Canad. Math. Bull. Vol. **64** (4), 2023, pp. 1274–1279
http://dx.doi.org/10.4153/S0008439523000358

Strong Shannon–McMillan–Breiman's theorem for locally compact groups

Behrang Forghani and May Nguyen

Abstract. We prove that for a vast class of random walks on a compactly generated group, the exponential growth of convolutions of a probability density function along almost every sample path is bounded by the growth of the group. As an application, we show that the almost sure and L^1 convergences of the Shannon–McMillan–Breiman theorem hold for compactly supported random walks on compactly generated groups with subexponential growth.

1 Introduction

One of the fundamental results to estimate asymptotic entropy in different contexts is the *Shannon–McMillan–Breiman theorem* [Bre, McM53, Sha48]. The analog of the Shannon–McMillan–Breiman theorem for stationary processes on uncountable spaces was developed over 20 years. It started in 60s by Moy [Moy], Perez [Per64], and Kieffer [Kie], and completed in mid-80s by Barron [Bar] and Algoet-Cover [AC]. In the context of random walks on countable groups, Derriennic [Der80] and Kaimanovich and Vershik [KV] proved the Shannon–McMillan–Breiman theorem. Their proofs heavily use Kingman's subadditive ergodic theorem, which fails for random walks on noncountable locally compact groups. Derriennic [Der80] asked if one can establish similar results in the context of random walks on noncountable locally compact groups. Although the analog of the Shannon–McMillan–Breiman theorem for stationary processes on continuous spaces was completed in 80s, its analog for random walks on noncountable locally compact groups remained unsolved in the last 40 years, until, in a recent result [FT22], Forghani and Tiozzo provided a weak version of the Shannon–McMillan–Breiman theorem for random walks on locally compact groups.

One of the questions remaining unsolved to complete the entropy theory of random walks on locally compact groups is proving (or disproving) strong versions of the Shannon–McMillan–Breiman theorem such as almost sure and L^1 convergences. The goal of this short note is to provide an upper bound for the exponential growth of convolutions of a probability density function along almost every sample path for a vast class of random walks on compactly generated groups, Theorem 1.2.

Received by the editors October 5, 2022; revised May 1, 2023; accepted May 1, 2023.
Published online on Cambridge Core May 5, 2023.
AMS subject classification: **60B15**, **60G50**, 37A35.
Keywords: Shannon, locally compact groups, entropy, random walks.

As a consequence of this result, we prove strong versions of the Shannon–McMillan–Breiman theorem for groups with subexponential growth, Theorem 1.3.

Let G be a locally compact group (that includes being second countable and Hausdorff) with the unique (up to a positive multiplicative constant) left Haar measure m. Let μ be a Borel probability measure on G. We say that μ is *absolutely continuous* if μ is absolutely continuous with respect to the left Haar measure m. We denote by $\frac{d\mu}{dm}$ the Radon–Nikodym derivative of μ with respect to m, and by μ^{*n} the n-fold convolution of μ. Note that when μ is absolutely continuous, μ^{*n} is also absolutely continuous. The n-fold convolution of μ is related to the random walk generated by μ on G. Let $(g_i)_{i\geq 1}$ be a sequence of independent identically μ-distributed random variables. The position of the random walk (G,μ) at time n is

$$x_n = g_1 g_2 \cdots g_n.$$

By the definition of the n-fold convolution, the distribution of the random variable x_n is μ^{*n}. A sequence of $\boldsymbol{x} = (x_n)_{n\geq 1}$ is called a sample path of the random walk (G,μ). Let $(\Omega,\mathbb{P})$ be the space of sample paths of the random walk (G,μ). The differential entropy of μ^{*n} with respect to m is

$$H_n = -\int_G \log \frac{d\mu^{*n}}{dm}(g)\, d\mu^{*n}(g).$$

Note that if G is not a countable group, then H_n could be negative. For example, let $G = \mathbb{R}$ and μ be uniformly distributed on the interval $[-1/4, 1/4]$, then $H_1 < 0$.

Conjecture 1.1 (Strong the Shannon–McMillan–Breiman) *Let G be a locally compact group. Let μ be an absolutely continuous probability measure on G with bounded density. If $H_n < \infty$ for all n, then for $\mathbb{P}$-almost every sample path (x_n), and*

$$-\frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n) \to h(\mu)$$

in $L^1(\Omega,\mathbb{P})$, where $h(\mu) = \lim_{n\to\infty} \frac{H_n}{n}$.

