

ASYMPTOTIC RÉNYI ENTROPIES OF RANDOM WALKS ON GROUPS

KIMBERLY GOLUBEVA, MINGHAO PAN, AND OMER TAMUZ

ABSTRACT. We introduce asymptotic Rényi entropies as a parameterized family of invariants for random walks on groups. These invariants interpolate between various well-studied properties of the random walk, including the growth rate of the group, the Shannon entropy, and the spectral radius. They furthermore offer large deviation counterparts of the Shannon-McMillan-Breiman Theorem. We prove some basic properties of asymptotic Rényi entropies that apply to all groups, and discuss their analyticity and positivity for the free group and lamplighter groups.

1. INTRODUCTION

The Avez entropy (or asymptotic Shannon entropy) of a random walk on a group is an essential tool for understanding its asymptotic properties, and in particular the Furstenberg-Poisson boundary [1, 8]. It is also useful for studying geometric properties of groups; for example, it is always positive for non-amenable groups, and zero for sub-exponential groups.

We introduce asymptotic Rényi entropies of a random walk on a finitely generated group. This is a family of invariants that generalizes the Avez entropy, as well as other useful invariants such as the growth rate of the group and the spectral radius of the walk. Rényi entropies originated in information theory as a general way to quantify randomness, beyond Shannon entropy [14]. They share some (but not all) of the useful properties of the Shannon entropy, including additivity for product measures and monotonicity under push-forwards, which makes them useful in the setting of random walks on groups.¹

Date: June 8, 2024.

This work was supported by a National Science Foundation CAREER award (DMS-1944153). The first author was supported by the Canadian Natural Sciences and Engineering Research Council (NSERC) PGS-D Scholarship [funding reference number 567723].

¹See [4, 12, 13] for axiomatic treatments of Rényi entropies and the related Rényi divergences. The axiomatization in the latter implies that Rényi divergences are the extreme points in the set of all divergences that are additive and monotone under push-forwards. A similar result applies to Rényi entropies.

Let ν be a finitely supported probability distribution on a countable set Ω . For $\alpha \in (0, \infty) \setminus \{1\}$, the α -Rényi entropy of ν is given by

$$H_\alpha(\nu) = \frac{1}{1-\alpha} \log \sum_{\omega \in \Omega} \nu(\omega)^\alpha.$$

Letting $H_1(\nu)$ be the Shannon entropy makes $\alpha \mapsto H_\alpha(\nu)$ a continuous map at $\alpha = 1$. Likewise, letting $H_0(\nu)$ be the logarithm of the size of the support and $H_\infty(\nu) = -\max_\omega \log \nu(\omega)$ extends this map to a continuous one defined on the domain $[0, \infty]$.

Let μ be a finitely supported probability measure on a group G , and denote by $\mu^{(n)}$ the n -fold convolution of μ . We say that μ is non-degenerate if its support generates G as a semigroup. For $\alpha \in [0, \infty]$, the *asymptotic α -Rényi entropy* of the μ -random walk on G is

$$h_\alpha(\mu) = \lim_{n \rightarrow \infty} \frac{1}{n} H_\alpha(\mu^{(n)}).$$

As we explain, this limit exists for every μ . For non-degenerate μ , it is easy to see that $h_0(\mu)$ is the exponential growth rate of G , $h_1(\mu)$ is the Avez entropy, and that for symmetric random walks, $h_\infty(\mu)$ is minus the logarithm of the spectral radius (Claim 2.3).

We begin by establishing some general properties that apply to all finitely supported measures on groups.

Theorem 1. *Let μ be a finitely supported probability measure on a group G , and consider the map $\alpha \mapsto h_\alpha(\mu)$.*

- (1) *For α in $[0, 1]$, $h_\alpha(\mu)$ is continuous and decreasing. It is strictly decreasing for all α such that $h_\alpha(\mu) > h_1(\mu)$.*
- (2) *For $\alpha \in (1, \infty]$, $h_\alpha(\mu)$ is continuous and decreasing.*
- (3) *For symmetric μ and $\alpha \in (2, \infty)$, $h_\alpha(\mu) = \frac{\alpha}{\alpha-1} h_\infty(\mu)$.*

A few observations are in order. Part (1) implies that $h_\alpha(\mu)$ interpolates continuously between the exponential growth rate $h_0(\mu)$ and the Avez entropy $h_1(\mu)$, yielding a non-trivial family of invariants indexed by $\alpha \in [0, 1]$. Parts (1) and (2) together imply that $h_\alpha(\mu)$ is everywhere (weakly) decreasing, and continuous except possibly at $\alpha = 1$. As we shall see, it is possible to have a discontinuity there. In particular, there will be a discontinuity for every symmetric, positive entropy random walk on an amenable group. We also note that by Theorem 3 below, while h_α is continuous on $(1, \infty)$, it is not always twice-differentiable in this range. We do not know if it is always differentiable. For symmetric random walks, part (3) shows that asymptotic Rényi entropies have a trivial form in the range $\alpha \geq 2$. The proof of part (1) uses two convexity properties of Rényi entropies (under re-parameterization), including a novel one which we show in Proposition 2.1.

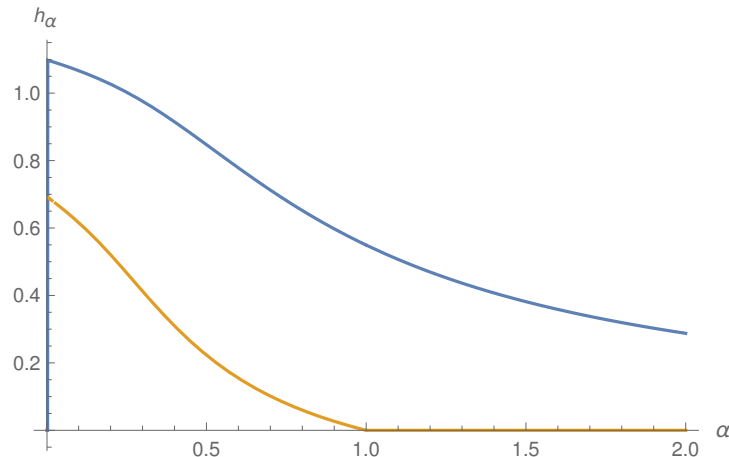


FIGURE 1. Rényi entropies for the simple random walk on the free group with two generators (blue, higher) and the switch-walk-switch walk on the lamplighter group (orange, lower). For both, h_0 is the exponential growth rate of the group, and h_1 is the Avez entropy. Elementary formulas for these graphs are given in (3.3) and (3.8).

Figure 1 shows the asymptotic Rényi entropies of the simple random walk on the free group and the switch-walk-switch walk on the lamplighter group; we calculate these explicitly below. This graph illustrates some properties of asymptotic Rényi entropies that hold more generally: (i) h_α is (weakly) decreasing in α , (ii) for symmetric random walks on non-amenable groups, h_α is positive for all $\alpha \in [0, \infty]$, (iii) for symmetric random walks on amenable groups, h_α vanishes on $(1, \infty]$ (Corollary 2.4). In both of these graphs, h_α is continuous. Theorem 1 shows that this is generally the case, except perhaps at $\alpha = 1$.

The asymptotic min-entropy. The Rényi entropy H_∞ is known as the min-entropy. The asymptotic min-entropy is $h_\infty = \lim_n -\frac{1}{n} \log \max_g \mu^{(n)}(g)$ is the exponential rate of decay of the largest atom in $\mu^{(n)}$. For symmetric random walks it is well-known that this maximum is achieved at the identity (at even times), and that h_∞ is minus the logarithm of the norm of the Markov operator (see, e.g., [11, Proposition 4.4.9]). Thus—for symmetric non-degenerate random walks—it follows from Kesten’s Theorem [10] that h_∞ vanishes if and only if the group is amenable. For non-symmetric random walks, we give an example of a walk on an amenable group for which $h_\infty > 0$ (Claim 3.6). For non-amenable groups we show that $h_\infty > 0$ for every non-degenerate random walk, including the non-symmetric ones (Claim 3.7).

The log-likelihood process. Fix a group G and a finitely supported random walk μ . Let X_1, X_2, \dots be random variables distributed i.i.d. μ , and let $Z_n = X_1 \cdot X_2 \cdots X_n$,

so that Z_n has distribution $\mu^{(n)}$, and Z_1, Z_2, \dots is the μ -random walk on G . Define the *log-likelihood* process by $L_n = -\log \mu^{(n)}(Z_n)$. Then the Shannon entropy of Z_n is the expectation of L_n , and the Shannon-McMillan-Breiman Theorem is the SLLN for the process $(L_n)_n$.

Rényi entropies are, up to a reparametrization, the cumulant generating function of L_n :

$$K_{L_n}(t) = \log \mathbb{E} [e^{tL_n}] = tH_{1-t}(\mu^{(n)}).$$

Hence the asymptotic Rényi entropies of the random walk capture the asymptotics of the moment generating functions of the log-likelihoods. It follows that the Rényi entropies are useful for establishing large deviation bounds for $\frac{1}{n}L_n$. In particular, whenever $(1 - \alpha)h_\alpha(\mu)$ is strictly convex its Legendre transform yields a rate function for large deviations of $\frac{1}{n}L_n$. Likewise, positivity of $h_\alpha(\mu)$ yields a Chernoff bound.

