^2}+O\left(\frac{1}{k^4}\right)\right)\right)$$
$$= \exp\left(\sum_{s=0}^{\infty} -\frac{\kappa^2(\alpha)(t^2-t)}{(2s+1)^2k}+O\left(\frac{1}{k^3}\right)\right) \qquad \left(\sum_{s=0}^{\infty}\frac{1}{2s+1} = \frac{\pi^2}{8}\right)$$
$$= \exp\left(-\frac{\kappa^2(\alpha)}{k}(t^2-t)\frac{\pi^2}{8}+O\left(\frac{1}{k^3}\right)\right),$$

Therefore,

$$\frac{\cos^k\left(\frac{\kappa(\alpha)\pi}{2k}t\right)\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}t\right)\Gamma\left(1-\frac{t}{k}\right)\Gamma\left(\frac{\alpha}{k}t\right)\right]^k}{\cos^{kt}\left(\frac{\kappa(\alpha)\pi}{2k}\right)\left[\frac{2}{\pi}\sin\left(\frac{\pi\alpha}{2k}\right)\Gamma\left(1-\frac{1}{k}\right)\Gamma\left(\frac{\alpha}{k}\right)\right]^{kt}}$$
$$= \exp\left(\frac{1}{k}\frac{\pi^2(t^2-t)}{24}\left(\alpha^2+2-3\kappa^2(\alpha)\right)+\frac{1}{k^2}\frac{t^3-t}{3}(1-\alpha^3)\zeta_3+O\left(\frac{1}{k^3}\right)\right)$$
$$= 1+\frac{1}{k}\frac{\pi^2(t^2-t)}{24}\left(\alpha^2+2-3\kappa^2(\alpha)\right)+\frac{1}{k^2}\frac{t^3-t}{3}(1-\alpha^3)\zeta_3$$
$$+\frac{1}{k^2}\frac{\pi^4(t^2-t)^2}{1152}\left(\alpha^2+2-3\kappa^2(\alpha)\right)^2+O\left(\frac{1}{k^3}\right)$$

E Proof of Lemma 4.7

First, we consider the right bound. From Lemma 4.6,

$$\frac{\epsilon^2}{G_{R,gm}} = C_R\log(1+\epsilon) - C_R\gamma_e(\alpha-1)$$
$$-\log\left(\cos\left(\frac{\kappa(\alpha)\pi C_R}{2}\right)\frac{2}{\pi}\Gamma\left(\alpha C_R\right)\Gamma\left(1-C_R\right)\sin\left(\frac{\pi\alpha C_R}{2}\right)\right),$$

and C_R is the solution to $g_1(C_R,\alpha,\epsilon) = 0$,

$$g_1(C_R,\alpha,\epsilon) = -\gamma_e(\alpha-1)+\log(1+\epsilon)+\frac{\kappa(\alpha)\pi}{2}\tan\left(\frac{\kappa(\alpha)\pi}{2}C_R\right)$$
$$-\frac{\alpha\pi/2}{\tan\left(\frac{\alpha\pi}{2}C_R\right)}-\psi\left(\alpha C_R\right)\alpha+\psi\left(1-C_R\right) = 0.$$

Using series representations in [8, 1.421.1,1.421.3,8.362.1]

$$\tan\left(\frac{\pi x}{2}\right) = \frac{4x}{\pi}\sum_{j=1}^{\infty}\frac{1}{(2j-1)^2-x^2}, \qquad \frac{1}{\tan(\pi x)} = \frac{1}{\pi x}+\frac{2x}{\pi}\sum_{j=1}^{\infty}\frac{1}{x^2-j^2},$$

$$\psi(x) = -\gamma_e-\sum_{j=0}^{\infty}\left(\frac{1}{x+j}-\frac{1}{j+1}\right) = -\gamma_e-\frac{1}{x}+x\sum_{j=1}^{\infty}\frac{1}{j(x+j)},$$