Note that under the assumptions in Conjecture 1.1, $h(\mu)$ exists and is finite. The invariant quantity $h(\mu)$ is called the *asymptotic entropy* of μ and plays a crucial role in understanding bounded harmonic functions and the Poisson boundary of a random walk. For instance, when the asymptotic entropy is finite, $h(\mu) = 0$ if and only if all bounded harmonic functions are constant (equivalently, the Poisson boundary is trivial; see [Der80, KV]. A version of the Shannon–McMillan–Breiman theorem is used to prove *ray and strip approximations*, fundamental criteria to identify the Poisson boundary and bounded harmonic functions of a random walk (see [FT22, Kai00] for more details).

Conjecture 1.1 has been solved when G is a countable group by Kaimanovich and Vershik [KV] and Derriennic [Der80] by using the subadditive ergodic theorem. In the recent development, Forghani and Tiozzo [FT22] established a weak version of Conjecture 1.1 for random walks on locally compact groups, that is, for $\mathbb{P}$-almost every

https://doi.org/10.4153/S0008439523000358 Published online by Cambridge University Press

sample path (x_n),

$$\liminf_{n\to\infty} -\frac{1}{n}\log\frac{d\mu^{*n}}{dm}(x_n) = h(\mu). \tag{1}$$

This paper is devoted to investigating Conjecture 1.1. Let K be a symmetric compact subset of G that generates G, that is, $G =< K >$. Hence, $\cup_{n=0}^{\infty} K^n = G$. The growth of K is

$$\nu(K) = \limsup_{n\to\infty} \frac{\log m(K^n)}{n}.$$

Note that $\nu(K)$ is finite (see, for instance, [FT22]). We say G has a *subexponential growth* if $\nu(K) = 0$ for one (equivalently, for every) compact symmetric generator K. For a probability measure μ with the compact support K, we define ν_μ to be the growth of its support, that is, $\nu_\mu = \nu(K)$.

Theorem 1.2 *Let G be a compactly generated locally compact group. Let μ be an absolutely continuous probability measure on G with an almost everywhere bounded density function. If the support of μ is compact, then for almost every sample path $\boldsymbol{x} = (x_n)$, we have*

$$h(\mu) \le \limsup_{n\to\infty} -\frac{1}{n}\log\frac{d\mu^{*n}}{dm}(x_n) \le \nu_\mu. \tag{2}$$

Note that the first inequality in (2) follows from (1). The inequality $h(\mu) \le \nu_\mu$ is a consequence of properties of differential entropy (see [Der80]). Our contribution is to show that the second inequality in (2) also holds. We prove this theorem in the next section. Even the fact that the limsup should be bounded is not straightforward. The proof is inspired by techniques in [FT22]. As an application of Theorem 1.2, we affirmatively answer Conjecture 1.1 when G has a subexponential growth.

Theorem 1.3 *Let G be a compactly generated locally compact group of subexponential growth. Let μ be an absolutely continuous probability measure on G with an almost everywhere bounded density function. If the support of μ is compact, then for almost every sample path $\boldsymbol{x} = (x_n)$, and*

$$\lim_{n\to\infty} \frac{1}{n}\log\frac{d\mu^{*n}}{dm}(x_n) = 0$$

in $L^1(\Omega, \mathbb{P})$.

Remark 1.4 A compactly generated group has polynomial growth if there exists $d > 0$ such that $m(K^n) = O(n^d)$ for some compact symmetric generator K. The class of subexponential groups includes groups with polynomial growth. By Gromov's result, a finitely generated group has polynomial growth if and only if it is virtually nilpotent. More generally, for compactly generated locally compact groups, Losert [Los01] proved that polynomial growth is equivalent to an existence of a normal series of normal closed subgroups $\{e\} \subset \cdots G_1 \subset G_n = G$ such that G_i/G_{i+1} is an $\overline{\text{FC}}$-group for $i = 0, \ldots, n$.

https://doi.org/10.4153/S0008439523000358 Published online by Cambridge University Press

Remark 1.5 Note that Erschler [Ers04] provided examples of countable groups with subexponential growth which admit symmetric probability measures with finite entropy such that the asymptotic entropy is positive. Indeed, those probability measures are infinitely supported (hence not compactly supported) and do not contradict Theorem 1.3.