These observations lead us to study the positivity and convexity of asymptotic Rényi entropies. Likewise, we are interested in deviations from continuity and analyticity, as representing phase transitions.

Positivity of asymptotic Rényi entropies. For symmetric random walks on amenable groups, we use standard results to show that $h_\alpha(\mu) = 0$ whenever $\alpha > 1$. For non-amenable groups and non-degenerate μ , regardless if the random walks are symmetric or not, $h_\alpha(\mu) > 0$ for all $\alpha \in [0, \infty]$.

Amenable groups with exponential growth will have $h_0(\mu) > 0$ for all non-degenerate μ . If the Avez entropy $h_1(\mu)$ is also positive, then $h_\alpha(\mu)$ will be positive on $[0, 1]$, since $h_\alpha(\mu)$ is decreasing. However, if $h_1(\mu) = 0$ then it is possible that $h_\alpha(\mu)$ vanishes already for some $\alpha < 1$. We offer this as an open question:

Question. *Does there exist a non-degenerate μ on a group with an exponential growth rate for which $h_\alpha(\mu) = 0$ for some $\alpha \in (0, 1)$?*

While a natural candidate would be the lamplighter group, we show that this is not the case.

Theorem 2. *For any non-degenerate, symmetric, finitely supported probability measure μ on the lamplighter group $G = L \wr \mathbb{Z}$ with lamps in a non-trivial finite group L , $h_\alpha(\mu) > 0$ for all $\alpha \in (0, 1)$.*

The proof of Theorem 2 involves “tilting” μ along the second coordinate, resulting in a random walk that has positive drift, and then relating its Rényi entropies to those of the original, untilted walk.²

²For μ -random walks with positive drift in the second coordinate (the location of the lamplighter) it is known that $h_1(\mu) > 0$ and hence, by the monotonicity of Rényi entropies, the claim also follows. For asymmetric μ with zero drift the claim is likewise true, using a similar proof to that of our Theorem 2.

Continuity and analyticity of asymptotic Rényi entropies. Theorem 1 states that h_α is continuous everywhere except perhaps at $\alpha = 1$. For symmetric random walks on amenable groups, since $h_\alpha = 0$ for all $\alpha > 1$, h_α is discontinuous at $\alpha = 1$ if and only if the Avez entropy h_1 is positive. It is natural to next ask for which random walks on non-amenable groups is h_α continuous at $\alpha = 1$. In particular, one may conjecture that this holds for any random walk on a hyperbolic group (see, e.g., [2] for results in this spirit).

As an example supporting this conjecture, we show that for the simple random walk on the free group, h_α is continuous at $\alpha = 1$. More generally, we show that h_α is (mostly) analytic.

Theorem 3. *Let μ be the simple random walk on a free group with at least two generators. Then $h_\alpha(\mu)$ is analytic on $(0, \infty) \setminus \{2\}$, but not at $\alpha = 2$, where it is not twice-differentiable.*

As another example we study the analyticity of the Rényi entropies of the “switch-walk-switch” (SWS) random walk on the lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}$. Since the Avez entropy vanishes for this walk, we know that $h_\alpha(\mu) = 0$ for all $\alpha \geq 1$.

Theorem 4. *Let μ be the SWS random walk on $\mathbb{Z}_2 \wr \mathbb{Z}$. Then $h_\alpha(\mu)$ is analytic on $(0, \infty) \setminus \{1\}$, but not at $\alpha = 1$.*

We prove Theorems 3 and 4 by explicitly calculating the Rényi entropies and showing that they are elementary functions.

We have so far considered only finitely supported μ . We end this section with a note about non-finitely supported μ . For such μ , $H_0(\mu) = \infty$, and thus $h_0(\mu) = \infty$. Nevertheless, it is still possible that $h_\alpha(\mu) < \infty$ for some $\alpha > 0$. For example, infinitely supported random walks with finite Shannon entropy are important in the study of groups of subexponential growth (see, e.g., [5, 6]). For $\alpha > 1$, $h_\alpha(\mu)$ is finite for any μ , and so asymptotic Rényi entropies might provide a tool to study random walks with heavy tails. We leave this question for future study.

2. PRELIMINARIES

2.1. Rényi entropy. Let ν be a finitely supported probability distribution on a countable set Ω . For $\alpha \in (0, 1) \cup (1, \infty)$, the α -Rényi entropy of ν is given by

$$H_\alpha(\nu) = \frac{1}{1 - \alpha} \log \sum_{\omega \in \Omega} \nu(\omega)^\alpha.$$

For $\alpha \in \{0, 1, \infty\}$ it is given by

$$\begin{aligned} H_0(\nu) &= \log |\{\omega : \nu(\omega) > 0\}| \\ H_1(\nu) &= \sum_{\omega \in \Omega} \nu(\omega) \log \frac{1}{\nu(\omega)} \\ H_\infty(\nu) &= \min_{\omega \in \Omega} \log \frac{1}{\nu(\omega)}. \end{aligned}$$

Hence H_0 is the logarithm of the size of the support, H_1 is the Shannon entropy, and H_∞ is minus the log of the mass of the largest atom. Under this definition, it is well known (see, e.g., [13, pp. 50–51]) or immediate that

- (i) The map $\alpha \mapsto H_\alpha(\nu)$ is continuous and (weakly) decreasing. If ν is not the uniform distribution on a subset of Ω then it is strictly decreasing.
- (ii) For every $\alpha \in [0, \infty]$ it holds that $H_\alpha(\nu_1 \times \nu_2) = H_\alpha(\nu_1) + H_\alpha(\nu_2)$.
- (iii) For every map $f: \Omega \rightarrow \Omega'$ and every $\alpha \in [0, \infty]$ it holds that $H_\alpha(f_*\nu) \leq H_\alpha(\nu)$.

Define the *log-likelihood* random variable $L: \Omega \rightarrow \mathbb{R}$ by $L(\omega) = -\log \nu(\omega)$. Let $K_\nu: \mathbb{R} \rightarrow \mathbb{R}$ be the cumulant generating function of L . That is, let

$$K_\nu(t) = \log \mathbb{E} [e^{tL}] = \log \sum_{\omega} e^{tL(\omega)} \nu(\omega).$$

Then for $\alpha \in (0, \infty)$,

$$(2.1) \quad K_\nu(1 - \alpha) = (1 - \alpha)H_\alpha(\nu).$$

Note that this implies that $(1 - \alpha)H_\alpha(\nu)$ is convex. As we show next, $H_\alpha(\nu)$ admits another re-parameterization which makes it convex for $\alpha \in (0, 1)$. A similar version for $\alpha \in (1, \infty)$ has been observed in [3].

Proposition 2.1. *For any finitely supported measure ν , $\beta \mapsto H_{1+\frac{1}{\beta}}(\nu)$ is convex for $\beta \in (-\infty, -1)$.*

Proof. Let ν be a finitely supported probability measure on a set Ω . For $f: \Omega \rightarrow \mathbb{R}$ and $p < 0$, we denote ℓ^p norms by $\|f\|_p^p := \sum_{\omega} |f(\omega)|^p \nu(\omega)$, where the sum is taken over all ω such that $f(\omega) \neq 0$.

Fixing $p_1, p_2 < 0$ and $\theta \in (0, 1)$, we let $p < 0$ be such that

$$\frac{1}{p} = \frac{\theta}{p_1} + \frac{1 - \theta}{p_2},$$

so that

$$\frac{\theta p}{p_1} + \frac{(1 - \theta)p}{p_2} = 1.$$

Then,

$$\|f\|_p^p = \sum_{\omega} |f(\omega)|^{\theta p} |f(\omega)|^{(1-\theta)p} \nu(\omega) \leq \|f^{\theta p}\|_{\frac{p_1}{\theta p}} \|f^{(1-\theta)p}\|_{\frac{p_2}{(1-\theta)p}} = \|f\|_{p_1}^{\theta p} \|f\|_{p_2}^{(1-\theta)p},$$

where the inequality follows from an application of the standard Hölder's inequality.

It then follows that

$$p \log \|f\|_p \leq \theta p \log \|f\|_{p_1} + (1-\theta)p \log \|f\|_{p_2},$$

and therefore

$$\log \|f\|_p \geq \theta \log \|f\|_{p_1} + (1-\theta) \log \|f\|_{p_2}.$$

Since $\frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}$, this shows that $p \mapsto \log \|f\|_{\frac{1}{p}}$ is concave for $p < 0$.