we re-write g_1 as

$$g_1 = -\gamma_e(\alpha-1)+\log(1+\epsilon)+\frac{\kappa\pi}{2}\frac{4\kappa C_R}{\pi}\sum_{j=1}^{\infty}\frac{1}{(2j-1)^2-(\kappa C_R)^2}$$
$$-\frac{\alpha\pi}{2}\left(\frac{2}{\pi\alpha C_R}+\frac{\alpha C_R}{\pi}\sum_{j=1}^{\infty}\frac{1}{(\alpha C_R/2)^2-j^2}\right)$$
$$-\alpha\left(-\gamma_e-\frac{1}{\alpha C_R}+\alpha C_R\sum_{j=1}^{\infty}\frac{1}{j(\alpha C_R+j)}\right)$$
$$+\left(-\gamma_e-\frac{1}{1-C_R}+(1-C_R)\sum_{j=1}^{\infty}\frac{1}{j(1-C_R+j)}\right)$$
$$= \log(1+\epsilon)+\kappa\sum_{j=1}^{\infty}\left(\frac{1}{2j+1-\kappa C_R}-\frac{1}{2j-1+\kappa C_R}\right)$$
$$+\alpha\sum_{j=1}^{\infty}\left(\frac{1}{2j-\alpha C_R}-\frac{1}{2j+\alpha C_R}\right)-\alpha\sum_{j=1}^{\infty}\left(\frac{1}{j}-\frac{1}{\alpha C_R+j}\right)$$
$$+\sum_{j=1}^{\infty}\left(\frac{1}{j}-\frac{1}{1-C_R+j}\right)+\frac{\kappa}{1-\kappa C_R}-\frac{1}{1-C_R}$$

We show that, as $\alpha \to 1$, i.e., $\kappa \to 1$, the term

$$\lim_{\alpha\to 1}\kappa\sum_{j=1}^{\infty}\left(\frac{1}{2j+1-\kappa C_R}-\frac{1}{2j-1+\kappa C_R}\right)-\alpha\sum_{j=1}^{\infty}\left(\frac{1}{j}-\frac{1}{\alpha C_R+j}\right)$$
$$+\alpha\sum_{j=1}^{\infty}\left(\frac{1}{2j-\alpha C_R}-\frac{1}{2j+\alpha C_R}\right)+\sum_{j=1}^{\infty}\left(\frac{1}{j}-\frac{1}{1-C_R+j}\right)$$
$$= \lim_{\alpha\to 1}\sum_{j=1}^{\infty}\left(\frac{\kappa}{2j+1-\kappa C_R}+\frac{\alpha}{2j-\alpha C_R}\right)-\alpha\sum_{j=1}^{\infty}\left(\frac{1}{j}-\frac{1}{\alpha C_R+j}\right)$$
$$-\sum_{j=1}^{\infty}\left(\frac{\kappa}{2j-1+\kappa C_R}+\frac{\alpha}{2j+\alpha C_R}\right)+\sum_{j=1}^{\infty}\left(\frac{1}{j}-\frac{1}{1-C_R+j}\right)$$
$$= \lim_{\alpha\to 1}\sum_{j=1}^{\infty}\frac{\kappa}{1+j-\kappa C_R}-\sum_{j=1}^{\infty}\frac{\kappa}{j+\kappa C_R}-\alpha\sum_{j=1}^{\infty}\left(\frac{1}{j}-\frac{1}{\alpha C_R+j}\right)$$
$$+\sum_{j=1}^{\infty}\left(\frac{1}{j}-\frac{1}{1-C_R+j}\right) = 0.$$

Lemma 4.6 has shown $g_1 = 0$ has a unique well-defined solution for $C_R \in (0,1)$. We need to analyze this term

$$\frac{\kappa}{1-\kappa C_R} - \frac{1}{1-C_R} = \frac{\kappa-1}{(1-\kappa C_R)(1-C_R)} = \frac{-\Delta}{(1-\kappa C_R)(1-C_R)},$$

which, as $\alpha \to 1$ (i.e., $\kappa \to 1$), must approach a finite limit. In other words, $C_R \to 1$, at the rate $O\left(\sqrt{\Delta}\right)$, i.e.,

$$C_R = 1 - \sqrt{\frac{\Delta}{\log(1+\epsilon)}} + o\left(\sqrt{\Delta}\right).$$