2 Proof of theorems

2.1 Proof of Theorem 1.2

The proof relies on the Borel–Cantelli lemma. Let K be the support of the probability measure μ. The support of μ^{*n} is K^n. For almost every sample path $\boldsymbol{x} = (x_n)$, define

$$F_n(\boldsymbol{x}) = \begin{cases} \left(\frac{d\mu^{*n}}{dm}(x_n)\right)^{-1}, & x_n \in K^n, \\ 0, & \text{otherwise.} \end{cases}$$

Note that $\frac{d\mu^{*n}}{dm}$ is a density function, so $\frac{d\mu^{*n}}{dm}(x) > 0$ for μ^{*n}-almost every x in G. Hence, $F_n(\boldsymbol{x})$ is well defined. Let

$$\limsup_{n\to\infty} \frac{1}{n} \log m(K^n) = v_\mu = v.$$

For $\varepsilon > 0$, define

$$B_n(\varepsilon) = B_n = \left\{\boldsymbol{x} \in \Omega \;:\; 1 \le e^{-n(\varepsilon+v)} F_n(\boldsymbol{x})\right\}.$$

The definition of the measurable set B_n implies that

$$\mathbb{P}(B_n) = \int_{B_n} d\mathbb{P} \le e^{-n(\varepsilon+v)} \int_\Omega F_n(\boldsymbol{x}) d\mathbb{P}(\boldsymbol{x}). \tag{3}$$

Because $F_n(\boldsymbol{x})$ only depends on the nth position of the sample path $\boldsymbol{x}$, the Markovian property of the random walk implies that

$$\int_\Omega F_n(\boldsymbol{x}) d\mathbb{P}(\boldsymbol{x}) = \int_{\{\boldsymbol{x}:\, x_n = g\}} F_n(\boldsymbol{x}) d\mu^{*n}(g). \tag{4}$$

Using the definition of F_n in the above equation yields

$$\int_\Omega F_n(\boldsymbol{x}) d\mathbb{P}(\boldsymbol{x}) = \int_{K^n} \left(\frac{d\mu^{*n}}{dm}(g)\right)^{-1} d\mu^{*n}(g) = \int_{K^n} 1 dm(g) = m(K^n). \tag{5}$$

By combining (3) and (5), we obtain

$$\mathbb{P}(B_n) \le e^{-n(\varepsilon+v)} m(K^n).$$

Because $\varepsilon + v > 0$, we deduce that $\sum_n \mathbb{P}(B_n) < \infty$, and applying the Borel–Cantelli lemma implies that

$$\mathbb{P}(\limsup_{n\to\infty} B_n) = 0.$$

https://doi.org/10.4153/S0008439523000358 Published online by Cambridge University Press

Therefore, the complement of B_n occurs for n sufficiently large. We conclude that for $\varepsilon > 0$, for n sufficiently large, and for almost every $\boldsymbol{x}$ in Ω,

$$e^{n(\varepsilon+\nu)} \geq F_n(\boldsymbol{x}) \implies n(\varepsilon+\nu) \geq \log F_n(\boldsymbol{x});$$

therefore, $\limsup_{n\to\infty} \frac{1}{n} \log F_n(\boldsymbol{x}) \leq (\varepsilon+\nu)$ for every $\varepsilon > 0$, which implies the desired result after letting ε decrease to 0.

2.2 Proof of Theorem 1.3: almost sure convergence

Note that for every g in G and for every natural number n, we can write

$$\frac{d\mu}{dm}^{*(n+1)}(g) = \int_G \frac{d\mu}{dm}(x^{-1}g)d\mu^{*n}(x) \leq \left\| \frac{d\mu}{dm} \right\|_\infty \int_G d\mu^{*n}(x) = \left\| \frac{d\mu}{dm} \right\|_\infty .$$

For every sample path $\boldsymbol{x} = (x_n)$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n) \leq \limsup_{n\to\infty} \frac{1}{n} \log \left\| \frac{d\mu}{dm} \right\|_\infty = 0. \tag{6}$$

Since G has subexponential growth, hence $\nu = 0$. Applying Theorem 1.2, we obtain for $\mathbb{P}$-almost every sample path $\boldsymbol{x}$ in Ω

$$\limsup_{n\to\infty} -\frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n) \leq 0. \tag{7}$$

Because $\limsup_{n\to\infty} -\frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n) = -\liminf_{n\to\infty} \frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n)$, combining with (6) yields

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n) \leq 0 \leq \liminf_{n\to\infty} \frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n),$$

and hence $\frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n) \to 0$ as $n \to \infty$ for almost every sample path $\boldsymbol{x} = (x_n)$.