Now let $f: \Omega \rightarrow \mathbb{R}$ be given by $f(\omega) = \nu(\omega)$. Then

$$\log \|f\|_{\frac{1}{p}} = p \log \sum_{\omega} \nu(\omega)^{1+\frac{1}{p}}.$$

Thus,

$$H_{1+\frac{1}{\beta}}(\nu) = -\beta \log \sum_{\omega} \nu(\omega)^{1+\frac{1}{\beta}} = -\log \|f\|_{\frac{1}{\beta}}$$

is a convex function for $\beta \in (-\infty, -1)$. □

We end this section with another simple observation.

Lemma 2.2. *For $\alpha > 1$ it holds that*

$$H_{\alpha}(\nu) \leq \frac{\alpha}{\alpha-1} H_{\infty}(\nu).$$

Proof. Since $\alpha > 1$, we note that the factor $1-\alpha$ is negative, and hence

$$H_{\alpha}(\nu) = \frac{1}{1-\alpha} \log \sum_{\omega} \nu(\omega)^{\alpha} \leq \frac{1}{1-\alpha} \log \max_{\omega} \nu(\omega)^{\alpha},$$

which, by the definition of H_{∞} , is equal to $\frac{\alpha}{\alpha-1} H_{\infty}(\nu)$. □

2.2. Random walks on groups. Let G be a finitely generated discrete group, and let μ be a finitely supported probability measure on G . Denote convolution by $*$, and the n -fold convolution of μ with itself by $\mu^{(n)}$. We say that μ is symmetric if $\mu(g) = \mu(g^{-1})$. We say that μ is non-degenerate if the support of μ generates G as a semi-group; equivalently, the μ -random walk on G is an irreducible Markov chain.

For $\alpha \in [0, \infty]$, define

$$h_{\alpha}(\mu) = \lim_n \frac{1}{n} H_{\alpha}(\mu^{(n)}).$$

We refer to the family $(h_\alpha)_\alpha$ as invariants, since if $\pi: G \rightarrow H$ is a group isomorphism that maps the probability measure μ on G to the probability measure ν on H , then $h_\alpha(\mu) = h_\alpha(\nu)$.

It follows from property (ii) of H_α that $H_\alpha(\mu^{(n)} \times \mu^{(m)}) = H_\alpha(\mu^{(n)}) + H_\alpha(\mu^{(m)})$. And since $\mu^{(n+m)} = \mu^{(n)} * \mu^{(m)}$, it follows from property (iii) of H_α that $H_\alpha(\mu^{(n+m)}) \leq H_\alpha(\mu^{(n)}) + H_\alpha(\mu^{(m)})$. Thus the map $n \mapsto H_\alpha(\mu^{(n)})$ is subadditive, and so the limit above exists and is finite. Moreover, this limit is equal to the infimum of $\frac{1}{n}H_\alpha(\mu^{(n)})$. Thus, for a given μ , $h_\alpha(\mu)$ is upper semi-continuous.

By definition,

$$h_1(\mu) = \lim_n \frac{1}{n} \sum_g \mu^{(n)}(g) \log \frac{1}{\mu^{(n)}(g)}$$

is the random walk entropy, or Avez entropy. Similarly, $h_0(\mu)$ is the exponential growth rate of the group generated by the support of μ :

$$h_0(\mu) = \lim_n \frac{1}{n} H_0(\mu^{(n)}) = \lim_n \frac{1}{n} \log |\{g : \mu^{(n)}(g) > 0\}| = \lim_n \frac{1}{n} \log |B_\mu(n)|.$$

Here $B_\mu(n)$ is the ball of radius n with respect to the word metric defined by the generating set given by the support of μ .

The proof of the next claim is standard; see, e.g., [11, Exercise 4.4.5], where it is shown that $H_\infty(\mu^{(2n)}) = -\log \mu^{(2n)}(e)$.

Claim 2.3. *Let μ be a finitely supported symmetric measure. Then $h_\infty(\mu)$ is the logarithm of the inverse of the spectral radius:*

$$h_\infty(\mu) = \lim_{n \rightarrow \infty} \frac{1}{2n} \log \frac{1}{\mu^{(2n)}(e)}.$$

There is an immediate corollary.

Corollary 2.4. *Let G be a finitely generated amenable group, and let μ be a finitely supported, symmetric, non-degenerate probability measure on G . Then $h_\alpha(\mu) = 0$ for all $\alpha > 1$.*

Proof. By Kesten's Theorem [9, 10], the assumptions on μ and the amenability of G imply that the spectral radius of the random walk is 1. Hence, by Claim 2.3, $h_\infty(\mu) = 0$.

By Lemma 2.2, if $\alpha > 1$ then $H_\alpha \leq \frac{\alpha}{\alpha-1} H_\infty$ and so we have that $h_\alpha(\mu) \leq \frac{\alpha}{\alpha-1} h_\infty(\mu) = 0$. \square

3. PROOFS

Proof of Theorem 1. Recall from (2.1) that for $\alpha \in (0, \infty)$, it holds that

$$(1 - \alpha)H_\alpha(\nu) = K_\nu(1 - \alpha).$$

The cumulant generating function $K_\nu(t)$ of $L(\omega) := -\log \nu(\omega)$ is a convex function for $t \in \mathbb{R}$. For a fixed t , it is straightforward to check that the following hold:

- (1) Additivity: $K_{\nu \times \nu'}(t) = K_\nu(t) + K_{\nu'}(t)$.
- (2) Monotonicity under push-forwards: $K_{f_*\nu}(t) \leq K_\nu(t)$.

By Fekete's lemma, the sequence of functions $t \mapsto \frac{1}{n}K_{\mu^{(n)}}(t)$ converges pointwise. The limit, which we denote by $k_\mu(t)$ is a convex function as well, and thus continuous. For $\alpha \geq 0$, we have $k_\mu(1-\alpha) = (1-\alpha)h_\alpha(\mu)$. Hence the restriction of $h_\alpha(\mu)$ to $[0, \infty) \setminus \{1\}$ is continuous. The function $\alpha \mapsto h_\alpha(\mu) = \inf_n \frac{1}{n}H_\alpha(\mu^{(n)})$ is upper semi-continuous as an infimum of continuous functions. Since $H_\alpha(\nu)$ is decreasing for any ν , so is $h_\alpha(\mu)$. Upper semi-continuity and monotonicity imply that $h_\alpha(\mu)$ is càglàd. Thus we can conclude that it is left continuous at 1.

To prove the continuity at ∞ , we note that $H_\alpha(\nu)$ is decreasing so that $H_\infty(\nu) = \inf_{\alpha \in (0, \infty)} H_\alpha(\nu)$. Therefore,

$$h_\infty(\mu) = \inf_n \frac{1}{n}H_\infty(\mu^{(n)}) = \inf_n \inf_{\alpha \in (0, \infty)} \frac{1}{n}H_\alpha(\mu^{(n)}) = \inf_{\alpha \in (0, \infty)} h_\alpha(\mu).$$

As $h_\alpha(\mu)$ is decreasing in α , we obtain continuity at ∞ .

We next show that $h_\alpha(\mu)$ is strictly decreasing on $[0, 1]$ whenever it is larger than $h_1(\mu)$. By Proposition 2.1, $H_{1+\frac{1}{\beta}}(\mu^{(n)})$ is a convex function for $\beta \in (-\infty, -1)$. Passing to the limit,

$$h_{1+\frac{1}{\beta}}(\mu) = \lim_n \frac{1}{n}H_{1+\frac{1}{\beta}}(\mu^{(n)})$$

is also a convex function for $\beta \in (-\infty, -1)$.

Denote $\bar{h}(\beta) = h_{1+\frac{1}{\beta}}(\mu)$, so that \bar{h} is convex for $\beta \in (-\infty, -1)$. Since the map $\beta \mapsto 1 + 1/\beta$ is strictly decreasing and $h_\alpha(\mu)$ is decreasing, \bar{h} is increasing. Since it is convex, it must be strictly increasing whenever it is not equal to its infimum. Hence, $h_\alpha(\mu) = \bar{h}(1/(\alpha - 1))$ is strictly decreasing in $(0, 1)$ whenever it is not equal to its infimum, $h_1(\mu)$.

We have so far established the first and second part of the claim. For the third, suppose that μ is symmetric, and let X_1, X_2, \dots be i.i.d. random variables taking value in G with law μ . Let $Z_n = X_1 \cdots X_n$ be the μ -random walk, so that the law of Z_n is $\mu^{(n)}$. Let Z'_1, Z'_2, \dots be an additional, independent μ -random walk. Then, since the random walk is symmetric,

$$h_2(\mu) = -\lim_n \frac{1}{n} \log \sum_g \mu^{(n)}(g)^2 = -\lim_n \frac{1}{n} \log \mathbb{P}[Z_n = Z'_n] = -\lim_n \frac{1}{n} \log \mu^{(2n)}(e).$$

Thus

$$h_2(\mu) = -\lim_n \frac{1}{n} \log \mu^{(2n)}(e) = -2 \lim_n \frac{1}{2n} \log \mu^{(2n)}(e) = 2h_\infty(\mu),$$

because for symmetric random walks the maximum probability is achieved at the identity. This shows the claim for $\alpha = 2$.