By Euler's reflection formula and series representations,

$$\frac{\epsilon^2}{G_{R,gm}} = C_R \log(1+\epsilon) - C_R \gamma_e(\alpha-1) + \log\left(\frac{\cos\left(\frac{\alpha\pi C_R}{2}\right)\Gamma(1-\alpha C_R)}{\cos\left(\frac{\kappa\pi C_R}{2}\right)\Gamma(1-C_R)}\right),$$

$$\begin{aligned}
&\frac{\cos\left(\frac{\alpha\pi C_R}{2}\right)\Gamma(1-\alpha C_R)}{\cos\left(\frac{\kappa\pi C_R}{2}\right)\Gamma(1-C_R)} \\
=&\exp(\gamma_e(\alpha-1)C_R)\frac{1-C_R}{1-\alpha C_R}\prod_{j=0}^{\infty}\left(1-\frac{\alpha^2C_R^2}{(2j+1)^2}\right)\left(1-\frac{\kappa^2C_R^2}{(2j+1)^2}\right)^{-1} \\
&\times\prod_{j=1}^{\infty}\exp\left(\frac{(1-\alpha)C_R}{j}\right)\left(1+\frac{1-C_R}{j}\right)\left(1+\frac{1-\alpha C_R}{j}\right)^{-1} \\
=&\exp(\gamma_e(\alpha-1)C_R)\frac{(1+\alpha C_R)(1-C_R)}{1-\kappa^2C_R^2}\prod_{j=1}^{\infty}\left(1-\frac{\alpha^2C_R^2}{(2j+1)^2}\right) \\
&\times\left(1-\frac{\kappa^2C_R^2}{(2j+1)^2}\right)^{-1}\exp\left(\frac{(1-\alpha)C_R}{j}\right)\left(1+\frac{1-C_R}{j}\right)\left(1+\frac{1-\alpha C_R}{j}\right)^{-1},
\end{aligned}$$

taking its logarithm yields

$$\begin{aligned}
&\log\frac{\cos\left(\frac{\alpha\pi C_R}{2}\right)\Gamma(1-\alpha C_R)}{\cos\left(\frac{\kappa\pi C_R}{2}\right)\Gamma(1-C_R)} = \gamma_e(\alpha-1)C_R + \log\frac{(1+\alpha C_R)(1-C_R)}{1-\kappa^2C_R^2} \\
&+\sum_{j=1}^{\infty}\log\frac{\left(1-\frac{\alpha^2C_R^2}{(2j+1)^2}\right)}{\left(1-\frac{\kappa^2C_R^2}{(2j+1)^2}\right)} + \left(\frac{(1-\alpha)C_R}{j}\right) + \log\frac{\left(1+\frac{1-C_R}{j}\right)}{\left(1+\frac{1-\alpha C_R}{j}\right)}.
\end{aligned}$$

If $\alpha < 1$, i.e., $\kappa = \alpha = 1-\Delta$, then

$$\begin{aligned}
&\log\frac{\cos\left(\frac{\alpha\pi C_R}{2}\right)\Gamma(1-\alpha C_R)}{\cos\left(\frac{\kappa\pi C_R}{2}\right)\Gamma(1-C_R)} \\
=&-\gamma_e\Delta C_R + \log\frac{1-C_R}{1-\alpha C_R} + \sum_{j=1}^{\infty}\left(\frac{(1-\alpha)C_R}{j}\right) + \log\frac{\left(1+\frac{1-C_R}{j}\right)}{\left(1+\frac{1-\alpha C_R}{j}\right)} \\
=&-\gamma_e\Delta C_R - \log\left(1+\frac{\Delta C_R}{1-C_R}\right) + \sum_{j=1}^{\infty}\frac{1}{2}\left(\frac{1-\alpha C_R}{j}\right)^2 - \frac{1}{2}\left(\frac{1-C_R}{j}\right)^2 \ldots \\
=&-\gamma_e\Delta C_R - \log\left(1+\frac{\Delta C_R}{1-C_R}\right) + \frac{\pi^2}{12}C_R\Delta(2-\alpha C_R - C_R) + \ldots
\end{aligned}$$

