



# Noise sensitivity of random walks on groups

Itai Benjamini and Jérémie Brioussel

Department of Mathematics, Weizmann Institute of Science, Rehovot, Israel

*E-mail address:* [itai.benjamini@gmail.com](mailto:itai.benjamini@gmail.com)

*URL:* <https://www.wisdom.weizmann.ac.il/~itai/>

Institut Montpellierain Alexander Grothendieck, Université de Montpellier, France

*E-mail address:* [jeremie.brieussel@umontpellier.fr](mailto:jeremie.brieussel@umontpellier.fr)

*URL:* <https://imag.umontpellier.fr/~brieussel/>

**Abstract.** A random walk on a group is noise sensitive if resampling every step independently with a small probability results in an almost independent output. We precisely define two notions:  $\ell^1$ -noise sensitivity and entropy noise sensitivity. Groups with one of these properties are necessarily Liouville. Homomorphisms to free abelian groups provide an obstruction to  $\ell^1$ -noise sensitivity. We also provide examples of  $\ell^1$  and entropy noise sensitive random walks.

Noise sensitivity raises many open questions which are described at the end of the paper.

## 1. Introduction

Physically, noise is a non-significant perturbation of an observation. In signal theory, noise is an unintentional perturbation of a message. Noise is an inherent phenomenon to physical observations and to communication. Its influence on a channel capacity was already taken into account by Shannon in his mathematical theory of communication [Shannon \(1948\)](#).

In probability theory, the noise of an event  $E(x_1, \dots, x_n)$  (i.e. a Boolean function) depending on a large number of variables can be modeled as the effect of replacing a (small) proportion  $\rho \in (0, 1)$  of the variables by random entries. An event is *noise sensitive* if the realisation of  $E$  gives no (or very little) information on what happens when the entries are perturbed by noise. For instance in bond percolation at critical probability  $\frac{1}{2}$ , having a left-right crossing on an  $n \times n$  square lattice is a noise sensitive event [Benjamini et al. \(1999\)](#); [Garban et al. \(2010\)](#). On the contrary, (weighted) majority functions are noise stable [Benjamini et al. \(1999\)](#). The effect of noise in percolation and voting games has been widely studied over the last decades. We refer to [Benjamini et al. \(1999\)](#); [Garban and Steif \(2015\)](#); [Garban et al. \(2010\)](#); [Kalai \(2018\)](#) and references therein for more on these topics.

In the present paper we investigate sensitivity to noise of random walks on groups. A random walk on a group  $G$  is a sequence of products  $X_n = s_1 \dots s_n$  of independent variables  $s_i$  following

---

*Received by the editors May 9th, 2022; accepted June 2nd, 2023.*

*2010 Mathematics Subject Classification.* 60B15, 05C81, 20F65.

*Key words and phrases.* Random walks on groups, noise sensitivity.

Second Author has been partially supported by grant ANR 22-CE40-0004 GoFR. .

identical distribution  $\mu$ . Given such a product  $X_n$ , we can noise it by resampling independently each increment  $s_i$  with a probability  $\rho \in (0, 1)$ , this provides a new variable  $Y_n^\rho$  depending on  $X_n$ . In the Cayley graph, the original word  $X_n$  can be interpreted as a sequence of instructions for a moving particle and  $Y_n^\rho$  as the effective trajectory if instructions are misread with probability  $\rho$ .

Broadly speaking, the random walk  $(G, \mu)$  is *noise sensitive* if  $Y_n^\rho$  generally seems independent of  $X_n$ . This vague statement can be specified in a number of ways and we refer to Section 2 for several precise definitions. We retain two principal notions of noise sensitivity in this context. The random walk  $(G, \mu)$  is  $\ell^1$ -noise sensitive if the law of the pair  $(X_n, Y_n^\rho)$  and the law of a pair  $(X_n, X'_n)$  of two *independent* samples are close in the sense that their  $\ell^1$ -distance tends to zero. The random walk is entropy noise sensitive if the ratio between the conditional entropy  $H(Y_n^\rho|X_n)$  and  $H(X_n)$  tends to one, which means that the average amount of information needed to describe  $Y_n^\rho$  once we already know  $X_n$  is (asymptotically) as big as the average amount of information needed to describe  $X_n$ .

These notions of noise sensitivity are relevant for infinite groups. They are trivially satisfied for finite groups by eventual equidistribution – see Theorem 5.1. For simplicity, we restricted our investigations to finitely generated groups and finitely supported probability measures, though the notions make sense in a wider setting.

We first point out two elementary obstructions to  $\ell^1$ -noise sensitivity.

**Theorem 1.1.** *Let  $G$  be a countable group and  $\mu$  a finitely supported probability measure on  $G$ .*

- (1) *If  $G$  admits a non-trivial homomorphism onto a free abelian group, then  $(G, \mu)$  is not  $\ell^1$ -noise sensitive.*
- (2) *If  $(G, \mu)$  is  $\ell^1$ -noise sensitive, then it is Liouville.*

This is a concatenation of Corollary 5.4 and Theorem 4.1. The first part essentially follows from the central limit theorem. For the second part, recall that the space of bounded harmonic functions on  $(G, \mu)$  is parametrized by the Poisson boundary Furstenberg (1971). The random walk  $(G, \mu)$  is Liouville if this boundary is reduced to a point, i.e. if all bounded harmonic functions are constant. In the non-Liouville case, the first increment  $s_1$  already carries information on the position of  $X_n$ . For instance in a free group, the first letter of a minimal representative word of  $X_n$  is correlated to the first increment. This prevents  $\ell^1$ -noise sensitivity of the free group. In fact, this obstruction can be strengthened to show that the free group is not even noise sensitive at large scale, see Proposition 4.2.

By Kaĭmanovich and Vershik (1983), Liouville property is also equivalent to the vanishing of the asymptotic entropy  $\lim \frac{1}{n} H(X_n)$  which measures the average amount of information in one increment of  $X_n$ . This provides an elementary obstruction to entropy noise sensitivity.

**Theorem 1.2.** *If  $(G, \mu)$  is entropy noise sensitive, then it is Liouville.*

This is a particular case of Theorem 4.3. It follows directly from the upper semi-continuity of the asymptotic entropy Amir et al. (2013, Proposition 4). By Frisch et al. (2019), any non-virtually nilpotent finitely generated group admits a symmetric probability measure of finite entropy (usually not finitely supported) for which it is non-Liouville. It is a fortiori neither  $\ell^1$ -noise sensitive, nor entropy noise sensitive.

Despite these obstructions, we provide examples of noise sensitive random walks on groups.

**Theorem 1.3.** *Regarding entropy noise sensitivity:*

- *finitely generated abelian groups are entropy noise sensitive with respect to any finitely supported measure,*
- *the lamplighter group  $\mathbb{Z}/2\mathbb{Z} \wr \mathbb{Z}$  is entropy noise sensitive with respect to the switch-walk-switch measure,*

- the permutational extension of the first Grigorchuk group described in [Bartholdi and Erschler \(2012\)](#) is partially entropy noise sensitive for some switch and walk measure, i.e.

$$\exists c > 0, \forall \rho \in (0, 1), \quad \liminf \frac{H(Y_n^\rho | X_n)}{H(X_n)} \geq c.$$

The first point is in Proposition 5.2, the second is Theorem 7.1 and the third point a particular case of Theorem 7.3.

**Theorem 1.4.** *On the infinite dihedral group  $D_\infty = \langle a, b | a^2, b^2 \rangle$ ,*

- the simple random walk is not  $\ell^1$ -noise sensitive,
- the lazy simple random walk is  $\ell^1$ -noise sensitive.

This is Theorem 6.1. It follows that  $\ell^1$ -noise sensitivity is not a geometric property, since it depends on the measure. Moreover as our notions of noise sensitivity are stable under taking direct products (of groups and measures), we obtain noise sensitive random walks on groups commensurable to  $\mathbb{Z}^d$  for any rank  $d$ .

We expect that morphisms to  $\mathbb{Z}$  and non-Liouville property (significance of the first increment) are the only two obstructions for a group to be noise sensitive. This raises the:

**Question 1.5.** *If  $(G, \mu)$  is Liouville and has no virtual morphism onto  $\mathbb{Z}$ , is it  $\ell^1$ -noise sensitive?*

Another (unexpected) type of obstruction to  $\ell^1$ -noise sensitivity would probably be very interesting and significant. As a particular case, we conjecture that:

**Conjecture 1.6.** *The first Grigorchuk group is  $\ell^1$ -noise sensitive for finitely supported  $\mu$ .*

As for entropy noise sensitivity, we believe it is widely spread and we conjecture:

**Conjecture 1.7.** *A random walk  $(G, \mu)$  is entropy noise sensitive if and only if it is Liouville.*

In particular, we believe that  $\ell^1$ -noise sensitivity implies entropy noise sensitivity – see Section 3.

Physically, an  $\ell^1$ -noise sensitive process can somewhat not be observed, since the observation  $Y_n^\rho$  does not provide any significant information on the actual output  $X_n$ . Speculatively, this could account for the rarity of Liouville (probability) groups in natural science. Indeed besides virtually nilpotent ones, all known Liouville groups are genuinely mathematical objects, introduced by mathematicians to solve their problems, without reference to physical sciences [Alešin \(1972\)](#); [Kaĭmanovich and Vershik \(1983\)](#); [Grigorchuk \(1985, 1986\)](#); [Amir et al. \(2013\)](#); [Brioussel \(2013\)](#); [Matte Bon \(2014\)](#); [Nekrashevych \(2018\)](#).

**Organization of the paper.** Precise definitions of noise sensitivity are given in Section 2 and some of their relationships are studied in Section 3. The effects of Liouville property on noise sensitivity are discussed in Section 4. Section 5 is devoted to finite and abelian groups, Section 6 to the infinite dihedral group. Wreath products are studied in Section 7. Some perspectives and open questions are presented in the final Section 8.

## 2. Notions of noise sensitivity

Let  $G$  be a countable group. We endow it with a probability measure  $\mu$  whose support generates  $G$  as a semi-group. The random walk  $(G, \mu)$  is the sequence of random variables  $X_n = s_1 \dots s_n$  where  $(s_k)_{k \geq 1}$  are independent of law  $\mu$ . The law of  $X_n$  is the  $n$ -fold convolution  $\mu_n := \mu^{*n}$ .

Given  $X_n$  and a noise parameter  $\rho \in [0, 1]$ , we consider another random variable where the increments  $s_k$  are *refreshed* (i.e. resampled independently according to  $\mu$ ) with probability  $\rho$ . More precisely, we define  $Y_n^\rho = r_1 \dots r_n$  where

$$r_k = \begin{cases} s_k & \text{with probability } 1 - \rho, \\ s'_k & \text{with probability } \rho, \end{cases}$$

where  $s'_k$  is an independent random variable of law  $\mu$ . We interpret  $Y_n^\rho$  as a version of  $X_n$  perturbed by some noise. Of course the law of  $Y_n^\rho$  alone is the same as that of  $X_n$ .

We denote by  $\pi_n^\rho$  the joint law of  $(X_n, Y_n^\rho)$  on  $G \times G$ . It is the distribution at time  $n$  of the random walk on  $G \times G$  with measure  $\pi^\rho := (1 - \rho)\mu^{\text{diag}} + \rho\mu^2$  where  $\mu^{\text{diag}}$  is the diagonal measure on the product taking values  $\mu^{\text{diag}}(x, y) = \mu(x)$  when  $x = y$  and  $\mu^{\text{diag}}(x, y) = 0$  otherwise, and  $\mu^2 := \mu \times \mu$ . In particular,  $\pi_n^\rho$  is symmetric in the two variables,  $\pi_n^1 = \mu_n^2$  is the law of a  $\mu^2$ -random walk on  $G \times G$  at time  $n$  and  $\pi_n^0 = \mu_n^{\text{diag}}$  is the original random walk embedded diagonally in  $G \times G$ .