Comparing this to the definition of $h_2(\mu)$, we see that the entire sum over $g \in G$ is dominated by $g = e$,

$$h_2(\mu) = -\lim_n \frac{1}{2n} \log \sum_g (\mu^{(2n)}(g))^2 = -\lim_n \frac{1}{2n} \log (\mu^{(2n)}(e))^2.$$

The same holds for $\alpha > 2$. We note that

$$\begin{aligned} \lim_n \frac{1}{2n} \log \mu^{(2n)}(e)^\alpha &= \lim_n \frac{1}{2n} \log \mu^{(2n)}(e)^2 + \lim_n \frac{1}{2n} \log \mu^{(2n)}(e)^{\alpha-2} \\ &= \lim_n \frac{1}{2n} \log \sum_g \mu^{(2n)}(g)^2 + \lim_n \frac{1}{2n} \log \mu^{(2n)}(e)^{\alpha-2} \\ &= \lim_n \frac{1}{2n} \log \sum_g \mu^{(2n)}(g)^2 \mu^{(2n)}(e)^{\alpha-2} \\ &\geq \lim_n \frac{1}{2n} \log \sum_g \mu^{(2n)}(g)^\alpha \\ &\geq \lim_n \frac{1}{2n} \log \mu^{(2n)}(e)^\alpha \end{aligned}$$

where the second equality follows from the $\alpha = 2$ case, and the second last inequality is because the maximum probability is achieved at the identity. Therefore, all the inequalities above are in fact equalities and

$$\alpha h_\infty(\mu) = -\lim_n \frac{1}{2n} \log \mu^{(2n)}(e)^\alpha = -\lim_n \frac{1}{2n} \log \sum_g \mu^{(2n)}(g)^\alpha = (\alpha - 1)h_\alpha(\mu).$$

□

3.1. The free group. Let $G = \mathbb{F}_d$ be the free group with $d \geq 2$ generators and let μ be the uniform distribution on the set of d generators and their inverses. As above, let X_1, X_2, \dots be i.i.d. random variables taking value in G with law μ . Let $Z_n = X_1 \cdots X_n$ be the μ -random walk, and let $D_n = |Z_n|$ be the distance between Z_n and the origin. Then D_1, D_2, \dots is a Markov chain, where $\mathbb{P}[D_n = 1 | D_{n-1} = 0] = 1$, and for $r > 0$, $\mathbb{P}[D_n = r - 1 | D_{n-1} = r] = \frac{1}{2d}$ and $\mathbb{P}[D_n = r + 1 | D_{n-1} = r] = 1 - \frac{1}{2d}$.

By the symmetry of the random walk, if $|g| = |h|$ then $\mu^{(n)}(g) = \mu^{(n)}(h)$. The number of elements $g \in \mathbb{F}_d$ with $|g| = k$ is equal to $2d \cdot (2d - 1)^{k-1}$. Hence

$$(3.1) \quad \mu^{(n)}(g) = \mathbb{P}[Z_n = g] = \frac{\mathbb{P}[D_n = k]}{2d \cdot (2d - 1)^{k-1}}.$$

It follows that when $\alpha \neq 1$,

$$\begin{aligned} H_\alpha(\mu^{(n)}) &= \frac{1}{1-\alpha} \log \sum_g \mathbb{P}[Z_n = g]^\alpha \\ &= \frac{1}{1-\alpha} \log \sum_{k=0}^n \sum_{|g|=k} \left(\frac{\mathbb{P}[D_n = k]}{2d \cdot (2d-1)^{k-1}} \right)^\alpha \\ &= \frac{1}{1-\alpha} \log \sum_{k=0}^n \mathbb{P}[D_n = k]^\alpha (2d \cdot (2d-1)^{k-1})^{1-\alpha}, \end{aligned}$$

and so

$$h_\alpha(\mu) = \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \sum_{k=0}^n \mathbb{P}[D_n = k]^\alpha (2d \cdot (2d-1)^{k-1})^{1-\alpha}.$$

And since the sum is bounded below by its maximum and above by $n+1$ times its maximum,

(3.2)

$$\begin{aligned} h_\alpha(\mu) &= \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \max_{k \in \{0, \dots, n\}} \mathbb{P}[D_n = k]^\alpha (2d-1)^{(1-\alpha)k} \\ &= \lim_n \frac{1}{n} \frac{1}{1-\alpha} \max_{k \in \{0, \dots, n\}} \alpha \log \mathbb{P}[D_n = k] + (1-\alpha)k \log(2d-1) \\ &= \lim_n \frac{1}{n} \frac{1}{1-\alpha} \max_{k \in \{\lfloor n/2 \rfloor, \dots, n\}} \alpha \log \mathbb{P}[D_n = 2k-n] + (1-\alpha)(2k-n) \log(2d-1), \end{aligned}$$

where the last equality is just a change of variables.

Let E_1, E_2, \dots be a random walk on \mathbb{Z} with $\mathbb{P}[E_{n+1} = E_n + 1] = 1 - \frac{1}{2d}$ and $\mathbb{P}[E_{n+1} = E_n - 1] = \frac{1}{2d}$. The only difference between E_n and D_n is that D_n is reflected at 0 while E_n is allowed to travel to the left of 0. We then try to understand the probability distribution of D_n from that of E_n .

Lemma 3.1. *For any $k \geq 0$ and any n , we have*

$$\frac{1}{n} \mathbb{P}[E_n = k] \leq \mathbb{P}[D_n = k] \leq 2 \mathbb{P}[E_n = k].$$

Proof. We begin by considering the case where $k = 0$. In particular, we will prove that for any n ,

$$\frac{1}{n} \mathbb{P}[E_n = 0] \leq \mathbb{P}[D_n = 0] \leq \frac{2d-1}{2d-2} \mathbb{P}[E_n = 0].$$

For this argument, we may assume that n is even (since, if n were odd, then both sides of the equation above would equal zero).

We define a coupling of D_i and E_i by first sampling E_i and then setting

$$\begin{aligned} D_{i+1} - D_i &= E_{i+1} - E_i & \text{if } D_i \geq 1 \\ D_{i+1} &= D_i + 1 & \text{if } D_i = 0. \end{aligned}$$

From above, we see that the random walks are synchronous but their positions may be shifted upon returning to 0, which gives $E_n \leq D_n$. Hence,

$$\mathbb{P}[D_n = 0] \leq \mathbb{P}[E_n \leq 0].$$

Then for $n = 2m$ even, we have

$$\begin{aligned} \mathbb{P}[E_{2m} \leq 0] &= \sum_{i=0}^m \mathbb{P}[E_{2m} = -2i] \\ &= \sum_{i=0}^m \left(\frac{2d-1}{2d}\right)^{m-i} \left(\frac{1}{2d}\right)^{m+i} \binom{2m}{m-i} \\ &\leq \sum_{i=0}^m \left(\frac{2d-1}{2d}\right)^{m-i} \left(\frac{1}{2d}\right)^{m+i} \binom{2m}{m} \\ &\leq \frac{1}{1 - \frac{1}{2d-1}} \left(\frac{2d-1}{2d}\right)^m \left(\frac{1}{2d}\right)^m \binom{2m}{m} \\ &= \frac{2d-1}{2d-2} \mathbb{P}[E_{2m} = 0] \end{aligned}$$

Thus, we have shown that when $k = 0$,

$$\mathbb{P}[D_n = 0] \leq \frac{2d-1}{2d-2} \mathbb{P}[E_n = 0].$$

For the lower bound, we first note that if E_n remains nonnegative then $D_n = E_n$, which follows from the coupling defined above. Therefore,

$$\{E_{2m} = 0 \text{ and } E_i \geq 0 \text{ for } i < 2m\} \subset \{D_{2m} = 0\}$$

and so $\mathbb{P}[E_{2m} = 0 \text{ and } E_i \geq 0 \text{ for } i < 2m] \leq \mathbb{P}[D_{2m} = 0]$. To show the desired lower bound, note that by the Bertrand's Ballot Theorem,

$$\mathbb{P}[D_{2m} = 0] \geq \mathbb{P}[E_{2m} = 0 \text{ and } E_i \geq 0 \text{ for } i < 2m] = \frac{1}{m+1} \mathbb{P}[E_{2m} = 0].$$

For $k > 0$, note that by again using the Bertrand's Ballot Theorem we have that

$$\mathbb{P}[D_n = k] \geq \mathbb{P}[E_n = k \text{ and } E_i > 0 \text{ for } i < n] = \frac{k}{n} \mathbb{P}[E_n = k] \geq \frac{1}{n} \mathbb{P}[E_n = k].$$

This shows the lower bound. For the upper bound, we will use induction on n to prove that for any k

$$\mathbb{P}[D_n = k] \leq 2\mathbb{P}[E_n = k].$$

The base case $n = 0$ is trivial. For the inductive step, we suppose that the above inequality holds for some n and all k . For $k > 1$, we have

$$\begin{aligned} \mathbb{P}[D_{n+1} = k] &= \frac{2d-1}{2d} \mathbb{P}[D_n = k-1] + \frac{1}{2d} \mathbb{P}[D_n = k+1] \\ &\leq 2 \left(\frac{2d-1}{2d} \mathbb{P}[E_n = k-1] + \frac{1}{2d} \mathbb{P}[E_n = k+1] \right) \\ &= 2\mathbb{P}[E_{n+1} = k]. \end{aligned}$$