Thus, for $\alpha < 1$, as $C_R = 1-\sqrt{\frac{\Delta}{\log(1+\epsilon)}} + o\left(\sqrt{\Delta}\right)$, we obtain

$$\begin{aligned}
\frac{\epsilon^2}{G_{R,gm}} =& C_R\log(1+\epsilon) - \frac{\Delta C_R}{1-C_R} + \frac{\pi^2}{12}C_R\Delta(2-\alpha C_R - C_R) + \ldots \\
=& \log(1+\epsilon) - 2\sqrt{\Delta\log(1+\epsilon)} + o\left(\sqrt{\Delta}\right)
\end{aligned}$$

If $\alpha > 1$, i.e., $\alpha = 1+\Delta$ and $\kappa = 1-\Delta$, then

$$\begin{aligned}
&\log\frac{\cos\left(\frac{\alpha\pi C_R}{2}\right)\Gamma(1-\alpha C_R)}{\cos\left(\frac{\kappa\pi C_R}{2}\right)\Gamma(1-C_R)} \\
=&\gamma_e\Delta C_R + \log\frac{(1+\alpha C_R)(1-C_R)}{1-\kappa^2C_R^2} + \sum_{j=1}^{\infty}\log\frac{\left(1-\frac{\alpha^2C_R^2}{(2j+1)^2}\right)}{\left(1-\frac{\kappa^2C_R^2}{(2j+1)^2}\right)} + \ldots
\end{aligned}$$

$$\begin{aligned}
&\log\frac{(1+\alpha C_R)(1-C_R)}{1-\kappa^2C_R^2} = \log\frac{1+\alpha C_R}{1+\kappa C_R} - \log\frac{1-\kappa C_R}{1-C_R} \\
=&\log\left(1+\frac{2\Delta C_R}{1+\kappa C_R}\right) - \log\left(1+\frac{\Delta C_R}{1-C_R}\right) = -\sqrt{\Delta\log(1+\epsilon)} + o\left(\sqrt{\Delta}\right),
\end{aligned}$$

$$\begin{aligned}
&\sum_{j=1}^{\infty}\log\frac{\left(1-\frac{\alpha^2C_R^2}{(2j+1)^2}\right)}{\left(1-\frac{\kappa^2C_R^2}{(2j+1)^2}\right)} = \sum_{j=1}^{\infty}\log\frac{1+\frac{\alpha C_R}{2j+1}}{1+\frac{\kappa C_R}{2j+1}} + \log\frac{1-\frac{\alpha C_R}{2j+1}}{1-\frac{\kappa C_R}{2j+1}} \\
=&\sum_{j=1}^{\infty}\log\left(1+\frac{\frac{2\Delta C_R}{2j+1}}{1+\frac{\kappa C_R}{2j+1}}\right) + \log\left(1-\frac{\frac{2\Delta C_R}{2j+1}}{1-\frac{\kappa C_R}{2j+1}}\right) = O(\Delta).
\end{aligned}$$

Therefore, for $\alpha > 1$, we also have

$$\frac{\epsilon^2}{G_{R,gm}} = \log(1+\epsilon) - 2\sqrt{\Delta\log(1+\epsilon)} + o\left(\sqrt{\Delta}\right).$$