In this paper, we investigate the notion of *noise sensitivity* of a random walk, that is how much  $Y_n^\rho$  for  $\rho \in (0, 1)$  can differ from  $X_n$ . Informally, we say that  $Y_n^\rho$  is *noise sensitive* if the couple  $(X_n, Y_n^\rho)$  resembles a couple  $(X_n, X'_n)$  of two independent samples of the random walk, i.e. if the two probability measures  $\pi_n^\rho$  and  $\mu_n^2$  are close. We present several precise quantitative ways of measuring noise sensitivity.

2.1. *Measure-theoretic notions of noise sensitivity.* Let us start with the most natural notion.

**Definition 2.1.**  *$\ell^1$ -noise sensitivity.* The random walk  $(G, \mu)$  is  *$\ell^1$ -noise sensitive* if

$$\forall \rho \in (0, 1), \quad \|\pi_n^\rho - \mu_n^2\|_1 \xrightarrow{n \rightarrow \infty} 0.$$

We recall that the  $\ell^1$ -distance (or twice total variation) between two probability measures  $\xi_1, \xi_2$  on a countable space  $E$  is

$$\|\xi_1 - \xi_2\|_1 := \sum_{x \in E} |\xi_1(x) - \xi_2(x)| \in [0, 2].$$

The  $\ell^1$ -distance is also characterised in terms of coupling:

$$\|\xi_1 - \xi_2\|_1 = \inf_{\nu \in \text{Coup}(\xi_1, \xi_2)} \nu(x \neq y),$$

where a coupling  $\nu$  of  $\xi_1$  and  $\xi_2$  is a probability measure on  $E \times E$  whose marginals are  $\xi_1$  and  $\xi_2$ .

Let us now define a weaker notion of noise sensitivity, related to entropy. Recall that the Shannon entropy of a random variable  $X$  of law  $\xi$  taking values in a countable set  $E$  is

$$H(X) = H(\xi) := - \sum_{x \in E} \xi(x) \log(\xi(x)) = \mathbb{E}_\xi I_\xi(x),$$

where  $I_\xi(x) = -\log(\xi(x))$  is the information function. Informally, the entropy of  $X$  is the average number of digits needed to describe the value of  $X$ . Moreover the conditional entropy of another random variable  $Y$  with respect to  $X$  is

$$H(Y|X) = \mathbb{E}_X H_X(Y) = \sum_{x \in X(E)} \mathbb{P}(X = x) \left( - \sum_{y \in E} \mathbb{P}(Y = y|X = x) \log \mathbb{P}(Y = y|X = x) \right)$$

i.e. the expectation with respect to  $X$  of the entropy of the law of  $Y$  conditioned by the value of  $X$ . Informally this is the average amount of information needed to describe  $Y$  when we already know  $X$ .

**Definition 2.2.** *Entropy noise sensitivity.* The random walk  $(G, \mu)$  is *entropy noise sensitive* if

$$\forall \rho \in (0, 1), \quad \frac{H(Y_n^\rho|X_n)}{H(X_n)} \xrightarrow{n \rightarrow \infty} 1.$$

The random walk is *partially entropy noise sensitive* if

$$\exists c > 0, \forall \rho \in (0, 1), \exists N, \forall n \geq N, \quad H(Y_n^\rho|X_n) \geq cH(X_n).$$

Informally speaking, partial entropy noise sensitivity ensures that no matter how small the noise parameter, there is always a fixed proportion of information that cannot be recovered from the noised sample.

It is well known that for any two random variables  $H((X, Y)) = H(Y|X) + H(X)$ . In the present context, we get  $H(Y_n^\rho|X_n) = H(\pi_n^\rho) - H(\mu_n)$ .

*2.2. Metric notions of noise sensitivity.* As the notion of  $\ell^1$ -convergence of measures is strong, it can be interesting to relax it. We propose here other notions of noise sensitivity, related to metrics on the group, and motivated by the usual interest in geometric group theory for the large scale aspects of the metric. They can be omitted in first reading.

Given a metric  $d$  on a space  $E$ , we may define other notions of convergence of measures, for instance via the Wasserstein distances, defined for  $p \in [1, \infty)$  by:

$$W_p(\xi_1, \xi_2) := \inf_{\nu \in \text{Coup}(\xi_1, \xi_2)} \mathbb{E}_\nu[d(x, y)^p]^{\frac{1}{p}}.$$

In practice, we will rather use a related notion based on the following quantity: given  $s > 0$  set

$$U^s(\xi_1, \xi_2) := \inf_{\nu \in \text{Coup}(\xi_1, \xi_2)} \nu(d(x, y) \geq s).$$

By Markov inequality, we have:

$$U^s(\xi_1, \xi_2)^{\frac{1}{p}} \leq \frac{W_p(\xi_1, \xi_2)}{s} \tag{2.1}$$

**Definition 2.3. Noise sensitivity at scale  $s_n$ .** Given a left-invariant distance on  $G \times G$  and a sequence  $(s_n)$ , the random walk  $(G, \mu)$  is *noise sensitive at scale  $s_n$*  if

$$\forall \rho \in (0, 1), \quad U^{s_n}(\pi_n^\rho, \mu_n^2) \xrightarrow{n \rightarrow \infty} 0.$$

From a mass transport point of view, this means that for  $n$  large, all but an  $\varepsilon$  proportion of the ‘‘sand pile’’ mass distribution of  $\mu_n^2$  can be obtained from  $\pi_n^\rho$ , moving sand by distance less than  $s_n$ .

Unless mentioned otherwise, we assume the distance on  $G \times G$  is a word distance, e.g.  $d_{G \times G}((x, y), (x', y')) = d_G(x, x') + d_G(y, y')$  where  $d_G$  is a word distance on  $G$ . We recall that all word distances on a finitely generated group are equivalent up to multiplicative constants.

We point out:

- Noise sensitivity at scale 1 is equivalent to  $\ell^1$ -noise sensitivity when the distance takes integer values, e.g. in discrete groups with word metric.
- If a Wasserstein distance satisfies  $W_p(\pi_n^\rho, \mu_n^2) = o(s_n)$ , then  $(G, \mu)$  is noise sensitive at scale  $s_n$  by (2.1).
- If  $s'_n \geq s_n$  and  $(G, \mu)$  is noise sensitive at scale  $s_n$ , then it is noise sensitive at scale  $s'_n$ , because  $U^{s'}(\xi_1, \xi_2) \leq U^s(\xi_1, \xi_2)$  whenever  $s' \geq s$ .
- The quantity  $U^s$  resembles a distance in the sense that

$$U^s(\xi_1, \xi_3) \leq U^{s_1}(\xi_1, \xi_2) + U^{s_2}(\xi_2, \xi_3) \quad \text{whenever } s_1 + s_2 = s. \tag{2.2}$$

The last point is proved along the same lines that  $W_1$  is a distance. Namely let  $\nu_{12}$  and  $\nu_{23}$  be appropriate couplings to get  $U^{s_1}(\xi_1, \xi_2)$  and  $U^{s_2}(\xi_2, \xi_3)$ . There exists  $\gamma$  a probability on  $E^3$  whose projections satisfy  $p_{12}\gamma = \nu_{12}$  and  $p_{23}\gamma = \nu_{23}$ . The coupling  $\nu_{13} := p_{13}\gamma$  gives (2.2).

**Definition 2.4. Noise sensitivity at large scale** The random walk  $(G, \mu)$  is *noise sensitive at large scale* if there exists a sequence such that  $s_n = o(\mathbb{E}d_G(X_n, X'_n))$  and  $(G, \mu)$  is noise sensitive at scale  $s_n$ .

The *spread*  $\mathbb{E}d_G(X_n, X'_n)$  of the random walk can be considered as the natural scale to describe the distribution  $\mu_n$ . Noise sensitivity at large scale essentially means that most refreshed samples  $Y_n^\rho$  look independent of  $X_n$  at the spread scale. For instance, the central limit theorem for virtually

abelian groups provides a gaussian description of the distribution of  $\mu_n$  at the spread scale  $\sqrt{n}$ . For such a group, noise sensitivity at large scale implies that the limit gaussian distribution of  $(X_n, Y_n^\rho)$  is the same as that of  $(X_n, X'_n)$ .

**Definition 2.5. Noise sensitivity in average distance.** Given a left-invariant distance  $d_G$  on  $G$  (e.g. a word distance when  $G$  is finitely generated), the random walk  $(G, \mu)$  is *noise sensitive in average distance* if

$$\forall \rho \in (0, 1), \quad \liminf_{n \rightarrow \infty} \frac{\mathbb{E}d_G(X_n, Y_n^\rho)}{\mathbb{E}d_G(X_n, X'_n)} \geq 1.$$

The random walk  $(G, \mu)$  is *partially noise sensitive in average distance* if

$$\exists c > 0, \forall \rho \in (0, 1), \exists N, \forall n \geq N, \quad \mathbb{E}d_G(X_n, Y_n^\rho) \geq c\mathbb{E}d_G(X_n, X'_n).$$

### 3. Relationships between notions of noise sensitivity

The  $\ell^1$  noise sensitivity seems to be the strongest among the notions defined above. It implies noise sensitivity at any (in particular at large) scale.

3.1.  *$\ell^1$ -noise sensitivity and entropy noise sensitivity.* We believe that  $\ell^1$ -noise sensitivity implies entropy noise sensitivity, which would be a consequence of Conjecture 1.7. We prove it under an additional assumption of homogeneity.

**Definition 3.1.** A sequence  $(\xi_n)$  of probability measures on  $G$  has *homogeneous entropy* if

$$h_n(\varepsilon) := \sup \left\{ \frac{1}{H(\xi_n)} \sum_{x \in A} \xi_n(x) \log(\xi_n(x)) \mid A \subset G, \xi_n(A) \leq \varepsilon \right\} \xrightarrow{\varepsilon \rightarrow 0} 0$$

uniformly in  $n$ .

Informally, this means that the smallest atoms do not contribute much to the entropy.

**Proposition 3.2.** *If a random walk  $(G, \mu)$  is  $\ell^1$ -noise sensitive and if the sequence of measures  $(\pi_n^\rho)_n$  has homogeneous entropy for each  $\rho \in (0, 1]$ , then  $(G, \mu)$  is entropy noise sensitive.*

*Proof:* The case of finite groups is trivial by Proposition 5.1. We assume  $G$  is infinite. Let  $\nu_n$  be a coupling between two sequences  $(\xi_n^1)$  and  $(\xi_n^2)$  with homogeneous entropy such that  $\nu_n(x_1 \neq x_2) \rightarrow 0$ . Then for small  $\varepsilon > 0$  and each  $i = 1, 2$ , denote  $A_\varepsilon^i := \{(1 - \varepsilon)\xi_n^i(x) \leq \nu_n(x, x) \leq \xi_n^i(x)\}$ . For  $n$  large enough,  $\xi_n^i(A_\varepsilon^i) \geq 1 - \varepsilon$ , thus  $\nu_n\{(x, x) : x \in A_\varepsilon^i\} \geq (1 - \varepsilon)^2 \geq 1 - 2\varepsilon$  and by taking marginal  $\xi_n^{3-i}(A_\varepsilon^i) \geq 1 - 2\varepsilon$ . It follows that  $\xi_n^i(A_\varepsilon^1 \cap A_\varepsilon^2) \geq 1 - 3\varepsilon$ .