For $k = 1$, we have

$$\begin{aligned} \mathbb{P}[D_{n+1} = 1] &= \mathbb{P}[D_n = 0] + \frac{1}{2d} \mathbb{P}[D_n = 2] \\ &\leq \frac{2d-1}{2d-2} \mathbb{P}[E_n = 0] + \frac{1}{2d} \cdot 2\mathbb{P}[E_n = 2] \\ &\leq 2 \left(\frac{2d-1}{2d} \mathbb{P}[E_n = 0] + \frac{1}{2d} \mathbb{P}[E_n = 2] \right) \\ &= 2\mathbb{P}[E_{n+1} = 1]. \end{aligned}$$

This completes the induction step. □

The new random walk E_n may be understood very well. First note that,

$$\mathbb{P}[E_n = 2k - n] = \binom{n}{k} \left(\frac{2d-1}{2d} \right)^k \left(\frac{1}{2d} \right)^{n-k}.$$

By a standard estimate,

$$\frac{1}{n+1} e^{nH(k/n)} \leq \binom{n}{k} \leq e^{nH(k/n)},$$

where $H(p) = -p \log p - (1-p) \log(1-p)$, and so

$$\log \mathbb{P}[E_n = 2k - n] = -n \log(2d) + nH(k/n) + k \log(2d-1) + o(n).$$

Furthermore, by Lemma 3.1, for $k \geq 0$, $\log \mathbb{P}[D_n = k] = \log \mathbb{P}[E_n = k] + o(n)$. Hence

$$\frac{1}{n} \log \mathbb{P}[D_n = 2k - n] = -\log(2d) + H(k/n) + k/n \log(2d-1) + o(1)$$

for $k \geq n/2$. Inserting this into (3.2) yields

$$\begin{aligned} h_\alpha(\mu) &= \lim_n \frac{1}{1-\alpha} \max_{k \in \{\lfloor n/2 \rfloor, \dots, n\}} \alpha(-\log(2d) + H(k/n) + k/n \log(2d-1)) \\ &\quad + (1-\alpha)(2k/n - 1) \log(2d-1). \end{aligned}$$

The expression being maximized is a function of k/n , which we denote by $f_\alpha: [1/2, 1] \rightarrow \mathbb{R}$:

$$f_\alpha(p) = \alpha(-\log(2d) + H(p) + p \log(2d - 1)) + (1 - \alpha)(2p - 1) \log(2d - 1).$$

Using this construction, we may now prove Theorem 3.

Proof of Theorem 3. For $\alpha \neq 1$, from the notation established above, we have

$$h_\alpha(\mu) = \frac{1}{1 - \alpha} \lim_n \max_{k \in \{\lceil n/2 \rceil, \dots, n\}} f_\alpha(k/n).$$

Since f_α is continuous,

$$h_\alpha(\mu) = \frac{1}{1 - \alpha} \max_{p \in [1/2, 1]} f_\alpha(p).$$

Since f_α is furthermore differentiable and concave, this maximum is achieved either at the end points of the interval, or else at the unique p_α^* at which the derivative of f_α vanishes. A simple calculation shows that this is given by

$$p_\alpha^* = \frac{(2d - 1)^{2/\alpha}}{(2d - 1) + (2d - 1)^{2/\alpha}}$$

for $\alpha \in (0, 2) \setminus \{1\}$. For these α , we have

$$(3.3) \quad h_\alpha(\mu) = \frac{1}{1 - \alpha} f_\alpha(p_\alpha^*),$$

where we recall that

$$f_\alpha(p) = \alpha(-\log(2d) + H(p) + p \log(2d - 1)) + (1 - \alpha)(2p - 1) \log(2d - 1).$$

This can be written out as an elementary (but unwieldy) function of α , which we omit. Importantly, this implies that $h_\alpha(\mu)$ is an analytic function of α on $(0, 2) \setminus \{1\}$. Moreover, $f_\alpha(p_\alpha^*)$, viewed as a complex function, is holomorphic around a neighborhood of $\alpha = 1$. Some computations show that $f_\alpha(p_\alpha^*)$ has a zero of order 1 at $\alpha = 1$ so $h_\alpha(\mu) = \frac{1}{1 - \alpha} f_\alpha(p_\alpha^*)$ is a meromorphic function whose singularity at $\alpha = 1$ can be removed. In other words, $h_\alpha(\mu)$ is real analytic at $\alpha = 1$.

For $\alpha \geq 2$, the Rényi entropy enters a difference phase $h_\alpha(\mu) = \frac{\alpha}{\alpha - 1} h_\infty(\mu)$. Comparing this with (3.3), we see that $h_\alpha(\mu)$ is first differentiable, but not second differentiable, at $\alpha = 2$.

□

3.2. The lamplighter group. Let $G = L \wr \mathbb{Z}$ be the lamplighter group on \mathbb{Z} with lamps in some non-trivial finite group L . We denote an element of G by $g = (f, z)$ where $f: \mathbb{Z} \rightarrow L$ has finite support, and $z \in \mathbb{Z}$. Denote by $\pi: G \rightarrow \mathbb{Z}$ be the group homomorphism $\pi(f, z) := z$ and denote by $\tau(f, z) := f$ the projection to the lamp configurations. Let μ be a symmetric, finitely supported measure on G .

We prove Theorem 2 by analyzing “tilted” versions of the μ -random walk. For $t \geq 0$ define the tilted measure μ_t by

$$\mu_t(g) = C_t^{-1} \mu(g) e^{t\pi(g)}$$

where

$$C_t = \sum_{g' \in G} \mu(g') e^{t\pi(g')}$$

is the moment generating function of $\pi_*\mu$, evaluated at t . Note that tilting commutes with convolution, so that $\mu_t^{(k)}$ is well-defined.

Note that $C_t'(0) = \sum_h \mu(h) \pi(h) = 0$ because μ is symmetric, and hence the expectation of $\pi_*\mu$ is equal to zero. The second derivative of C_t at zero is the variance $v > 0$ of $\pi_*\mu$, and hence $C_t = 1 + vt^2 + o(t^2)$.

Let Z_1, Z_2, \dots be a μ_t -random walk, and let $Q_n^t = \{\pi(Z_1), \dots, \pi(Z_n)\}$ be the set of locations visited by the lamplighter.

Lemma 3.2. *There is a constant $q > 0$ such that, for all $t > 0$ small enough, $\liminf_n \frac{1}{n} |Q_n^t| \geq qt$ almost surely.*

Proof. The tilted random walk has a positive drift of

$$\begin{aligned} \Delta_t &:= C_t^{-1} \sum_g \mu(g) \exp(t\pi(g)) \pi(g) \\ &= C_t^{-1} \left(\sum_{g:\pi(g)<0} \mu(g) \exp(t\pi(g)) \pi(g) + \sum_{\pi(g)>0} \mu(g) \exp(t\pi(g)) \pi(g) \right) \\ &\geq C_t^{-1} \left(\sum_{g:\pi(g)<0} \mu(g) \pi(g) + \sum_{\pi(g)>0} \mu(g) (1 + t\pi(g)) \pi(g) \right) \\ &\geq C_t^{-1} t \sum_{\pi(g)>0} \mu(g) (\pi(g))^2. \end{aligned}$$

As $C_t = 1 + o(t)$, we have

$$\Delta_t > \beta t$$

for $t > 0$ small enough, where we set $\beta := \frac{1}{2} \sum_{\pi(g)>0} \mu(g) (\pi(g))^2$. Note that $\beta > 0$.

Since μ has finite support, there exists some $M > 0$ such that $|\pi(g)| < M$ for any g in the support of μ . Then the step size under μ_t is also bounded by M . Let Z_1, Z_2, \dots be a μ_t -random walk. For any $\varepsilon > 0$, by law of large numbers, we have

$$\liminf_n \mathbb{P}(\pi(Z_n) > (\Delta_t - \varepsilon)n) = 1.$$

Note that in order for the lamplighter to arrive at some position N at time n , the lamplighter must have visited at least N/M many positions on the way. Thus

$$|Q_n^t| \geq \frac{\pi(Z_n)}{M}.$$

Therefore, for $t > 0$ small enough,

$$\liminf_n \frac{1}{n} |Q_n^t| \geq \frac{\Delta_t - \varepsilon}{M} > \frac{\beta t - \varepsilon}{M}.$$

Set $q := \frac{\beta}{M}$. As ε is arbitrary, we have

$$\liminf_n \frac{1}{n} |Q_n^t| \geq qt.$$

□

For the tilted random walk, the lamp status at the origin is a non-trivial tail random variable so the random walk has positive asymptotic Shannon entropy: $h_1(\mu_t) > 0$. The next proposition shows that this entropy grows at least linearly with t , for small t .