2.3 Proof of Theorem 1.3: L^1 convergence

The proof follows from Scheffé's lemma (see, for example, [Wil91, 5.10]). We have $\frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n) \to 0$ for almost every sample path $\boldsymbol{x} = (x_n)$. Define

$$b_n = \frac{1}{n} \log \left\| \frac{d\mu}{dm} \right\|_\infty - \frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n).$$

We have $b_n \geq 0$ and

$$\int_\Omega b_n d\mathbb{P} = \frac{1}{n} \log \left\| \frac{d\mu}{dm} \right\|_\infty + \frac{1}{n} H_n \to 0.$$

Thus, $b_n \to 0$ in $L^1(\Omega, \mathbb{P})$. Also, $\frac{1}{n} \log \left\| \frac{d\mu}{dm} \right\|_\infty \to 0$ in $L^1(\Omega, \mathbb{P})$. Therefore, $\frac{1}{n} \log \frac{d\mu^{*n}}{dm}(x_n) \to 0$ in $L^1(\Omega, \mathbb{P})$.

https://doi.org/10.4153/S0008439523000358 Published online by Cambridge University Press

References

[AC] P. H. Algoet and T. M. Cover, *A sandwich proof of the Shannon–McMillan–Breiman theorem*. Ann. Probab. **16**(1988), no. 2, 899–909.

[Bar] A. R. Barron, *The strong ergodic theorem for densities: generalized Shannon–McMillan–Breiman theorem*. Ann. Probab. **13**(1985), no. 4, 1292–1303.

[Bre] L. Breiman, *The individual ergodic theorem of information theory*. Ann. Math. Stat. **28**(1957), 809–811.

[Der80] Y. Derriennic, *Quelques applications du théorème ergodique sous-additif*. In: Conference on random walks (Kleebach, 1979), Astérisque, 74, Société mathématique de France, Paris, 1980, pp. 183–201 (in French).

[Ers04] A. Erschler, *Boundary behavior for groups of subexponential growth*. Ann. of Math. (2) **160**(2004), no. 3, 1183–1210.

[FT22] B. Forghani and G. Tiozzo, *Shannon's theorem for locally compact groups*. Ann. Probab. **50**(2022), no. 1, 61–89.

[Kai00] V. A. Kaimanovich, *The Poisson formula for groups with hyperbolic properties*, Ann. of Math. (2) **152**(2000), no. 3, 659–692.

[KV] V. A. Kaimanovich and A. M. Vershik, *Random walks on discrete groups: Boundary and entropy*. Ann. Probab. **11**(1983), no. 3, 457–490.

[Kie] J. C. Kieffer, *A simple proof of the Moy–Perez generalization of the Shannon–McMillan theorem*. Pacific J. Math. **51**(1974), 203–206.

[Los01] V. Losert, *On the structure of groups with polynomial growth II*. J. Lond. Math. Soc. (2) **63**(2001), no. 3, 640–654.

[McM53] B. McMillan, *The basic theorems of information theory*. Ann. Math. Stat. **24**(1953), 196–219.

[Moy] S.-t. C. Moy, *Generalizations of Shannon–McMillan theorem*. Pacific J. Math. **11**(1961), 705–714.

[Per64] A. Perez, *Extensions of Shannon–McMillan's limit theorem to more general stochastic processes*. In: Transactions of the third Prague conference on information theory, statistical decision functions, random processes (Liblice, 1962), Academia Publishing House of the Czech Academy of Sciences, Prague, 1964, pp. 545–574.

[Sha48] C. E. Shannon, *A mathematical theory of communication*. Bell System Tech. J. **27**(1948), 379–423, 623–656.

[Wil91] D. Williams, *Probability with martingales*, Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, 1991.

Department of Mathematics, The College of Charleston, 66 George Street, Charleston, SC 29403, USA
e-mail: forghanib@cofc.edu

Department of Statistics, George Mason University, 4400 University Drive, Fairfax, VA 22030, USA
e-mail: mnguy22@gmu.edu

https://doi.org/10.4153/S0008439523000358 Published online by Cambridge University Press