As the modulus of continuity of the function  $x \log(x)$  is  $\omega(\varepsilon) = \varepsilon |\log \varepsilon|$ , one has for all  $x, y \in (0, 1]$

$$\begin{aligned} \left| \frac{x}{y} - 1 \right| \leq \varepsilon &\iff |x - y| \leq \varepsilon y \implies |x \log x - y \log y| \leq \omega(\varepsilon y) \leq \varepsilon y |\log \varepsilon + \log y| \\ &\implies \left| \frac{x \log x}{y \log y} - 1 \right| \leq \frac{\varepsilon |\log \varepsilon|}{|\log y|} + \varepsilon \end{aligned}$$

Using the homogeneity assumption:

$$\begin{aligned} H(\xi_n^i) &= - \sum_x \xi_n^i(x) \log(\xi_n^i(x)) = - \sum_{x \in A_\varepsilon^1 \cap A_\varepsilon^2} \xi_n^i(x) \log(\xi_n^i(x)) + h_n^i(3\varepsilon)H(\xi_n^i) \\ &= - \sum_{x \in A_\varepsilon^1 \cap A_\varepsilon^2} \nu_n(x, x) \log(\nu_n(x, x)) \frac{\xi_n^i(x) \log(\xi_n^i(x))}{\nu_n(x, x) \log(\nu_n(x, x))} + h_n^i(3\varepsilon)H(\xi_n^i) \\ &= -(1 + \psi_n^i(\varepsilon)) \sum_{x \in A_\varepsilon^1 \cap A_\varepsilon^2} \nu_n(x, x) \log(\nu_n(x, x)) + h_n^i(3\varepsilon)H(\xi_n^i), \end{aligned}$$

where  $|\psi_n^i(\varepsilon)| \leq \frac{\varepsilon|\log \varepsilon|}{\log 2} + \varepsilon \xrightarrow{\varepsilon \rightarrow 0} 0$  provided all the atoms of  $\nu_n$  have mass at most  $\frac{1}{2}$ . We apply these equalities for  $\xi_n^1 = \pi_n^\rho$  and  $\xi_n^2 = \mu_n^2 = \pi_n^1$  (the condition on the atoms of  $\nu_n$  is satisfied for  $n$  large because  $G$  is infinite). We get that  $\sum_{x \in A_\varepsilon^1 \cap A_\varepsilon^2} \nu_n(x, x) \log(\nu_n(x, x)) = H(\pi_n^\rho) + o(H(\pi_n^\rho)) = H(\mu_n^2) + o(H(\mu_n^2))$ . It follows that  $\frac{H(\pi_n^\rho)}{H(\mu_n^2)} \rightarrow 1$ , which is equivalent to entropy noise sensitivity.  $\square$

*Remark 3.3.* For random walks on groups, homogeneous entropy, as well as homogeneous spread (see Definition 3.4 below), are related to the tail decay of the distribution  $\mu_n$ . They are not easy to check on a given group, for it is usually hard to compute entropy or spread in the first place. However they are satisfied for random walks on virtually abelian groups by gaussian decay. They are also satisfied by symmetric random walks on the lamplighter group  $\mathbb{Z}/2\mathbb{Z} \wr \mathbb{Z}$ , because of the gaussian decay of the range of a random walk on  $\mathbb{Z}$  – see formula (7.1) in the proof of Theorem 7.1. Such argument on the range can probably be adapted to the case of iterated lamplighter groups, and possibly to the case of diagonal products of lamplighter groups as in Briussel and Zheng (2021). However detailed computations are out of the scope of the present article. Non-Liouville random walks also have homogeneous entropy and spread by Shannon’s theorems, see e.g. Kaïmanovich and Vershik (1983). We do not know if there exist random walks on groups with non-homogeneous entropy or non-homogeneous spread.

3.2. *Noise sensitivity at large scale and noise sensitivity in average distance.* We also believe that if  $(G, \mu)$  is noise sensitive at large scale, then it is noise sensitive in average (word) distance. Remark 4.4 shows that the converse is not true. Again, we prove this under an additional assumption.

**Definition 3.4.** We say the random walk  $(G, \mu)$  has *homogeneous spread* if

$$f_n(\varepsilon) := \sup \left\{ \frac{1}{\mathbb{E}d_G(X_n, X'_n)} \sum_{(x,y) \in A} d_G(x, y) \mu_n^2(x, y) \mid A \subset G \times G, \mu_n^2(A) \leq \varepsilon \right\} \xrightarrow{\varepsilon \rightarrow 0} 0$$

uniformly in  $n$ .

Informally, this means that far away points do not contribute to the spread.

**Proposition 3.5.** *If a random walk  $(G, \mu)$  is noise sensitive at large scale and has homogeneous spread, then it is noise sensitive in average word distance.*

*Proof:* There exists a sequence  $s_n = o(\mathbb{E}d_G(X_n, X'_n))$  such that  $U^{s_n}(\pi_n^\rho, \mu_n^2) \rightarrow 0$ . So there exists a probability  $\nu_n \in \text{Coup}(\pi_n^\rho, \mu_n^2)$  such that  $\nu_n(A_n) \rightarrow 1$  where  $A_n$  is the event  $d_{G^2}((x, y), (x', y')) \leq s_n$ . We compute

$$\begin{aligned} \mathbb{E}d_G(X_n, Y_n^\rho) &= \sum_{G^2} \pi_n^\rho(x, y) d_G(x, y) = \sum_{G^4} \nu_n(x, y, x', y') d_G(x, y) \\ &\geq \sum_{A_n} \nu_n(x, y, x', y') d_G(x', y') - 2s_n \quad \text{by triangle inequality,} \\ &= \sum_{G^2} \nu_n((G^2 \times \{(x', y')\}) \cap A_n) d_G(x', y') - 2s_n. \end{aligned}$$

Given  $\varepsilon > 0$ , for  $n$  large enough there exist  $B_n \subset G^2$  such that  $\mu_n^2(B_n) \geq 1 - \varepsilon$  and

$$\forall (x', y') \in B_n, \quad \nu_n((G^2 \times \{(x', y')\}) \cap A_n) \geq (1 - \varepsilon) \mu_n^2(x', y').$$

Then

$$\begin{aligned} \mathbb{E}d_G(X_n, Y_n^\rho) &\geq (1 - \varepsilon) \sum_{B_n} \mu_n^2(x', y') d_G(x', y') - 2s_n \\ &\geq (1 - \varepsilon)(1 - f_n(\varepsilon)) \mathbb{E}_{\mu_n^2} d_G(x', y') - 2s_n. \end{aligned}$$



Noise sensitivity in average distance now follows from the homogeneity assumption.  $\square$

#### 4. Noise sensitivity and Liouville property

A function on a group  $G$  is  $\mu$ -harmonic if  $f(y) = \sum_{x \in G} \mu(x)f(yx)$  for all  $y$ . A random walk  $(G, \mu)$  is Liouville if there are no non-constant bounded  $\mu$ -harmonic functions. This is equivalent to the fact that the Poisson boundary is reduced to a point [Furstenberg \(1971\)](#) and also equivalent to the fact that the entropy  $H(X_n)$  of the random walk is sublinear [Kaĭmanovich and Vershik \(1983\)](#).

**Theorem 4.1.** *If  $(G, \mu)$  is  $\ell^1$ -noise sensitive, then  $(G, \mu)$  is Liouville.*

The converse is not true because of infinite abelian groups, see [Proposition 5.2](#).

*Proof:* First observe that  $\ell^1$ -noise sensitivity implies

$$\|\mu^{\text{diag}} * \mu_{n-1}^2 - \mu_n^2\|_1 \xrightarrow{n \rightarrow \infty} 0. \quad (4.1)$$

Indeed, taking the convex decomposition of the first factor  $\pi^\rho = (1 - \rho)\mu^{\text{diag}} + \rho\mu^2$ , the  $\ell^1$ -noise sensitivity first implies  $(1 - \rho)\|\mu^{\text{diag}} * \pi_{n-1}^\rho - \mu_n^2\|_1 \rightarrow 0$ . On the other hand  $\|\mu^{\text{diag}} * \pi_{n-1}^\rho - \mu^{\text{diag}} * \mu_{n-1}^2\|_1 \rightarrow 0$  by left multiplication. Convergence [\(4.1\)](#) follows.

Assume by contradiction that  $(G, \mu)$  is not Liouville, i.e. admits a non-trivial Poisson boundary  $(\Pi, \nu)$ . We claim that there exists a subset  $A \subset \Pi$  such that

$$\nu(A)^2 < \sum_{z \in G} \mu(z)\nu(z^{-1}A)^2. \quad (4.2)$$

Indeed, as  $\nu$  is stationary  $\nu(A) = \sum_{z \in G} \mu(z)\nu(z^{-1}A)$  and the above inequality [\(4.2\)](#) holds with  $\leq$  in place of  $<$  by Jensen. As the square function is strictly convex, equality implies  $\nu(z^{-1}A) = \nu(A)$  for all  $z$  in the support of  $\mu$  and a fortiori for all  $z$ . Therefore equality for all  $A$  would imply triviality of the Poisson boundary, whence the claim.

Now by stationarity of  $\nu^2$  with respect to  $\mu^2$ , we have for each  $n$  that

$$\nu(A)^2 = \sum_{(x,y) \in G \times G} \mu_n^2(x,y)\nu(x^{-1}A)\nu(y^{-1}A)$$

and

$$\sum_{z \in G} \mu(z)\nu(z^{-1}A)^2 = \sum_{(x,y) \in G \times G} \mu^{\text{diag}} * \mu_{n-1}^2(x,y)\nu(x^{-1}A)\nu(y^{-1}A)$$

Convergence [\(4.1\)](#) implies that the difference between right hand sides tends to zero, raising a contradiction to [\(4.2\)](#) that  $\nu(A)^2 = \sum_{z \in G} \mu(z)\nu(z^{-1}A)^2$ .  $\square$

For the free groups, this theorem can be improved.

**Theorem 4.2.** *Finitely supported random walks on free groups are not noise sensitive at large scale.*

*Proof:* We follow the same line of reasoning as in the previous proof. We are now given a sublinear sequence  $s_n$  such that  $U^{s_n}(\pi_n^\rho, \mu_n^2) \rightarrow 0$ . By diagonal coupling of the first factor, we deduce  $U^{s_{n-1}}(\pi_n^\rho, \pi^\rho * \mu_{n-1}^2) \rightarrow 0$ . By [\(2.2\)](#) in [Section 2.2](#), we obtain:

$$U^{s_n + s_{n-1}}(\pi^\rho * \mu_{n-1}^2, \mu_n^2) \rightarrow 0. \quad (4.3)$$

Let us consider a subset  $A$  satisfying [\(4.2\)](#) in the Poisson boundary which is now the geometric boundary of the Cayley tree, i.e. the set of infinite geodesic rays out of the neutral element  $e$ . Up to taking a close enough approximation, we may assume that  $A$  is open-closed, i.e. a finite union of "cylinder" sets. Recall that "cylinder" sets  $C_h$  are indexed by  $h \in \mathbb{F}$ , and  $C_h$  is the collection of geodesic rays out of  $e$  crossing  $h$ . Moreover, the point  $h$  partitions  $\mathbb{F} \setminus \{h\}$  into the set  $\hat{C}_h$  consisting



of points  $g$  from which there is a geodesic ray towards  $C_h$  not going through  $h$  and its complement  $\hat{C}_h^e$ .