Proposition 3.3. *Suppose μ is non-degenerate. There is a constant $c > 0$ such that $h_1(\mu_t) \geq ct$ for $t > 0$ small enough.*

Proof. We first show that we can assume that $\mu(e) > 0$. Let $\eta = \frac{1}{2}\mu + \frac{1}{2}\delta_e$, where δ_e is the point mass at the identity of G . Define the tilted measures η_t similarly to the definition of μ_t . Then

$$\eta_t(g) = \frac{1}{\frac{1}{2}C_t + \frac{1}{2}} \left(\frac{1}{2}\mu(g)e^{t\pi(g)} + \frac{1}{2}\delta_e(g) \right) = \frac{1}{\frac{1}{2}C_t + \frac{1}{2}} \left(\frac{1}{2}C_t\mu_t(g) + \frac{1}{2}\delta_e(g) \right),$$

and so

$$\eta_t = \alpha_t\mu_t + (1 - \alpha_t)\delta_e$$

for some α_t that tends continuously to $1/2$ as t tends to zero.

Since $h_1(\alpha\mu_t + (1 - \alpha)\delta_e) = \alpha h_1(\mu_t)$, it follows that if $h_1(\eta_t) > ct$ then $h_1(\mu_t) > ct$ for all t small enough. It hence suffices to prove the claim for η , so that it follows for μ . We thus assume without loss of generality that $\mu(e) > 0$.

Denote by $f^\ell: \mathbb{Z} \rightarrow L$ the function given $f^\ell(0) = \ell$ and $f^\ell(z) = e$ for all $z \neq 0$, and where, by slight notation overloading, e is the identity of L .

Since μ is non-degenerate, and since $\mu(e) > 0$, there is some k large enough and $\varepsilon > 0$ such that $\mu^{(k)}(f^\ell, 0) > 2\varepsilon$ for all $\ell \in L$. We hence again assume without loss of generality that this already holds for μ , since $h(\mu_t^{(k)}) = kh(\mu_t)$. For all t small enough we will have that $\mu_t(f^\ell, 0) > \varepsilon$ for all $\ell \in L$. Let ν be the uniform distribution over $\tilde{L} = \{(f^\ell, 0) : \ell \in L\}$. Then we can write $\mu_t = \varepsilon\nu + (1 - \varepsilon)\bar{\mu}_t$ where

$$\bar{\mu}_t = \frac{\mu_t - \varepsilon\nu}{1 - \varepsilon}$$

is a probability measure.

Let $\bar{X}_1, \bar{X}_2, \dots$ be i.i.d. random variables with distribution $\bar{\mu}_t$. Let W_n be i.i.d. random variables with distribution ν . Let B_1, B_2, \dots be i.i.d. Bernoulli random variables with $\mathbb{P}[B_n = 1] = \varepsilon$. Let

$$X_n = \begin{cases} \bar{X}_n & \text{if } B_n = 0 \\ W_n & \text{if } B_n = 1. \end{cases}$$

Then X_n has distribution μ_t . Let $Z_n = X_1 \cdot X_2 \cdots X_n$ be a μ_t -random walk on G .

We would like to show that $H_1(Z_n) \geq nct$ for some c , all t small enough and all n large enough.

Let

$$P_n = \{z \in Z : \exists k \leq n, \pi(Z_k) = z, B_k = 1\}.$$

This is the random set of locations at which the lamplighter was at times in which $B_k = 1$. Since $\pi(W_n) = 0$, the probability that $\pi(Z_n) \in P_n$ is at least ε . Hence $\mathbb{E}[|P_n|] \geq \varepsilon \mathbb{E}[|Q_n^t|]$, and by Lemma 3.2, there is a constant c such that for all t small enough, $\mathbb{E}[|P_n|] \geq nct$ for all n large enough.

Denote $F_n = \tau(Z_n)$ the lamp configuration at time n . Suppose for a moment that $\bar{X}_1, \dots, \bar{X}_n$ and B_1, \dots, B_n are given. As $\pi(W_k) = 0$, we will know the location of the lamplighter at all times up to n and so we can deduce P_n . We will also know the state of the lamps outside of P_n . On P_n , the conditional distribution of lamps is i.i.d. uniform, since W_k are uniform. Hence, the conditional entropy of F_n is exactly $|P_n| \log |L|$ and

$$\begin{aligned} H_1(Z_n) &\geq H_1(Z_n | \bar{X}_1, \dots, \bar{X}_n, B_1, \dots, B_n) \\ &\geq H_1(F_n | \bar{X}_1, \dots, \bar{X}_n, B_1, \dots, B_n) \\ &= \mathbb{E}[|P_n|] \log |L| \\ &\geq nct. \end{aligned}$$

□

With the above construction, we are now in position to prove Theorem 2.

Proof of Theorem 2. Suppose by contradiction, for some $0 < \alpha' < 1$ we have

$$h_{\alpha'}(\mu) = \lim \frac{1}{n} \frac{1}{1 - \alpha'} \log \sum_g \mu^{(n)}(g)^{\alpha'} = 0.$$

Equivalently,

$$(3.4) \quad \sum_g \mu^{(n)}(g)^{\alpha'} = \exp(o(n)).$$

Let $\alpha = \frac{1+\alpha'}{2} < 1$. For a fixed $t > 0$, let $A > 0$ be chosen later. Then by Cauchy-Schwarz inequality,

$$(3.5) \quad \left(\sum_{\pi(g) \geq An} \mu^{(n)}(g)^\alpha \exp(\alpha t \pi(s)) \right)^2 \leq \left(\sum_{\pi(g) \geq An} \mu^{(n)}(g)^{\alpha'} \right) \left(\sum_{\pi(g) \geq An} \mu^{(n)}(g) \exp(2\alpha t \pi(g)) \right).$$

For the first term in the multiplication, we know that

$$\sum_{\pi(g) \geq An} \mu^{(n)}(g)^{\alpha'} \leq \sum_s \mu^{(n)}(g)^{\alpha'} = \exp(o(n)).$$

For the second term, we note that $\pi_*\mu$ is a zero-drift random walk on \mathbb{Z} , since μ is symmetric. By the Hoeffding bound, there exists a constant $c > 0$ only depending on $\pi_*\mu$ such that

$$\pi_*\mu^{(n)}(k) = \sum_{\pi(g)=k} \mu^{(n)}(g) \leq \sum_{\pi(g) \geq k} \mu^{(n)}(g) \leq \exp\left(-c \frac{k^2}{n}\right).$$

Taking a weighted sum of the above inequality, we have

$$\begin{aligned} \sum_{\pi(g) \geq An} \mu^{(n)}(g) \exp(2\alpha t \pi(g)) &= \sum_{k=\lceil An \rceil}^{\infty} \sum_{\pi(g)=k} \mu^{(n)}(g) \exp(2\alpha t k) \\ &\leq \sum_{k=\lceil An \rceil}^{\infty} \exp\left(-c \frac{k^2}{n}\right) \exp(2\alpha t k) \\ &\leq \sum_{k=\lceil An \rceil}^{\infty} \exp(-ckA) \exp(2\alpha t k) \\ &\leq \frac{M}{cA - 2\alpha t} \exp(An(-cA + 2\alpha t)) \end{aligned}$$

for some M large enough, and given that $cA - 2\alpha t > 0$. Sending $n \rightarrow \infty$, the last line goes to 0 so

$$(3.6) \quad \lim_{n \rightarrow \infty} \sum_{\pi(g) \geq An} \mu^{(n)}(g) \exp(2\alpha t \pi(g)) \rightarrow 0.$$

(3.5) together with (3.4) and (3.6) yields

$$(3.7) \quad \sum_{\pi(g) \geq An} \mu^{(n)}(g)^\alpha \exp(\alpha t \pi(g)) = \exp(o(n)).$$

For t small enough, by Proposition 3.3, there exists a constant $c' > 0$ such that

$$c't \leq h(\mu_t) = \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \sum_g C_t^{-n\alpha} \mu^{(n)}(g)^\alpha \exp(\alpha t \pi(g)).$$

As $C_t = 1 + O(t^2)$, we have

$$ct + O(t^2) \leq \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \sum_g \mu^{(n)}(g)^\alpha \exp(\alpha t \pi(g)).$$

Because the tail is exponentially small as in (3.7), we get

$$\begin{aligned} ct + O(t^2) &\leq \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \sum_{\pi(g) \leq An} \mu^{(n)}(g)^\alpha \exp(\alpha t \pi(g)) \\ &\leq \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \sum_{\pi(g) \leq An} \mu^{(n)}(g)^\alpha \exp(\alpha t An) \\ &\leq \frac{\alpha t A}{1-\alpha} + \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \sum_{\pi(g) \leq An} \mu^{(n)}(g)^\alpha \\ &\leq \frac{\alpha t A}{1-\alpha} + \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \sum_g \mu^{(n)}(g)^\alpha \end{aligned}$$