Next, we consider the left bound. From Lemma 4.6,

$$\mathbf{Pr}\left(\hat{F}_{(\alpha),gm,b} - F_{(\alpha)} \le -\epsilon F_{(\alpha)}\right) \le \exp\left(-k\frac{\epsilon^2}{G_{L,gm}}\right),$$

$$\begin{aligned}
\frac{\epsilon^2}{G_{L,gm}} =& -C_L\log(1-\epsilon) + C_L\gamma_e(\alpha-1) + \log\alpha \\
&-\log\left(\cos\left(\frac{\kappa(\alpha)\pi}{2}C_L\right)\Gamma(C_L)\right) + \log\left(\Gamma(\alpha C_L)\cos\left(\frac{\pi\alpha C_L}{2}\right)\right).
\end{aligned}$$

and C_L is the solution to $g_2(C_L,\alpha,\epsilon) = 0$,

$$\begin{aligned}
g_2(C_L,\alpha,\epsilon) =& \log(1-\epsilon) - \gamma_e(\alpha-1) - \frac{\kappa(\alpha)\pi}{2}\tan\left(\frac{\kappa(\alpha)\pi}{2}C_L\right) \\
&+\frac{\alpha\pi}{2}\tan\left(\frac{\alpha\pi}{2}C_L\right) - \psi(\alpha C_L)\alpha + \psi(C_L) = 0.
\end{aligned}$$

Using series representations, we re-write g_2 as

$$\begin{aligned}
g_2 =& -\gamma_e(\alpha-1) + \log(1-\epsilon) - \frac{\kappa\pi}{2}\frac{4\kappa C_L}{\pi}\sum_{j=1}^{\infty}\frac{1}{(2j-1)^2-(\kappa C_L)^2} \\
&+\frac{\alpha\pi}{2}\frac{4\alpha C_L}{\pi}\sum_{j=1}^{\infty}\frac{1}{(2j-1)^2-(\alpha C_L)^2} \\
&-\alpha\left(-\gamma_e - \frac{1}{\alpha C_L} + (\alpha C_L)\sum_{j=1}^{\infty}\frac{1}{j(\alpha C_L + j)}\right) \\
&+\left(-\gamma_e - \frac{1}{C_L} + C_L\sum_{j=1}^{\infty}\frac{1}{j(C_L+j)}\right) \\
=&\log(1-\epsilon) - \kappa\sum_{j=1}^{\infty}\left(\frac{1}{2j-1-\kappa C_L} - \frac{1}{2j-1+\kappa C_L}\right) \\
&+\alpha\sum_{j=1}^{\infty}\frac{1}{2j-1-\alpha C_L} - \frac{1}{2j-1+\alpha C_L} \\
&+(1-\alpha)C_L\sum_{j=1}^{\infty}\frac{\alpha C_L + j(1+\alpha)}{j(\alpha C_L+j)(C_L+j)}.
\end{aligned}$$

We first consider $\alpha = 1+\Delta > 1$. In order for $g_2 = 0$ to have a meaningful solution, we must make sure that

$$\frac{-\kappa}{1-\kappa C_L} + \frac{\alpha}{1-\alpha C_L} = \frac{2\Delta}{(1-\kappa C_L)(1-\alpha C_L)} = \frac{2\Delta}{1-2C_L+C_L^2-\Delta^2C_L^2}$$

converges to a finite value as $\alpha \to 1$, i.e., $C_L \to 1$ also. This provides an approximation for C_L when $\alpha > 1$:

$$C_L = 1 - \sqrt{\frac{2\Delta}{-\log(1-\epsilon)}} + o\left(\sqrt{\Delta}\right).$$