Recall that the speed (or drift)  $\lambda := \lim \frac{1}{n} \mathbb{E}d(e, X_n) > 0$  is positive. Given  $\alpha \in (0, \lambda)$ , we define  $[x]_n$  to be the point at distance  $\alpha n$  from  $e$  on the geodesic ray from  $e$  to  $x$ , when  $x$  is at distance at least  $\alpha n$  from the identity. Otherwise  $[x]_n = e$ . We will use the fact that

$$\mathbb{P}(d(e, X_n) \geq \alpha n) \xrightarrow{n \rightarrow \infty} 1. \tag{4.4}$$

We claim that

$$\left| \sum_{G^2} \mu_n^2(g) g\nu(A) - \sum_{G^2} \mu_n^2(g) [g]_n \nu(A) \right| \xrightarrow{n \rightarrow \infty} 0 \tag{4.5}$$

where we write  $g\nu(A) := \nu(x^{-1}A)\nu(y^{-1}A)$  for  $g = (x, y)$  and  $[g]_n = ([x]_n, [y]_n)$ . Convoluting by  $\pi^\rho$  which has finite support, we get

$$\left| \sum_{G^2} \pi^\rho * \mu_{n-1}^2(g) g\nu(A) - \sum_{G^2} \pi^\rho * \mu_{n-1}^2(g) [g]_n \nu(A) \right| \xrightarrow{n \rightarrow \infty} 0. \tag{4.6}$$

Now let  $\nu_n$  denote appropriate couplings to get (4.3), we have:

$$\begin{aligned} & \left| \sum_{G^2} \mu_n^2(g) [g]_n \nu(A) - \sum_{G^2} \pi^\rho * \mu_{n-1}^2(g) [g]_n \nu(A) \right| \\ &= \left| \sum_{(g_1, g_2) \in G^4} \nu_n(g_1, g_2) ([g_1]_n \nu(A) - [g_2]_n \nu(A)) \right| \\ &\leq 2 \sum_{(g_1, g_2) \in G^4} \nu_n(g_1, g_2) \mathbf{1}_{\{[g_1]_n \neq [g_2]_n\}} \\ &\leq U^{s_n + s_{n-1}} (\pi^\rho * \mu_{n-1}^2, \mu_n^2) + 4\mathbb{P}[d_G(e, X_n) \leq \alpha n + s_n + s_{n-1}] \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

where the last inequality is due to the geometry of the tree that if both  $x, x'$  are at distance  $\geq \alpha n + s$  from identity and  $d(x, x') \leq s$ , then  $[x]_n = [x']_n$ . The convergence to 0 follows by (4.3) and (4.4) for  $\alpha < \alpha' < \lambda$  as  $\alpha n + s_n + s_{n-1} < \alpha' n$  for  $n$  large. Together with (4.5) and (4.6), we deduce

$$\left| \sum_{G^2} \mu_n^2(g) g\nu(A) - \sum_{G^2} \pi^\rho * \mu_{n-1}^2(g) g\nu(A) \right| \xrightarrow{n \rightarrow \infty} 0$$

raising the same contradiction as at the end of the proof of Theorem 4.1.

There remains to prove the claim (4.5). As  $A$  is a finite union of cylinder sets, it is enough to prove that for any  $h_1, h_2$  in  $\mathbb{F}$ , one has

$$\left| \sum_{(x,y) \in G^2} \mu_n(x) x\nu(C_{h_1}) \mu_n(y) y\nu(C_{h_2}) - \sum_{(x,y) \in G^2} \mu_n(x) [x]_n \nu(C_{h_1}) \mu_n(y) [y]_n \nu(C_{h_2}) \right| \xrightarrow{n \rightarrow \infty} 0 \tag{4.7}$$

We use the facts that  $x\nu(A) = \nu(x^{-1}A)$  is the probability that a random walk started in  $x$  tends to a boundary point in  $A$ , and that for any  $\varepsilon > 0$  there exists  $N$  such that when  $d(x, z) \geq N$ , the probability that a random walk started in  $x$  ever hits  $z$  is at most  $\varepsilon$ . Using union bounds, we can deduce that for any  $R \geq 0$ , there exists  $N'$  such that when  $d(x, h_1) \geq N'$ , a random walk started at  $x$  hits the ball  $B(h_1, R)$  with probability at most  $\varepsilon$ . We use it for  $R$  large enough that a random walk trajectory cannot move from  $\hat{C}_{h_1}$  to its complement (or vice-versa) without stepping in  $B(h_1, R)$ .

It follows that for  $d(e, x) \geq \alpha n \geq N' + d(e, h_1)$  we have

$$\begin{aligned} &\text{if } x \in \hat{C}_{h_1}, \text{ then } x\nu(C_{h_1}) \geq 1 - \varepsilon \text{ and } [x]_n\nu(C_{h_1}) \geq 1 - \varepsilon, \\ &\text{if } x \in \hat{C}_{h_1}^c, \text{ then } x\nu(C_{h_1}) \leq \varepsilon \text{ and } [x]_n\nu(C_{h_1}) \leq \varepsilon. \end{aligned}$$

Finally, let us denote  $\Lambda_n$  the set of points  $(x, y)$  in  $G^2$  such that  $d(e, x) \geq \alpha n$  and  $d(e, y) \geq \alpha n$  where  $n$  is large enough to have  $d(e, x) \geq \alpha n \geq N' + d(e, h_1)$  and the analogous property for  $y, h_2$  in place of  $x, h_1$ . For  $(x, y) \in \Lambda_n$ , one has

$$|x\nu(C_{h_1})y\nu(C_{h_2}) - [x]_n\nu(C_{h_1})[y]_n\nu(C_{h_2})| \leq 5\varepsilon$$

by checking the four cases. Then the expression in (4.7) is

$$\left| \sum_{(x,y) \in G^2} \mu_n(x)\mu_n(y) (x\nu(C_{h_1})y\nu(C_{h_2}) - [x]_n\nu(C_{h_1})[y]_n\nu(C_{h_2})) \right| \leq 2\mu_n^2(\Lambda_n^c) + 5\varepsilon\mu_n^2(\Lambda_n) \leq 4\varepsilon + 5\varepsilon$$

where  $\mu_n^2(\Lambda_n^c)$  is bounded above using (4.4) for  $n$  large. □

The above proof can probably be generalised to groups with hyperbolic properties. However in full generality, it is not clear if sublinearly close points have close actions on the harmonic measure.

**Theorem 4.3.** *If  $(G, \mu)$  is partially entropy noise sensitive, then  $(G, \mu)$  is Liouville.*

This is one direction of Conjecture 1.7.

*Proof:* Assume  $(G, \mu)$  is not Liouville, which means that the associated Poisson boundary is not trivial. By [Kaĭmanovich and Vershik \(1983\)](#), this implies

$$h_\infty(\mu) := \lim_{n \rightarrow \infty} \frac{H(\mu_n)}{n} > 0.$$

Thus  $H(X_n) = H(\mu_n) = h_\infty(\mu)n + o(n)$ . Similarly

$$h_\infty(\pi^\rho) := \lim_{n \rightarrow \infty} \frac{H(\pi_n^\rho)}{n} \geq h_\infty(\mu)$$

because  $H(\pi_n^\rho) = H(X_n, Y_n^\rho) = H(Y_n^\rho | X_n) + H(X_n) \geq H(X_n) = H(\mu_n)$ . It follows that  $H(\pi_n^\rho) = h_\infty(\pi^\rho)n + o(n)$ .

Moreover  $\pi^\rho = (1 - \rho)\mu^{\text{diag}} + \rho\mu^2$  tends to  $\mu^{\text{diag}}$  weakly as  $\rho \rightarrow 0$ , and  $H(\pi^\rho) \rightarrow H(\mu^{\text{diag}}) = H(\mu)$  as these measures have support in the finite set  $\text{supp}(\mu)^2$ . The upper semi-continuity of asymptotic entropy [Amir et al. \(2013, Proposition 4\)](#) implies

$$\lim_{\rho \rightarrow 0} h_\infty(\pi^\rho) = h_\infty(\mu).$$

Therefore we have

$$\lim_{n \rightarrow \infty} \frac{H(Y_n^\rho | X_n)}{H(X_n)} = \frac{h_\infty(\pi^\rho) - h_\infty(\mu)}{h_\infty(\mu)} \xrightarrow{\rho \rightarrow 0} 0,$$

ruling out partial entropy noise sensitivity. □

*Remark 4.4.* Noise sensitivity in average distance does not imply Liouville property. Indeed for the simple random walk on the free group  $\mathbb{F}_2$ , one easily computes  $\mathbb{E}d_{\mathbb{F}_2}(X_n, X'_n) \sim n \sim \mathbb{E}d_{\mathbb{F}_2}(X_n, Y_n^\rho)$  for any  $\rho > 0$ . It follows that average distance noise sensitivity does not imply large scale noise sensitivity.

### 5. Finite and abelian groups

#### 5.1. Finite groups.

**Proposition 5.1.** *Any finite group is  $\ell^1$ -noise sensitive, entropy noise sensitive and noise sensitive in average distance with respect to any generating probability measure.*

*Proof:* The measure  $\pi^\rho := (1 - \rho)\mu_{\text{diag}} + \rho\mu^2$  is generating of  $G^2$ , so  $\pi_n^\rho$  satisfies  $\|\pi_n^\rho - \text{unif}_{G^2}\|_1 \leq \beta_1^n$  for some  $\beta_1 < 1$  (see e.g. [Saloff-Coste \(2004\)](#)). Similarly  $\|\mu_n^2 - \text{unif}_{G^2}\|_1 \leq \beta_2^n$  for some  $\beta_2 < 1$ . Therefore the random walk in  $G$  is  $\ell^1$ -noise sensitive, with total variation converging exponentially fast to zero.

Since the set  $G$  is finite, it follows that  $H(\pi_n^\rho) \rightarrow H(\text{unif}_{G^2}) = 2H(\text{unif}_G)$  and  $H(\mu_n) \rightarrow H(\text{unif}_G)$ . This implies entropy noise sensitivity.

Moreover let  $Z$  denote a uniformly random variable in  $G$ , it is immediate that  $|\mathbb{E}d_G(X_n, X'_n) - \mathbb{E}d_G(e, Z)| \leq \text{diam}(G)\|\mu_n - \text{unif}_G\|_1$  and

$$|\mathbb{E}d_G(X_n, Y_n^\rho) - \mathbb{E}d_G(e, Z)| \leq \text{diam}(G) (\|\mu_n - \text{unif}_G\|_1 + \|\pi_n^\rho - \text{unif}_{G^2}\|_1).$$

By  $\ell^1$ -noise sensitivity, both converge to zero, whence distance noise sensitivity. □

This result is not surprising since we defined noise sensitivity only asymptotically. An interesting further question is whether it is possible that noise sensitivity manifests itself before cut-off, that is before the random walk seems equidistributed. However we will not pursue in this direction.

#### 5.2. Infinite abelian groups.

In abelian groups, computations are explicit, based on the local central limit theorem.

**Proposition 5.2.** *Let  $G$  be an infinite finitely generated abelian group and  $\mu$  be any finitely supported probability measure. Then*

- $(G, \mu)$  is entropy noise sensitive,
- $(G, \mu)$  is not, even partially, noise sensitive in average word distance,
- $(G, \mu)$  is not  $\ell^1$ -noise sensitive.

The proof of the first point is based on the following folklore lemma.