We set $A = (4\alpha t)/c$ so that $cA - 2\alpha t > 0$. Then for t small,

$$\frac{\alpha t A}{1-\alpha} = \frac{\alpha}{1-\alpha} \frac{4\alpha}{c} t^2 < ct + O(t^2).$$

Thus,

$$h_\alpha(\mu) = \lim_n \frac{1}{n} \frac{1}{1-\alpha} \log \sum_g \mu^{(n)}(g)^\alpha > 0.$$

However, it follows that $h_{\alpha'}(\mu) \geq h_\alpha(\mu) > 0$ as $\alpha > \alpha'$, a contradiction to our assumption on α' . This concludes our proof. \square

Remark 3.4. *The proof of Theorem 2 should work more generally for other groups which have \mathbb{Z} as a quotient group and \mathbb{Z} acts interestingly on the group, such as one-dimensional Baumslag-Solitar groups.*

We next consider a particular case of the lamplighter group where $L = \mathbb{Z}_2$. We calculate the Rényi entropy of the “switch-walk-switch” (SWS) random walk on the lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}$, i.e., the μ -random walk where $\mu = \eta * \sigma * \eta$, $\eta(f^1, 0) = \eta(0, 0) = 1/2$, and $\sigma(0, -1) = \sigma(0, 1) = 1/2$. Thus, at every step the lamplighter switches the lamp at the current location with probability one half, takes a step of the simple random walk and then again flips the lamp at the current location with probability one half.

Theorem 3.5. *Let μ be the SWS walk on $\mathbb{Z}_2 \wr \mathbb{Z}$. Then*

$$(3.8) \quad h_\alpha(\mu) = \varphi_\alpha \left(\frac{4^{1/\alpha} - 4}{4^{1/\alpha} + 4} \right)$$

for all $\alpha \in (0, 1]$, where

$$(3.9) \quad \varphi_\alpha(p) = p \log 2 - \frac{\alpha}{2(1-\alpha)} [(1-p) \log(1-p) + (1+p) \log(1+p)].$$

Note that Theorem 4 is an immediate consequence, since Theorem 3.5 shows that $h_\alpha(\mu)$ is an elementary function on $(0, 1]$.

Proof of Theorem 3.5. Fix n . We first express the asymptotic Rényi entropies in terms of the number of positions the lamplighter ever visits up to time n . Let $S_m = \pi(Z_m)$ be the position of the lamplighter at time m . Note that S_m is a simple random walk on \mathbb{Z} . Let $L := \inf \{S_m, 0 \leq m \leq n\}$ and $R := \sup \{S_m, 0 \leq m \leq n\}$. The number of positions the lamplighter ever visits up to time n is then $R - L + 1$. As $L, R \in [-n, n]$, for each k there exists ℓ_k, r_k with $r_k - \ell_k + 1 = k$ such that

$$(3.10) \quad \mathbb{P}[L = \ell_k, R = r_k] \geq \frac{1}{5n^2} \mathbb{P}[R - L + 1 = k].$$

We slightly abuse notation and let $\mu^{(n)}(f) := \sum_x \mu^{(n)}(f, x)$ denote the probability that the lamp configuration at time n is equal to f . Fix $\alpha \in (0, 1)$. Then

$$\log \sum_g \mu^{(n)}(g)^\alpha = \log \sum_f \sum_x \mu^{(n)}(f, x)^\alpha \geq \log \sum_f \mu^{(n)}(f)^\alpha,$$

since $q^\alpha + p^\alpha \geq (p+q)^\alpha$ for all $p, q \in [0, 1]$. For each k , we can further bound this from below by restricting the sum:

$$\log \sum_g \mu^{(n)}(g)^\alpha \geq \log \sum_{\text{supp}(f) \subseteq [\ell_k, r_k]} \mu^{(n)}(f)^\alpha.$$

We note that conditional on L and R , the lamp configuration follows a uniform distribution over $\{0, 1\}^{[L, R]} \times \{0\}^{(-\infty, \infty) \setminus [L, R]}$. Hence, for $\text{supp}(f) \subseteq [\ell_k, r_k]$,

$$\mu^{(n)}(f) = \mathbb{P}[\tau(Z_n) = f] \geq \mathbb{P}[\tau(Z_n) = f, R = r_k, L = \ell_k] = 2^{-k} \mathbb{P}[R = r_k, L = \ell_k].$$

Thus, since the sum over all such f has 2^k summands,

$$\begin{aligned} \log \sum_g \mu^{(n)}(g)^\alpha &\geq \max_k \log [(\mathbb{P}[L = \ell_k, R = r_k] \cdot 2^{-k})^\alpha 2^k] \\ (3.11) \qquad \qquad \qquad &= \max_k \log [(\mathbb{P}[R - L + 1 = k] \cdot 2^{-k})^\alpha 2^k] + o(n). \end{aligned}$$

For the other direction,

$$\sum_g \mu^{(n)}(g)^\alpha = \sum_f \sum_{x \in [-n, n]} \mu^{(n)}(f, x)^\alpha \leq \sum_f \sum_{x \in [-n, n]} \mu^{(n)}(f)^\alpha = (2n+1) \sum_f \mu^{(n)}(f)^\alpha$$

since $\mu^{(n)}(f, x) \leq \mu^{(n)}(f)$. For a given configuration $f : \mathbb{Z} \rightarrow \mathbb{Z}/2\mathbb{Z}$, let $r(f)$ and $\ell(f)$ denote the rightmost and leftmost position on which lamps are on. Then

$$\begin{aligned} \mu^{(n)}(f) &= \mathbb{P}[\tau(Z_n) = f] = \sum_{m, s} \mathbb{P}[R = m, L = s, \tau(Z_n) = f] \\ &= \sum_{m \geq r(f), s \leq \ell(f)} \mathbb{P}[R = m, L = s] 2^{-(m-s+1)}. \end{aligned}$$

Hence

$$\begin{aligned} \sum_g \mu^{(n)}(g)^\alpha &\leq (2n+1) \sum_f \left(\sum_{m \geq r(f), s \leq \ell(f)} \mathbb{P}[R = m, L = s] \cdot 2^{-(m-s+1)} \right)^\alpha \\ &\leq (2n+1) \sum_f \sum_{m \geq r(f), s \leq \ell(f)} \mathbb{P}[R = m, L = s]^\alpha \cdot 2^{-(m-s+1)\alpha}, \end{aligned}$$

where the second inequality again uses $(p+q)^\alpha \leq p^\alpha + q^\alpha$.

We exchange the order of summation and note that for each m, s there are 2^{m-s+1} terms, to get

$$\sum_f \sum_x \mu^{(n)}(f, x)^\alpha \leq (2n+1) \sum_{m, s} \mathbb{P}[R = m, L = s]^\alpha 2^{(m-s+1)(1-\alpha)}.$$

Since this sum has $(n+1)^2$ summands,

$$\begin{aligned} \sum_f \sum_x \mu^{(n)}(f, x)^\alpha &\leq 4n^3 \max_{m, s} \mathbb{P}[R = m, L = s]^\alpha 2^{(m-s+1)(1-\alpha)} \\ &\leq 4n^3 \max_k \mathbb{P}[R - L + 1 = k]^\alpha \cdot 2^{k(1-\alpha)}. \end{aligned}$$

Together with (3.11), we have

$$(3.12) \quad \log \sum_g \mu^{(n)}(g)^\alpha = \max_k \log [\mathbb{P}[R - L = k]^\alpha \cdot 2^{k(1-\alpha)}] + o(n).$$

We claim that we can replace $\mathbb{P}[R - L = k]$ in the equation above with $\mathbb{P}[R - L \geq k]$. On the one hand, we trivially have $\mathbb{P}[R - L = k] \leq \mathbb{P}[R - L \geq k]$, yielding one direction. For the other direction, we note that for any $k' \geq 0$ there exists $k^* \geq k'$ such that

$$\mathbb{P}[R - L = k^*] \geq \frac{1}{n} \mathbb{P}[R - L \geq k']$$

by the pigeonhole principle. Therefore, for any k' there is a $k^* \geq k'$ such that

$$\begin{aligned} \max_k \log [\mathbb{P}[R - L = k]^\alpha \cdot 2^{k(1-\alpha)}] &\geq \log [\mathbb{P}[R - L = k^*]^\alpha \cdot 2^{k^*(1-\alpha)}] \\ &\geq \log \left[\frac{1}{n^\alpha} \mathbb{P}[R - L \geq k']^\alpha \cdot 2^{k'(1-\alpha)} \right] \\ &= \log [\mathbb{P}[R - L \geq k']^\alpha \cdot 2^{k'(1-\alpha)}] + \log \frac{1}{n^\alpha}. \end{aligned}$$

Taking the supremum over k' yields

$$\max_k \log [\mathbb{P}[R - L = k]^\alpha \cdot 2^{k(1-\alpha)}] \geq \max_k \log [\mathbb{P}[R - L \geq k]^\alpha \cdot 2^{k(1-\alpha)}] + o(n).$$