Using series representations, we obtain

$$
\begin{aligned}
& C_L\gamma_e(\alpha-1)+\log\alpha+\log\frac{\Gamma\left(\alpha C_L\right)\cos\left(\frac{\pi\alpha C_L}{2}\right)}{\cos\left(\frac{\kappa(\alpha)\pi}{2}C_L\right)\Gamma\left(C_L\right)}\\
=& \log\left(\prod_{s=1}^{\infty}\frac{1+\frac{C_L}{s}}{1+\frac{\alpha C_L}{s}}\exp\left(\frac{\Delta C_L}{s}\right)\prod_{s=0}^{\infty}\frac{1-\frac{\alpha^2C_L^2}{(2s+1)^2}}{1-\frac{\kappa^2C_L^2}{(2s+1)^2}}\right)\\
=& \sum_{s=1}^{\infty}\left(-\frac{\Delta C_L}{s+C_L}+\frac{\Delta C_L}{s}+o\left(\Delta\right)\right)+\log\left(\frac{1-\alpha^2C_L^2}{1-\kappa^2C_L^2}\right)\\
&+\sum_{s=1}^{\infty}\log\frac{1-\frac{\alpha^2C_L^2}{(2s+1)^2}}{1-\frac{\kappa^2C_L^2}{(2s+1)^2}}=-\sqrt{-2\Delta\log(1-\epsilon)}+O\left(\Delta\right).
\end{aligned}
$$

Therefore, for $\alpha > 1$

$$
G_{L,gm}=\frac{\epsilon^2}{-\log(1-\epsilon)-2\sqrt{-2\Delta\log(1-\epsilon)}+o\left(\sqrt{\Delta}\right)}.
$$

Finally, we need to consider $\alpha < 1$. In this case,

$$
\begin{aligned}
g_2 =& \log(1-\epsilon)+\Delta C_L\sum_{j=1}^{\infty}\frac{\alpha C_L+j(1+\alpha)}{j(\alpha C_L+j)(C_L+j)}\\
=& \log(1-\epsilon)+\Delta C_L\left(\sum_{j=1}^{\infty}\frac{1}{j(j+C_L)}+\sum_{j=1}^{\infty}\frac{1}{(1+C_L)^2}\right)+o\left(\Delta\right).
\end{aligned}
$$

Using properties of Riemann's Zeta function and Bernoulli numbers[8, 9.511,9.521.1,9.61]

$$
\begin{aligned}
& \sum_{j=1}^{\infty}\frac{1}{(j+C_L)^2}=-\frac{1}{C_L^2}+\int_0^{\infty}\frac{t\exp(-C_Lt)}{1-\exp(-t)}dt\\
=& -\frac{1}{C_L^2}+\int_0^{\infty}\left(1+\frac{t}{2}+\frac{t^2}{12}+\ldots\right)\exp(-C_Lt)dt=\frac{1}{C_L}+O\left(\frac{1}{C_L^2}\right).
\end{aligned}
$$

Using the integral relation[8, 0.244.1] and treating C_L as a positive integer (which does not affect the asymptotics)

$$
\begin{aligned}
& \sum_{j=1}^{\infty}\frac{1}{j(j+C_L)}=\frac{1}{C_L}\int_0^1\frac{1-t^{C_L}}{1-t}dt\\
=& \frac{1}{C_L}\int_0^1 t^{C_L-1}+t^{C_L-2}+\ldots+1dt=\frac{1}{C_L}\sum_{j=1}^{C_L}\frac{1}{j}\\
=& \frac{1}{C_L}\left(\gamma_e+\log C_L+O\left(C_L^{-1}\right)\right).
\end{aligned}
$$

Thus, the solution to $g_2 = 0$ can be approximated by

$$
\log(1-\epsilon)+\Delta\left(1+\gamma_e+\log C_L\right)+o(\Delta)=0.
$$

Again, using series representations, we obtain

$$
\begin{aligned}
& C_L\gamma_e(\alpha-1)+\log\alpha+\log\frac{\Gamma\left(\alpha C_L\right)}{\Gamma\left(C_L\right)}\\
=& \log\left(\prod_{j=1}^{\infty}\frac{1+\frac{C_L}{j}}{1+\frac{\alpha C_L}{j}}\exp\left(-\frac{\Delta C_L}{j}\right)\right)\\
=& \sum_{j=1}^{\infty}\left(\frac{\Delta C_L}{j+C_L}-\frac{\Delta C_L}{j}+\ldots\right)=-\Delta C_L\left(\gamma_e+\log C_L\right)+\ldots
\end{aligned}
$$