**Lemma 5.3.** *Let  $X_n$  be a random walk on  $\mathbb{Z}^d$  with step distribution  $\mu$  finitely supported and generating. Then*

$$H(X_n) = \frac{d}{2} \log n + o(\log n).$$

*Proof of Lemma 5.3:* We denote  $m = \mathbb{E}X_1$ . We fix  $\varepsilon > 0$ . The central limit theorem gives  $K$  such that  $\mathbb{P}(|X_n - nm| \leq K\sqrt{n}) \geq 1 - \varepsilon$ . Then

$$\begin{aligned} H(X_n) &= H(X_n | |X_n - nm| \leq K\sqrt{n})\mathbb{P}(|X_n - nm| \leq K\sqrt{n}) \\ &\quad + H(X_n | |X_n - nm| > K\sqrt{n})\mathbb{P}(|X_n - nm| > K\sqrt{n}) + h_1, \end{aligned} \tag{5.1}$$

where  $0 \leq h_1 \leq \log 2$  is the entropy of the conditioning partition. The local central limit theorem [Davis and McDonald \(1995\)](#); [Lawler and Limic \(2010\)](#) provides two constants  $c_1, c_2 > 0$  depending only on  $K$  such that for all  $n$  large enough, for all  $|x - nm| \leq K\sqrt{n}$ ,

$$\frac{c_1}{n^{\frac{d}{2}}} \leq \mathbb{P}(X_n = x | |X_n - nm| \leq K\sqrt{n}) \leq \frac{c_2}{n^{\frac{d}{2}}} \tag{5.2}$$

It follows that for  $n$  large enough

$$\frac{d}{2} \log(n) + \log \frac{1 - \varepsilon}{c_2} = \log \frac{(1 - \varepsilon)n^{\frac{d}{2}}}{c_2} \leq H(X_n | |X_n - nm| \leq K\sqrt{n}) \leq \frac{d}{2} \log(n) + \log \frac{1 - \varepsilon}{c_1} \tag{5.3}$$

and the lower bound together with (5.1) give  $\liminf \frac{H(X_n)}{\log n} \geq \frac{d}{2}$ .

On the other hand, as  $X_n$  is confined in a ball of radius  $Cn$  and of volume at most  $C'n^d$ , we have

$$H(X_n | |X_n - nm| > K\sqrt{n}) \leq \log(C'n^d) = d \log(n) + \log(C').$$

For  $n$  large enough, we deduce with (5.1) and (5.3)

$$H(X_n) \leq (1 - \varepsilon) \frac{d}{2} \log(n) + \varepsilon d \log(n) + C''$$

which gives  $\limsup \frac{H(X_n)}{\log n} \leq \frac{d}{2}$ . □

*Proof of Proposition 5.2:* Such a group  $G$  is isomorphic to  $\mathbb{Z}^d \times F$  for some finite abelian group  $F$ . Lemma 5.3 still holds in this more general setting for the finite group  $F$  contributes at most a constant to entropy.

We first prove entropy noise sensitivity, we have to show  $\lim \frac{H(Y_n^\rho | X_n)}{\log n} = \frac{d}{2}$ . The inequality  $H(Y_n^\rho | X_n) \leq H(X_n)$  provides the upper bound.

Let  $\ell$  be the number of refreshed variables in  $Y_n^\rho$ . Given  $\ell$ , we can write  $(X_n, Y_n^\rho) = (X_{n-\ell} X_\ell, X_{n-\ell} X'_\ell)$  where  $X_\ell$  and  $X'_\ell$  are independent, of law  $\mu_\ell$ . They are also independent of  $X_{n-\ell}$  which has law  $\mu_{n-\ell}$ . As  $\ell$  follows a binomial law  $\mathcal{B}(n, \rho)$ , we have for any  $\varepsilon > 0$

$$\mathbb{P}(A_\varepsilon) = \mathbb{P}(\ell \in [(\rho - \varepsilon)n, (\rho + \varepsilon)n]) \xrightarrow{n \rightarrow \infty} 1 \tag{5.4}$$

where we denote  $A_\varepsilon = \{\ell \in [(\rho - \varepsilon)n, (\rho + \varepsilon)n]\}$ . As further conditioning does not increase entropy, we have the lower bound

$$H(Y_n^\rho | X_n) = H(X_{n-\ell} X'_\ell | X_n) \geq H(X_{n-\ell} X'_\ell | X_n, X_{n-\ell}, \ell) = H(X'_\ell | X_n, X_{n-\ell}, \ell) = H(X'_\ell | \ell).$$

It follows that

$$H(Y_n^\rho | X_n) \geq \sum_{k=0}^n H(X'_k) \mathbb{P}(\ell = k) \geq \sum_{k=(\rho-\varepsilon)n}^n H(X'_k) \mathbb{P}(\ell = k) \geq H(X'_{(\rho-\varepsilon)n}) \mathbb{P}(\ell \geq (\rho - \varepsilon)n)$$

and thus by Lemma 5.3 and (5.4)

$$H(Y_n^\rho | X_n) \geq \left( \frac{d}{2} \log((\rho - \varepsilon)n) + o(\log n) \right) (1 + o(1)) = \frac{d}{2} \log n + o(\log n)$$

which is the required lower bound to get entropy noise sensitivity.

Regarding the spread, there exist two constants  $c_3, c_4 > 0$  depending only on  $\mu$  such that

$$c_3 \sqrt{n} \leq \mathbb{E}d_G(X_n, X'_n) \leq c_4 \sqrt{n}. \tag{5.5}$$

Conditioning as above by the number  $\ell$  of refreshed variables, we have

$$\mathbb{E}d_G(X_n, Y_n^\rho) = \mathbb{E}d_G(X_{n-\ell} X_\ell, X_{n-\ell} X'_\ell) = \mathbb{E}d_G(X_\ell, X'_\ell) = \sum_{k=0}^{Cn} \mathbb{E}(d_G(X_\ell, X'_\ell) | \ell = k) \mathbb{P}(\ell = k)$$

where  $C = 2 \max\{d_G(e, x) | x \in \text{supp}\mu\}$ . By splitting the sum at  $(\rho + \varepsilon)n$  for arbitrary  $\varepsilon > 0$  and using (5.4), we get

$$\mathbb{E}d_G(X_n, Y_n^\rho) \leq c_4 \sqrt{(\rho + \varepsilon)n} \mathbb{P}(\ell \leq (\rho + \varepsilon)n) + c_4 \sqrt{Cn} \mathbb{P}(\ell > (\rho + \varepsilon)n) = c_4 \sqrt{\rho + \varepsilon} \sqrt{n} (1 + o(1)).$$

With (5.5) we deduce that

$$\limsup_{n \rightarrow \infty} \frac{\mathbb{E}d_G(X_n, Y_n^\rho)}{\mathbb{E}d_G(X_n, X'_n)} \leq \frac{c_4}{c_3} \sqrt{\rho} \xrightarrow{\rho \rightarrow 0} 0.$$

This rules out partial noise sensitivity in average word distance, and thus  $\ell^1$ -noise sensitivity by Proposition 3.5. □

**Corollary 5.4.** *Let  $(G, \mu)$  be a group with a surjective homomorphism  $G \twoheadrightarrow \mathbb{Z}$  and  $\mu$  finitely supported, then  $(G, \mu)$  is not  $\ell^1$ -noise sensitive.*

*Proof:* It follows from Proposition 5.2 because  $\ell^1$ -noise sensitivity is preserved under taking quotients  $G \rightarrow \bar{G}$  (for the induced measure): if  $\mu_1, \mu_2$  are two measures on  $G$ , then  $\|\bar{\mu}_1 - \bar{\mu}_2\|_1 \leq \|\mu_1 - \mu_2\|_1$ .  $\square$

### 5.3. Product groups.

**Proposition 5.5.** *Let  $G_1, G_2$  be two finitely generated groups with probabilities  $\mu_1$  and  $\mu_2$  respectively. If both  $(G_1, \mu_1)$  and  $(G_2, \mu_2)$  are noise sensitive in the sense of one definition of section 2, then  $(G_1 \times G_2, \mu_1 \times \mu_2)$  is noise sensitive as well in this definition.*

*Proof:* It follows straightforwardly from the definitions.  $\square$

The choice of product probability measure is important here, because noise sensitivity notions may depend on the choice of probability. It is the case for  $\ell^1$ -noise sensitivity and noise sensitivity in average distance by Corollary 6.3.

## 6. The infinite dihedral group

Let  $D_\infty = \langle a, b \mid a^2 = b^2 = 1 \rangle$  denote the infinite dihedral group. The random walk on  $D_\infty$  driven by  $\frac{1}{2}(\delta_a + \delta_b)$  is called simple. The random walk driven by  $\frac{1}{3}(\delta_e + \delta_a + \delta_b)$  is called lazy simple.

**Theorem 6.1.** *Let  $D_\infty$  be the infinite dihedral group.*

- *The simple random walk on  $D_\infty$  is not  $\ell^1$ -noise sensitive.*
- *The lazy simple random walk on  $D_\infty$  is  $\ell^1$ -noise sensitive.*

The Cayley graph of  $D_\infty$  with respect to the generating set  $\{a, b\}$  is a line, so it coincides with the Cayley graph of  $\mathbb{Z}$ . Therefore random walks on  $D_\infty$  are related to random walks on the integers, which explains the first statement. However the edge labelings are very different. This difference will be key to the second statement.

**Corollary 6.2.** *The simple random walk on  $D_\infty$  is not noise sensitive in average word distance. The lazy simple random walk on  $D_\infty$  is noise sensitive in average word distance.*

*Proof of Corollary 6.2:* The first statement is justified below in the proof of Theorem 6.1. The second statement follows from Proposition 3.5.  $\square$

Note that both simple and lazy simple random walks on  $D_\infty$  are entropy noise sensitive by the same argument as in abelian case.

**Corollary 6.3.** *The  $\ell^1$ -noise sensitivity and the noise sensitivity in average word distance are not group properties but depend on the probability measure. A fortiori they are not preserved under quasi-isometries.*

**Corollary 6.4.** *For any positive integer  $d$ , there exists a group commensurable with  $\mathbb{Z}^d$  which is  $\ell^1$ -noise sensitive for some probability measure.*

*Proof:* Take a direct product  $D_\infty^d$  and use Proposition 5.5.  $\square$

*Remark 6.5.* Theorem 6.1 shows that virtually abelian groups may be  $\ell^1$  noise sensitive. Informally, this is due to the noise sensitivity of the action of the finite quotient on the torsion free subgroup. It is obvious that if we are given a trajectory in the streets of New York by a sequence of moves North-South-East-West and we misread one instruction, we will still end up close to the aim. However if the instructions are given in terms of Forward-Backward-Left-Right and we miss a turn, we will most likely end up very far from the aim. It would be interesting to understand precisely when a virtually abelian group is noise sensitive or not.

6.1. *Couplings of (lazy or not) simple random walks on  $\mathbb{Z}$ .* We record here three lemmas that will be used in the proof of Theorem 6.1. The first one is very classical and the other two are easy consequences.

**Lemma 6.6.** *Let  $S_n$ , resp.  $S'_n$ , denote a (lazy or not) simple random walk on  $\mathbb{Z}$  started at 0, resp. at  $x\sqrt{n}$ . There exists  $\sigma = \sigma_x \in \{0, 1\}$ , such that*

$$\limsup_{n \rightarrow \infty} \|\text{law}(S_n) - \text{law}(S'_n - \sigma)\|_1 \xrightarrow{x \rightarrow 0} 0.$$

*For lazy simple random walks, this holds true with  $\sigma = 0$ .*

*Proof:* Let us start with the case of (non lazy) simple random walk. We assume  $x\sqrt{n}$  is even and take  $\sigma = 0$ . The case  $x\sqrt{n}$  odd and  $\sigma = 1$  is similar. We consider the following coupling of  $S_n$  and  $S'_n$  : while  $S_k \neq S'_k$ , sample the increments independently, once  $S_k = S'_k$  take the same increments. This ensures that  $S_k = S'_k$  for any  $k \geq T = \min\{k : S_k = S'_k\}$ . We have  $\|\text{law}(S_n) - \text{law}(S'_n - \sigma)\|_1 \leq 2\mathbb{P}(T > n)$ .