Now (3.12) becomes

$$\log \sum_g \mu^{(n)}(g)^\alpha = \max_k \log [\mathbb{P}[R - L \geq k]^\alpha \cdot 2^{k(1-\alpha)}] + o(n).$$

Citing Theorem 1 and Remark 1 in [7], we have that for any $x \in [0, 1]$

$$\lim_n \frac{1}{n} \log \mathbb{P}[R - L \geq nx] = \psi(x),$$

where $\psi(x) = -\frac{1}{2}(1+x)\log(1+x) - \frac{1}{2}(1-x)\log(1-x)$. We note that when a sequence of decreasing functions converges pointwise to a continuous function, the sequence also converges uniformly. Therefore, we can write

$$\frac{1}{n} \log \mathbb{P}[R - L \geq k] = \psi(k/n) + o(1),$$

where $o(1)$ vanishes uniformly in k . We have

$$\begin{aligned} \frac{1}{n} \log \sum_g \mu^{(n)}(g)^\alpha &= \frac{1}{n} \max_k \log [\mathbb{P}[R - L \geq k]^\alpha \cdot 2^{k(1-\alpha)}] + o(1) \\ &= \frac{1}{n} \max_k [\alpha \log \mathbb{P}[R - L \geq k] + k(1 - \alpha) \log 2] + o(1) \\ &= \max_k \left[\alpha \psi\left(\frac{k}{n}\right) + \frac{k}{n}(1 - \alpha) \log 2 \right] + o(1). \end{aligned}$$

Finally, let $p := \frac{k}{n}$. Then

$$\begin{aligned} h_\alpha(\mu) &= \frac{1}{1 - \alpha} \lim \frac{1}{n} \log \sum_g \mu^{(n)}(g)^\alpha \\ &= \frac{1}{1 - \alpha} \lim_n \max_{p \in \{0/n, 1/n, \dots, n/n\}} [\alpha \psi(p) + p(1 - \alpha) \log 2] \\ &= \max_{p \in [0,1]} \left[\frac{\alpha}{1 - \alpha} \psi(p) + p \log 2 \right] \\ &= \max_{p \in [0,1]} \varphi_\alpha(p) \end{aligned}$$

where

$$\begin{aligned} \varphi_\alpha(p) &= \frac{\alpha}{1 - \alpha} \psi(p) + p \log 2 \\ &= p \log 2 - \frac{\alpha}{2(1 - \alpha)} [(1 - p) \log(1 - p) + (1 + p) \log(1 + p)]. \end{aligned}$$

Some computations show that the above expression is maximized when

$$p_\alpha^* = \frac{4^{1/\alpha} - 4}{4^{1/\alpha} + 4}.$$

The continuity of the Rényi entropy from the left extends the result to h_1 . This completes the proof of Theorem 3.5. \square

3.3. The asymptotic min-entropy. The next claim provides an example of a non-symmetric random walk on an amenable group for which the asymptotic min-entropy h_∞ is positive. Consider a drifting SWS walk μ_β on $\mathbb{Z}_2 \wr \mathbb{Z}$. That is, let μ_β be given by $\mu_\beta = \eta * \sigma * \eta$, where $\eta(f^1, 0) = \eta(0, 0) = 1/2$, and $\sigma(0, -1) = \frac{1}{2}(1 - \beta) = 1 - \sigma(0, 1)$ for some $\beta \in (-1, 1)$. The drift of the walker is $\sum_g \mu_\beta(g) \pi(g) = \beta$.

Claim 3.6. *If $\beta \neq 0$ then $h_\infty(\mu_\beta) > 0$.*

Proof. By symmetry, we can assume without loss of generality that $\beta > 0$. Let Z_1, Z_2, \dots be the μ_β -random walk. Then, by the Chernoff bound, there is some $r > 0$ such that $\mathbb{P}[\pi(Z_n) < n\beta/2] \leq e^{-rn}$. In particular, if $\pi(g) \leq n\beta/2$ then $\mu_\beta(g) \leq e^{-rn}$.

Now, conditioned on $\pi(Z_n) = k \geq 0$, it holds by the definition of the SWS walk that the lamp configuration restricted to $\{0, 1, \dots, k-1\}$ is distributed uniformly, so that each configuration has probability 2^{-k} . Thus, for $k \geq n\beta/2$ and g such that $\pi(g) = k$,

$$\mu_\beta(g) \leq 2^{-n\beta/2} = e^{-\frac{1}{2} \log(2)\beta n}.$$

It follows that if we let $c = \min\{\frac{1}{2} \log(2)\beta, r\}$, then $\mu_\beta(g) \leq e^{-cn}$ for all $g \in G$, and $h_\infty(\mu) \geq c$. \square

Next, we show that for non-degenerate random walks on non-amenable groups the asymptotic min-entropy is always positive. Note that this follows from Kesten's Theorem for symmetric μ .

Claim 3.7. *Let μ be a finitely supported, non-degenerate probability measure on a non-amenable group G . Then $h_\infty(\mu) > 0$.*

To prove this claim we recall some basic definitions. Let $\ell^2(G)$ be the Hilbert space of real square-integrable functions on G , equipped with the standard inner product and norm. For $h \in G$, let $R_h: \ell^2(G) \rightarrow \ell^2(G)$ be the right shift operator given by $[R_h(\varphi)](g) = \varphi(gh)$. Then R_g is an orthogonal linear operator and $h \mapsto R_h$ is the right regular representation of G . Let $M = \sum_h \mu(h)R_h$ be the Markov operator of the μ -random walk. As is well known (see, e.g., [15, Theorem 12.5]) when G is non-amenable and μ is non-degenerate, then the operator norm $\rho(M)$ is strictly less than 1, i.e., $\rho(M) := \sup\{\|M\varphi\| : \|\varphi\| = 1\} < 1$. Note that this holds even when μ is not symmetric.

Proof of Claim 3.7. Note that $[M^n\varphi](g) = \sum_h \mu^{(n)}(h)\varphi(gh)$, and so $[M^n\delta_e](g) = \mu^{(n)}(g^{-1})$. Since $\rho(M^n) \leq \rho(M)^n$, $\|M^n\delta_e\| \leq \rho(M)^n$, and so

$$\rho(M)^{2n} \geq \sum_g \mu^{(n)}(g^{-1})^2 \geq \max_g \mu^{(n)}(g)^2.$$

Hence

$$-\log \rho(M) \leq -\frac{1}{n} \log \max_g \mu^{(n)}(g).$$

Taking the limit as n tends to infinity yields that

$$-\log \rho(M) \leq h_\infty(\mu),$$

and so, since $\rho(M) < 1$, we have proved the claim. \square

REFERENCES

- [1] A. Avez, *Théorème de Choquet-Deny pour les groupes à croissance non exponentielle*, CR Acad. Sci. Paris Sér. A **279** (1974), 25–28.
- [2] Adrien Boulanger, Pierre Mathieu, Cagri Sert, and Alessandro Sisto, *Large deviations for random walks on hyperbolic spaces* (2021).
- [3] Filipp Buryak and Yuliya Mishura, *Convexity and robustness of the rényi entropy*, 2021, pp. 387–412.
- [4] Imre Csiszár, *Axiomatic characterizations of information measures*, Entropy **10** (2008), no. 3, 261–273.
- [5] Anna Erschler and Tianyi Zheng, *Growth of periodic grigorchuk groups*, Inventiones mathematicae **219** (2020), no. 3, 1069–1155.
- [6] Joshua Frisch, Yair Hartman, Omer Tamuz, and Pooya Vahidi Ferdowsi, *Choquet-deny groups and the infinite conjugacy class property*, Annals of Mathematics **190** (2019), no. 1, 307–320.
- [7] Yuji Hamana and Harry Kesten, *Large deviations for the range of an integer valued random*, Annales de l’Institut Henri Poincaré **38** (2002), 17–58.
- [8] Vadim A Kaimanovich and Anatoly M Vershik, *Random walks on discrete groups: boundary and entropy*, The annals of probability (1983), 457–490.
- [9] Harry Kesten, *Full Banach mean values on countable groups*, Mathematica Scandinavica (1959), 146–156.
- [10] ———, *Symmetric random walks on groups*, Transactions of the American Mathematical Society **92** (1959), no. 2, 336–354.
- [11] Steven P Lalley, *Random walks on infinite groups*, Vol. 297, Springer Nature, 2023.
- [12] Xiaosheng Mu, Luciano Pomatto, Philipp Strack, and Omer Tamuz, *From Blackwell dominance in large samples to Rényi divergences and back again*, Econometrica **89** (2021), no. 1, 475–506. <https://arxiv.org/pdf/1906.02838>.
- [13] Jose C Principe, *Information theoretic learning: Rényi’s entropy and kernel perspectives*, Springer Science & Business Media, 2010.
- [14] Alfréd Rényi, *On measures of entropy and information*, Proceedings of the fourth berkeley symposium on mathematical statistics and probability, volume 1: Contributions to the theory of statistics, 1961.
- [15] Wolfgang Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, Cambridge University Press, 2000.