Combining the results, we obtain, when $\alpha < 1$ and $\Delta \to 0$,

$$
G_{L,gm}=\frac{\epsilon^2}{\Delta\left(\exp\left(\frac{-\log(1-\epsilon)}{\Delta}-1-\gamma_e\right)\right)+o\left(\Delta\exp\left(\frac{1}{\Delta}\right)\right)}.
$$

References

[1] C.C. Aggarwal *Data Streams: Models and Algorithms*. Springer, 2007.

[2] N. Alon, Y. Matias, and M. Szegedy. The space complexity of approximating the frequency moments. In *STOC'96*, 20–29.

[3] B. Babcock, S. Babu, M. Datar, R. Motwani, and J. Widom. Models and issues in data stream systems. In *PODS'02*, 1–16.

[4] Z. Bar-Yossef, T. S. Jayram, R. Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *JCSS*, 68(4):702-732, 2004.

[5] J. M. Chambers, C. L. Mallows, and B. W. Stuck. A method for simulating stable random variables. *JASA*, 71(354):340–344, 1976.

[6] J. Feigenbaum, S. Kannan, M. Strauss, and M. Viswanathan. An approximate l_1-difference algorithm for massive data streams. In *FOCS'99*, 501–511.

[7] S. Ganguly and G. Cormode. On estimating frequency moments of data streams. In *APPROX-RANDOM'07*, 479–493.

[8] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 2000.

[9] N. J. A. Harvey, J. Nelson, and K. Onak. Sketching and streaming entropy via approximation theory. In *FOCS'08*.

[10] N. J. A. Harvey, J. Nelson, and K. Onak. Streaming algorithms for estimating entropy. In *ITW'08*.

[11] M. R. Henzinger, P. Raghavan, and S. Rajagopalan. *Computing on Data Streams*. American Mathematical Society, 1999.

[12] P. Indyk. Stable distributions, pseudorandom generators, embeddings, and data stream computation. *Journal of ACM*, 53(3):307–323, 2006.

[13] P. Indyk and D. P. Woodruff. Optimal approximations of the frequency moments of data streams. In *STOC'05*, 202–208.

[14] C. D. Hardin Jr. Skewed stable variables and processes. Technical Report 79, University of North Carolina, 1984.

[15] P. Li. Estimators and tail bounds for dimension reduction in l_α $(0 < \alpha \leq 2)$ using stable random projections. In *SODA'08*, 10 – 19.

[16] P. Li. A very efficient scheme for estimating entropy of data streams using compressed counting. `arxiv.org/PS_cache/arxiv/pdf/0808/0808.1771v2.pdf`, 2008.

[17] P. Li, D. Paul, R. Narasimhan, and J. Cioffi. On the distribution of SINR for the MMSE MIMO receiver and performance analysis. *IEEE Trans. Inform. Theory*, 52(1):271–286, 2006.

[18] S. Muthukrishnan. Data streams: Algorithms and applications. *Foundations and Trends in Theoretical Computer Science*, 1:117–236, 2 2005.

[19] A. Rényi. On measures of information and entropy. In *The 4th Berkeley Symposium on Mathematics, Statistics and Probability 1960*, 547–561, 1961.

[20] M. Saks and X. Sun. Space lower bounds for distance approximation in the data stream model. In *STOC'02*, 360–369.

[21] C. Tsallis. Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52:479–487, 1988.

[22] D. P. Woodruff. Optimal space lower bounds for all frequency moments. In *SODA'04*, 167–175.

[23] H. Zhao, A. Lall, M. Ogihara, O. Spatscheck, J. Wang, and J. Xu. A data streaming algorithm for estimating entropies of OD flows. In *IMC'07*.

[24] V. M. Zolotarev. *One-dimensional Stable Distributions*. American Mathematical Society, 1986.