Now while  $k \leq T$ , the random variable  $S'_k - S_k$  has the law of a random walk with step distribution  $\frac{1}{2}\delta_0 + \frac{1}{4}\delta_2 + \frac{1}{4}\delta_{-2}$ . Then  $\mathbb{P}(T > n)$  is the probability that such a random walk, starting at the even integer  $x\sqrt{n}$ , does not hit 0 by time  $n$ . The later tends to 0 with  $x$  by Révész (2013, Theorem 2.13).

For lazy simple random walks, the random variable  $S'_k - S_k$  has a step distribution charging positively each increment in  $\{-2, -1, 0, 1, 2\}$ , and it can hit 0 as above without parity issue.  $\square$

We will actually rather use the proof than the lemma, and we call the coupling above the standard coupling between two simple random walks on  $\mathbb{Z}$  started at different positions.

**Lemma 6.7.** *Let  $S_n$ , resp.  $S'_n$ , denote a (lazy or not) simple random walk on  $\mathbb{Z}$  started at 0, resp. at  $x\sqrt{n}$ , and let  $\delta \in \mathbb{R}$ . There exists  $\sigma \in \{0, 1\}$  such that*

$$\limsup_{n \rightarrow \infty} \|\text{law}(S_n) - \text{law}(S'_{(1+\delta)n} - \sigma)\|_1 \rightarrow 0$$

*when both  $x$  and  $\delta$  tend to 0. This holds true with  $\sigma = 0$  for lazy simple random walks.*

*Proof:* This is a consequence of the previous lemma. We consider only the (non lazy) simple random walk and assume that  $\delta > 0$  and that both  $x\sqrt{n}$  and  $\delta n$  are even. Similar arguments apply in the other cases. We first sample  $S'_{\delta n}$ . By the central limit theorem, for any  $\varepsilon > 0$ , there exists  $A > 0$  such that for  $n$  large

$$\mathbb{P}\left[|S'_{\delta n}| \leq A\sqrt{\delta n}\right] \geq 1 - \varepsilon.$$

Conditioning on this event, we apply Lemma 6.6 to  $S_n$  and  $S'^{-1}_{\delta n} S'_{(1+\delta)n}$  which is starting at  $x'\sqrt{n}$  with  $x - A\delta \leq x' \leq x + A\delta$ . The probability to achieve their standard coupling is greater than  $1 - \varepsilon$  as long as  $x$  and  $\delta$  are small enough. It follows that the probability to couple  $S_n$  and  $S'_{(1+\delta)n}$  is at least  $1 - 2\varepsilon$  for  $n$  large.  $\square$

In order to ease notations, we write  $[a \pm b]$  as a shortcut for the interval  $[a - b, a + b]$ .

**Lemma 6.8.** *Assume  $\ell^+, \ell^- \in [(\frac{4}{9} \pm \delta)n]$  and  $\ell \in [(1 \pm \delta)n]$ , then for any given integer  $s$ ,*

$$\limsup_{n \rightarrow \infty} \left\| \left( \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1} \right)^{* \ell^+} * \left( \frac{1}{2}\delta_0 + \frac{1}{4}\delta_1 + \frac{1}{4}\delta_{-1} \right)^{* \ell^-} - \delta_s * \left( \frac{1}{3}\delta_0 + \frac{1}{3}\delta_1 + \frac{1}{3}\delta_{-1} \right)^{* \ell} \right\|_1 \xrightarrow{\delta \rightarrow 0} 0.$$

We could replace  $s$  by  $x\sqrt{n}$  with  $x$  tending to 0 but this will not be necessary. There is no shift  $\sigma \neq 0$  because of the lazy aspect of the random walk.

*Proof:* Fix  $\varepsilon > 0$ . Set  $\mu_S = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$  and  $\mu_{LS} = \frac{1}{3}\delta_0 + \frac{1}{3}\delta_1 + \frac{1}{3}\delta_{-1}$ . Then

$$\left( \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1} \right)^{* \ell^+} * \left( \frac{1}{2}\delta_0 + \frac{1}{4}\delta_1 + \frac{1}{4}\delta_{-1} \right)^{* \ell^-} = \mu_S^{* \ell^+} * \left( \frac{3}{4}\mu_{LS} + \frac{1}{4}\delta_0 \right)^{* \ell^-} = \mu_S^{* \ell^+} * (\mu_{LS})^{* m}$$

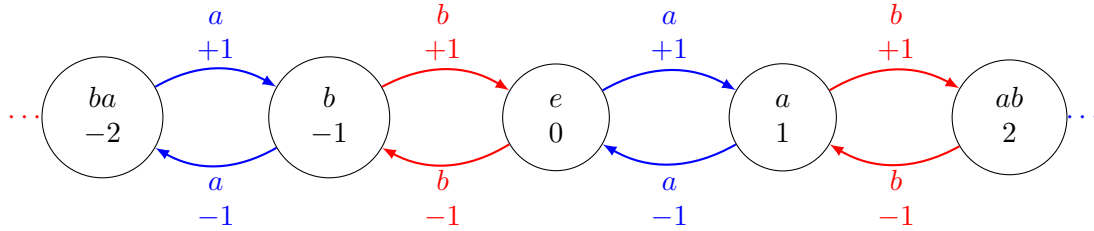


FIGURE 6.1. Cayley graph of  $D_\infty$  and identification with  $\mathbb{Z}$ .

with  $1 - \varepsilon \leq \mathbb{P} [m \in [(\frac{3}{4} \pm \delta) \ell^-]] \leq \mathbb{P} [m \in [(\frac{1}{3} \pm 2\delta) n]]$  for  $n$  large. Similarly we have

$$\delta_s * \mu_{\text{LS}}^{*\ell} = \delta_s * \mu_{\text{LS}}^{*\frac{2}{3}n} * \mu_{\text{LS}}^{*\ell - \frac{2}{3}n} = \delta_s * \left( \frac{2}{3}\mu_S + \frac{1}{3}\delta_0 \right)^{*\frac{2}{3}n} * \mu_{\text{LS}}^{*\ell - \frac{2}{3}n} = \delta_s * \mu_S^{*m'} * \mu_{\text{LS}}^{*\ell - \frac{2}{3}n}$$

with  $1 - \varepsilon \leq \mathbb{P} [m' \in [(\frac{2}{3} \pm \delta) \frac{2}{3}n]] \leq \mathbb{P} [m' \in [(\frac{4}{9} \pm \delta) n]]$  for  $n$  large. Under these generic conditions on  $m$  and  $m'$ , Lemma 6.7 gives a probability no less than  $1 - \varepsilon$  that the two random walks with respective laws  $\mu_S^{*\ell^+}$  and  $\delta_s * \mu_S^{*m'}$  achieve a standard coupling (up to adding 1), provided  $\delta$  is small enough, and  $n$  large. Conditioning by such a success, Lemma 6.7, applied to random walks of respective laws  $(\mu_{\text{LS}})^{*m}$  and  $\mu_{\text{LS}}^{*\ell - \frac{2}{3}n}$  started at most 1 apart, gives a probability no less than  $1 - \varepsilon$  that the two considered random walks achieve coupling. We conclude that

$$\left\| \left( \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1} \right)^{*\ell^+} * \left( \frac{1}{2}\delta_0 + \frac{1}{4}\delta_1 + \frac{1}{4}\delta_{-1} \right)^{*\ell^-} - \delta_s * \left( \frac{1}{3}\delta_0 + \frac{1}{3}\delta_1 + \frac{1}{3}\delta_{-1} \right)^{*\ell} \right\|_1 \leq 4\varepsilon$$

provided  $\delta$  is small enough and  $n$  large. □

6.2. *Proof of Theorem 6.1.* Let us now describe precisely the labelling of the Cayley graphs of  $\mathbb{Z}$  and  $D_\infty$ , pictured in Figure 6.1. In the integers, from each vertex there is an edge to the right labelled by +1 and an edge to the left labelled by -1. In the dihedral group, let us say that the words  $(ab)^k$  correspond to even positions and the words  $(ab)^ka$  to odd positions. At each even position, there is an edge to the left labelled by  $a$  and an edge to the right labelled by  $b$ . At each odd position, this is the converse.

*Proof of the first part of Theorem 6.1:* For the simple random walk, the key observation is that position  $X_n$  and time  $n$  have the same parity. The path parametrized by  $X_n$  corresponds to a random walk on the integers where the increments at odd times are +1 for  $a$  and -1 for  $b$ , and the increments at even times are +1 for  $b$  and -1 for  $a$ . This provides a simple random walk on the integers, which is neither  $\ell^1$ -noise sensitive nor noise sensitive in average word distance by Proposition 5.2. □

For lazy random walks, time and position no longer have the same parity. Let us first observe the effect of refreshing one increment. Denote  $X_n = v_0sv_1$  and  $Y_n = v_0rv_1$  with  $s, r$  independent,  $\mu$ -distributed in  $\{a, b, e\}$ . By abuse of notation, we denote  $X_n = v_0 + s + v_1$  the corresponding path in the integers, as above.

When  $s \in \{a, b\}$  and  $r = e$  (or vice-versa), it corresponds in the integers to  $s = \pm 1$  (the sign depends on the parity of  $v_0$ ) being replaced by  $r = 0$  (or vice-versa). Thus  $v_0s$  and  $v_0r$  have different parity, so the next moves of the random walk, described by the word  $v_1$ , will be mirrored of the moves in the original walk. We write  $Y_n = v_0 + r - v_1$ .

When  $s = a$  and  $r = b$  (or vice-versa), then  $s = \pm 1$  (depending on  $v_0$ ) is replaced by  $r = \mp 1$ . The parity is not modified and we write  $Y_n = v_0 + r + v_1$ .



		$L_0$	$\alpha_i = -\alpha_{i-1}$				$\alpha_i = \alpha_{i-1}$			
	$s_i \in D_\infty$	$e$	$a$	$b$	$e$	$e$	$a$	$a$	$b$	$b$
	$r_i \in D_\infty$	$e$	$e$	$e$	$a$	$b$	$a$	$b$	$a$	$b$
even position	$s_i \in \mathbb{Z}$	0	1	-1	0	0	1	1	-1	-1
	$r_i \in \mathbb{Z}$	0	0	0	1	-1	1	-1	1	-1
odd position	$s_i \in \mathbb{Z}$	0	-1	1	0	0	-1	-1	1	1
	$r_i \in \mathbb{Z}$	0	0	0	-1	1	-1	1	-1	1

FIGURE 6.2. Correspondence of the effect of noising in the dihedral and integer model. The position is the integer  $v_0 + \dots + s_{j_{i-1}} + v_{i-1}$ , as in (6.1).

Now denote  $X_n = v_0 s_{j_1} v_1 s_{j_2} \dots s_{j_m} v_m$  a lazy simple random walk where  $s_{j_i}$  are the increments to be refreshed and  $v_i = s_{j_{i+1}} \dots s_{j_{i+1}-1}$  are words. With the notations above, it corresponds to a sum

$$X_n = v_0 + s_{j_1} + v_1 + s_{j_2} + v_2 \dots + s_{j_m} + v_m \tag{6.1}$$

in the integers. The refreshed sample  $Y_n^\rho = v_0 r_{j_1} v_1 \dots r_{j_m} v_m$  corresponds to a sum

$$Y_n^\rho = \alpha_0 v_0 + r_{j_1} + \alpha_1 v_1 + r_{j_2} + \alpha_2 v_2 \dots + r_{j_m} + \alpha_m v_m \tag{6.2}$$

where  $\alpha_i = \pm 1$  are given by  $\alpha_0 = 1$  and  $\alpha_i = -\alpha_{i-1}$  if and only if  $s_i \in \{a, b\}$  and  $r_i = e$  or vice-versa, i.e. there is a change of parity, which occurs with probability  $4/9$ . This is summarised in Table 6.2.

At this stage, we can get an intuition of the result because the sum  $\sum r_j$  is actually independent of the sum  $\sum s_j$  and the Lyapunov central limit theorem applied to the random variables  $\pm v_i$  ensures that the sums  $\sum v_i$  and  $\sum \alpha_i v_i$  are essentially independent. Making this precise would prove noise sensitivity at large scale.

However, to obtain  $\ell^1$  convergence, we need to construct a coupling between  $(X_n, Y_n^\rho)$  and  $(X_n, X'_n)$ . For this, we will use the additive model of expressions (6.1) and (6.2).

*Proof of the second part of Theorem 6.1:* Let us first explain how we sample  $(X_n, Y_n^\rho)$  in the integer model (6.1,6.2). We first sample the locations  $L$  of the refreshed increments, according to a Bernoulli law of parameter  $\rho$ . Among  $L$ , we sample the locations  $L_0$  where  $(s_i, r_i) = (e, e)$  (or  $(s_i, r_i) = (0, 0)$  in the integer model), according to a Bernoulli law of parameter  $\frac{1}{9}$ . Then we sample the  $v_i$ , i.e. the increments  $s_i$  for  $i$  not in  $L$  according to  $\frac{1}{3}(\delta_e + \delta_a + \delta_b)$ . Then we sample the values of  $\alpha_i$  according to a Bernoulli law of parameter  $\frac{1}{2}$ . This partitions  $L \setminus L_0$  into  $L_\alpha^+ \sqcup L_\alpha^-$ . Finally, we sample the pairs  $(s_i, r_i)$  according to  $\frac{1}{4}(\delta_{(1,-1)} + \delta_{(1,1)} + \delta_{(-1,1)} + \delta_{(-1,-1)})$  for  $i$  in  $L_\alpha^+$  and according to  $\frac{1}{4}(\delta_{(1,0)} + \delta_{(-1,0)} + \delta_{(0,1)} + \delta_{(0,-1)})$  for  $i$  in  $L_\alpha^-$ . In view of Table 6.2, this yields a sample of  $(X_n, Y_n^\rho)$  in the integer model.

Now we explain how to couple  $(X_n, Y_n^\rho)$  with  $(X_n, X'_n)$ . More precisely, we will couple the conditioned law  $(Y_n^\rho | X_n)$  with  $X'_n$  (which is equal to  $(X'_n | X_n)$  by independence) on a large enough subset.

We first sample the subsets  $L$  and  $L_0$ , thus we have integer expressions of  $X_n$  and  $Y_n^\rho$  as in (6.1,6.2). We will mimic these expressions for  $X'_n$  and write it in the form

$$X'_n = v'_0 + s'_{j_1} + v'_1 + s'_{j_2} + v'_2 \dots + s'_{j_m} + v'_m + \sum_{i \in L_0} s'_i$$

where  $v'_i = \sum_{i \in \{j_i+1, \dots, j_{i+1}-1\} \setminus L_0} s'_i$ . In other words, we pull out the terms in  $L_0$ . Then we will call length of  $v_i$ , and denote by  $\ell(v_i)$ , the cardinal of the set  $\{j_i+1, \dots, j_{i+1}-1\} \setminus L_0$ . This choice makes sure that  $\ell(v_i) = \ell(v'_i)$  for all  $i$ . This length follows a geometric law of parameter  $\frac{1-\frac{8}{9}\rho}{1-\frac{1}{9}\rho}$  (taking into

account the specificity of  $L_0$ ). It follows that for any integer  $k \geq 0$  there is a constant  $c_k > 0$  such that for any  $\delta > 0$

$$\mathbb{P} [\#\{i : \ell(v_i) = k\} \in [(c_k \pm \delta)n]] \xrightarrow{n \rightarrow \infty} 1. \tag{6.3}$$

In order to construct a coupling between  $(X_n, Y_n^\rho)$  and  $(X_n, X'_n)$ , we will take advantage of commutativity of the addition of integers in order to decompose the conditioned process  $(Y_n^\rho | X_n)$  and  $X'_n$  into a finite sequence of independent rescaled simple random walks, run during close times. For this, we fix an integer  $K$  to be determined later and define:

$$X_n^{\text{long}} = \sum_{\{i:\ell(v_i)>K\}} v_i \quad \text{and} \quad X_n^{\text{short}} = \sum_{k=0}^K \sum_{\{i:\ell(v_i)\leq K, |v_i|=k\}} v_i = \sum_{k=0}^K k S_{m(k)}^{(k)} \tag{6.4}$$

and similarly

$$Y_n^{\text{long}} = \sum_{\{i:\ell(v_i)>K\}} \alpha_i v_i \quad \text{and} \quad Y_n^{\text{short}} = \sum_{k=0}^K \sum_{\{i:\ell(v_i)\leq K, |v_i|=k\}} \alpha_i v_i = \sum_{k=0}^K k S_{m(k)}^{\prime\prime(k)} \tag{6.5}$$

where  $m(k)$  is the cardinal of the subset  $\{i : \ell(v_i) \leq K, |v_i| = k\}$ . The sum of the  $v_i$ 's on the latter subset has the law of a simple random walk, which we denote  $S_{m(k)}^{(k)}$ , rescaled by the multiplicative factor  $k$ . As the  $\alpha_i$ 's are independent uniform on  $\{\pm 1\}$ , the sum of the  $\alpha_i v_i$ 's indexed by this very subset is an independent simple random walk  $S_{m(k)}^{\prime\prime(k)}$ . We also denote

$$X_n^{\alpha+} = \sum_{j \in L_\alpha^+} s_j, \quad X_n^{\alpha-} = \sum_{j \in L_\alpha^-} s_j, \quad Y_n^{\alpha+} = \sum_{j \in L_\alpha^+} r_j, \quad Y_n^{\alpha-} = \sum_{j \in L_\alpha^-} r_j,$$

in order to have

$$X_n = X_n^{\text{long}} + X_n^{\text{short}} + X_n^{\alpha+} + X_n^{\alpha-} \quad \text{and} \quad Y_n^\rho = Y_n^{\text{long}} + Y_n^{\text{short}} + Y_n^{\alpha+} + Y_n^{\alpha-}.$$

Similarly we decompose  $X'_n = X_n^{\text{long}} + X_n^{\text{short}} + X_n^{\prime\alpha+} + X_n^{\prime\alpha-} + X_n^{\prime 0}$  into

$$X_n^{\prime\text{long}} = \sum_{\{i:\ell(v'_i)>K\}} v'_i, \quad X_n^{\prime\text{short}} = \sum_{k=0}^K \sum_{\{i:\ell(v'_i)\leq K, |v'_i|=k\}} v_i = \sum_{k=0}^K k S_{m'(k)}^{\prime(k)},$$

and

$$X_n^{\prime\alpha+} = \sum_{j \in L_\alpha^+} s'_j, \quad X_n^{\prime\alpha-} = \sum_{j \in L_\alpha^-} s'_j, \quad X_n^{\prime 0} = \sum_{j \in L_0} s'_j,$$

Using (6.3) and the distributions of the  $v_i$ 's conditioned by their lengths, we obtain for each  $k \geq 0$  a constant  $c'_k > 0$  such that for any  $\delta > 0$  and  $n$  large

$$\mathbb{P} [m(k) \in [(c'_k \pm \delta)n]] \xrightarrow{n \rightarrow \infty} 1 \tag{6.6}$$

and the same holds for  $m'(k)$ . Observe that  $c'_1 > c_1$ . Also note that by (6.3) and tail decay of the geometric law, there exists a sequence  $\varepsilon_K \xrightarrow{K \rightarrow \infty} 0$  such that

$$\mathbb{P} [\ell(X_n^{\text{long}}) \leq \varepsilon_K n] \xrightarrow{n \rightarrow \infty} 1 \tag{6.7}$$

where  $\ell(X_n^{\text{long}}) = \sum_{i:\ell(v_i)>K} \ell(v_i)$  is the length of  $X_n^{\text{long}}$  as a word (we removed the trivial increments in  $L_0$ ).

We are now ready to describe the required coupling. We fix some  $\varepsilon > 0$ . By Lemma 6.7, there exists  $x, \delta > 0$  such that any two simple random walks started distance  $\leq x\sqrt{n}$  apart and run for times in  $[(1 \pm \delta)n]$  achieve a standard coupling with probability at least  $1 - \varepsilon$ , provided  $n$  is large

enough. Moreover by the central limit theorem applied to a (lazy or not) simple random walk  $S_n$ , there exists  $A > 0$  such that for  $n$  large

$$\mathbb{P} [|S_n| \leq A\sqrt{n}] \geq 1 - \varepsilon.$$

We choose  $K$  large enough that  $\frac{4A\sqrt{\varepsilon K}}{c_1} \leq x$ . The central limit theorem and (6.7) show that

$$\mathbb{P} [|X_n^{\text{long}}| \leq A\sqrt{\varepsilon_K n}] \geq 1 - \varepsilon$$

and the same inequality holds with  $Y_n^{\text{long}}$  or  $X_n^{\text{long}}$  in place of  $X_n^{\text{long}}$ . Let  $B_0 = \{|Y_n^{\text{long}}| > A\sqrt{\varepsilon_K n}\} \cup \{|X_n^{\text{long}}| > A\sqrt{\varepsilon_K n}\}$  denote the “bad” set at step 0. We have  $\mathbb{P}(B_0) \leq 2\varepsilon$ .

We also choose  $\delta_1 > 0$  such that  $\frac{\delta_1}{c_1} \leq \delta$ . We set  $B_1 = \{m(1) \notin [(c'_1 \pm \delta_1)n]\} \cup \{m'(1) \notin [(c'_1 \pm \delta_1)n]\}$ . Applying (6.6) with  $k = 1$ , we have  $\mathbb{P}(B_1) \leq \frac{2\varepsilon}{K}$  provided  $n$  is large enough.

Conditioned on  $X_n$  and on  $B_0^c \cap B_1^c$ , the random variables  $Y_n^{\text{long}} + S_{m(1)}''(1)$  and  $X_n^{\text{long}} + S_{m'(1)}''(1)$  are simple random walks started distance  $|Y_n^{\text{long}} - X_n^{\text{long}}| \leq 2A\sqrt{\varepsilon_K}\sqrt{n}$  apart and run for times  $m(1), m'(1) \in [(c'_1 \pm \delta_1)n] \subset [(1 \pm \delta)c'_1 n]$  by choice of  $\delta_1$ . The choice of  $K$  permits to use Lemma 6.7 and we obtain for  $n$  large

$$\mathbb{P} [|Y_n^{\text{long}} + S_{m(1)}''(1) - (X_n^{\text{long}} + S_{m'(1)}''(1))| \leq 1 | B_0^c \cap B_1^c, X_n] \geq 1 - \varepsilon. \tag{6.8}$$

Now we will use Lemma 6.7 repeatedly as follows. For  $\varepsilon' = \frac{\varepsilon}{K} > 0$ , there exists  $\delta' > 0$  such that two simple random walks started distance at most 1 apart and run for times in  $[(1 \pm \delta')n]$  achieve a standard coupling with probability at least  $1 - \varepsilon'$ . For each  $2 \leq k \leq K$ , we choose  $\delta_k > 0$  with  $\frac{\delta_k}{c'_k} \leq \delta'$ . We denote  $B_k$  the “bad” event that  $\{m(k), m'(k)\}$  is not included in  $[(c'_k \pm \delta_k)n]$ . For  $n$  large enough, we have  $\mathbb{P}(B_k) \leq 2\varepsilon'$  by (6.6), so the “bad” event  $B = \cup_{k=0}^K B_k$  has probability  $\mathbb{P}(B) \leq 4\varepsilon$  and thus  $\mathbb{P}(B^c) \geq 1 - 4\varepsilon$ .

Let us denote  $G_k$  the “good” event at step  $k$ , still conditioned by  $X_n$ , that

$$G_k = \left\{ \left| Y_n^{\text{long}} + S_{m(1)}''(1) + \dots + kS_{m(k)}''(k) - \left( X_n^{\text{long}} + S_{m'(1)}''(1) + \dots + kS_{m'(k)}''(k) \right) \right| \leq 1 | X_n \right\}$$

Inequality (6.8) implies  $\mathbb{P}(G_1 | B^c, X_n) \geq 1 - \varepsilon$ . We also have:

$$\forall 2 \leq k \leq K, \quad \mathbb{P}(G_k | B^c, G_{k-1}, X_n) \geq 1 - \varepsilon'.$$

Indeed under the conditions  $X_n, G_{k-1}$  and  $B^c$ , the event  $G_k$  has the probability that two simple random walks started at most 1 apart and run for times  $m(k), m'(k) \in [(c'_k \pm \delta_k)n] \subset [(1 \pm \delta')c'_k n]$  achieve a standard coupling, which holds for  $n$  large with probability no less than  $1 - \varepsilon'$  by choice of  $\delta'$ . It follows that

$$\mathbb{P}(G_K | B^c, X_n) \geq 1 - \varepsilon - (K - 1)\varepsilon' \geq 1 - 2\varepsilon. \tag{6.9}$$

There remains to consider the contributions of the pairs  $(s_i, r_i)$  for  $i \in L \setminus L_0$  and of  $s'_i$  for  $i \in L$ . Note that as the values of  $\alpha_i$ 's have been sampled (which guaranteed the independence of  $S_{m(k)}^{(k)}$  and  $S_{m(k)}''(k)$ ), the locations  $L_\alpha^+$  and  $L_\alpha^-$  are determined already. For some  $\delta_\alpha > 0$ , let us denote

$$B_\alpha = \left\{ \#L_\alpha^+ \notin \left[ \left( \frac{4\rho}{9} \pm \delta_\alpha \right) n \right] \right\} \cup \left\{ \#L_\alpha^- \notin \left[ \left( \frac{4\rho}{9} \pm \delta_\alpha \right) n \right] \right\} \cup \left\{ \#L_0 \notin \left[ \left( \frac{\rho}{9} \pm \delta_\alpha \right) n \right] \right\}$$

the “bad” event that the partition  $L = L_\alpha^+ \sqcup L_\alpha^- \sqcup L_0$  is not “generic”. We have for any  $\delta > 0$  and  $n$  large that  $\mathbb{P}(B_\alpha^c) \geq 1 - \varepsilon$ .

According to Table 6.2, the laws of  $Y_n^{\alpha+}$  and  $Y_n^{\alpha-}$  conditioned by  $B_\alpha^c, X_n, L_\alpha^+, L_\alpha^-$  are  $(\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1})^{*\#L_\alpha^+}$  and  $(\frac{1}{2}\delta_0 + \frac{1}{4}\delta_1 + \frac{1}{4}\delta_{-1})^{*\#L_\alpha^-}$  respectively. The law of  $X_n^{\alpha+} + X_n^{\alpha-} + X_n^0$  is  $(\frac{1}{3}\delta_0 + \frac{1}{3}\delta_1 + \frac{1}{3}\delta_{-1})^{*\#L}$ .

Note that  $G_K = \left\{ \left| Y_n^{\text{long}} + Y_n^{\text{short}} - \left( X_n^{\text{long}} + X_n^{\text{short}} \right) \right| \leq 1 | X_n \right\}$ . We apply Lemma 6.8 with  $|\sigma| \leq 1$ , and obtain, provided  $\delta_\alpha$  is small enough and  $n$  large,

$$\mathbb{P} \left[ Y_n^\rho = X'_n \mid G_K, B_\alpha^{\mathbb{C}}, X_n \right] \geq 1 - \varepsilon.$$

Combined with (6.9), we get

$$\mathbb{P} \left[ Y_n^\rho = X'_n \mid B^{\mathbb{C}}, B_\alpha^{\mathbb{C}}, X_n \right] \geq \mathbb{P} \left[ Y_n^\rho = X'_n \mid G_K, B^{\mathbb{C}}, B_\alpha^{\mathbb{C}}, X_n \right] \mathbb{P} \left[ G_K \mid B^{\mathbb{C}}, B_\alpha^{\mathbb{C}}, X_n \right] \geq 1 - 3\varepsilon.$$

Integrating over the condition  $X_n$ , we get  $\mathbb{P} \left[ (X_n, Y_n^\rho) = (X_n, X'_n) \mid B^{\mathbb{C}}, B_\alpha^{\mathbb{C}} \right] \geq 1 - 3\varepsilon$ . Finally as  $\mathbb{P} \left( B^{\mathbb{C}} \right) \leq 4\varepsilon$  and  $\mathbb{P} \left( B_\alpha^{\mathbb{C}} \right) \leq \varepsilon$ , we obtain for  $n$  large that  $\mathbb{P} \left[ (X_n, Y_n^\rho) = (X_n, X'_n) \right] \geq 1 - 8\varepsilon$   $\square$

Observe that in the proof above the parity issue between  $X_n$  and  $Y_n^\rho$  can only be solved using the final “lazy”  $Y_n^{\alpha-}$  part, because  $X_n^{\text{long}}$ , respectively  $X_n^{\text{short}}$ ,  $X_n^{\alpha+}$ , has the same parity as  $Y_n^{\text{long}}$ , respectively  $Y_n^{\text{short}}$ ,  $Y_n^{\alpha+}$ .

### 7. Wreath products

Let  $G$  and  $\Lambda$  be two groups. Assume  $G$  acts transitively on a set  $S$ . The permutational wreath product of  $G$  and  $\Lambda$  over  $S$  is the semi-direct product

$$\Lambda \wr_S G := \left( \bigoplus_S \Lambda \right) \rtimes G.$$

Its elements are pairs  $(f, g)$  where  $f : S \rightarrow \Lambda$  is finitely supported and  $g$  belongs to  $G$ . The action of  $G$  on finitely supported functions is by translations  $gf(\cdot) = f(\cdot g)$ . Given an arbitrary point  $o$  in  $S$ , a natural generating set is the union of elements  $(\text{id}, g)$  for  $g$  in some generating set of  $G$  together with elements  $(\lambda\delta_o, \text{id})$  for  $\lambda$  in some generating set of  $\Lambda$  where  $\lambda\delta_o(x) = \lambda$  if  $x = o$  and  $\lambda\delta_o(x) = \text{id}$  otherwise.

When  $S = G$  is acted upon by the right regular representation, we recover the (usual) wreath product and we simply denote it by  $\Lambda \wr G$ .

**7.1. The lamplighter group.** The lamplighter group is the wreath product  $\mathbb{Z}/2\mathbb{Z} \wr \mathbb{Z}$ . Its elements are pairs  $(f, t)$  where  $t$  is an integer and  $f : \mathbb{Z} \rightarrow \mathbb{Z}/2\mathbb{Z}$  is a finitely supported function. As the action is by shift on  $\mathbb{Z}$ , the product is  $(f, t)(f', t') = (f(\cdot) + f'(\cdot + t), t + t')$ . One can think of  $t$  as a position of a lighter and  $f$  as a space of configurations of lamps on or off. The group is generated by  $(0, \pm 1)$ , which correspond to moves of the lighter, and  $(1\delta_0, 0)$ , where  $1\delta_0(x)$  takes value 1 for  $x = 0$  and value 0 otherwise, which corresponds to switching on or off the lamp at the lighter’s position. Let  $\mu_1$  be equidistributed on  $\{(0, 0), (1\delta_0, 0)\}$  and  $\mu_2$  be equidistributed on  $(0, \pm 1)$ , then the measure  $\mu := \mu_1 * \mu_2 * \mu_1$  is called the “switch-walk-switch” measure.

**Theorem 7.1.** *The lamplighter group with switch-walk-switch measure is entropy noise sensitive and partially noise sensitive in average word distance.*

This group is not  $\ell^1$ -noise sensitive by Corollary 5.4.

*Proof:* Let  $X_n = s_1 \dots s_n = (f_n, x_n)$  denote a sample. The projection  $x_n$  to the integers is a simple random walk. We denote

$$\text{Loc}(x, n) := \{0 \leq t \leq n : x_t = x\}$$

the local time at  $x$  (both the set and its cardinal are called local time by a slight abuse). Conditioned on the local time, the state of the lamp at  $x$  is given by:

$$f_n(x) = \prod_{t \in \text{Loc}(x,n)} \alpha_t, \quad \text{where } \alpha_t \text{ are i.i.d. uniform in } \mathbb{Z}/2\mathbb{Z}.$$

Let us denote  $R_n := \{x \in \mathbb{Z} : \text{Loc}(x, n) > 0\}$  the range of the random walk projected to the integers. The entropy of  $X_n$  is given by

$$H(X_n) = \mathbb{E}|R_n| + O(\log n) \tag{7.1}$$

because to describe  $X_n$  we need to provide the lamp configuration on the range and the position (which has only logarithmic entropy). Note that we use logarithm in base 2. It is well-known that  $\mathbb{E}|R_n| \asymp \sqrt{n}$ . Moreover, by gaussian decay, the expected size of range is homogeneous in the sense that for  $n$  large, one has  $\mathbb{E}\mathbb{1}_A|R_n|/\mathbb{E}|R_n| \leq f(\mathbb{P}(A))$  for some function  $f(\varepsilon) \xrightarrow{\varepsilon \rightarrow 0} 0$ .

Now consider a refreshed sample  $Y_n^\rho = r_1 \dots r_n = (g_n, y_n)$ . Again conditioning by the trajectory projected onto the integers, it appears that the conditional entropy satisfies

$$H(Y_n^\rho|X_n) \geq \mathbb{E}|R_n^{\text{ref}}(Y_n^\rho)|, \tag{7.2}$$

where  $R_n^{\text{ref}}(Y_n^\rho) := \{x \in R_n(Y_n^\rho) : \exists t \in \text{Loc}(x, n), r_t = s'_t \text{ was refreshed}\}$ , because for  $x$  in this set of refreshed lamps, the value  $g_n(x)$  is independent of the sample  $X_n$ . Moreover, for a given  $x$  in the range of  $Y_n^\rho$ , the probability that the lamp is not refreshed is precisely  $(1 - \rho)^{\text{Loc}(x,n)}$ , so our conditional entropy is related to the distribution of local time. The Ray-Knight theorem guarantees that

$$\mathbb{P}[\#\{x \in R_n : \text{Loc}(x, n) \leq \delta\sqrt{n}\} \leq \delta\sqrt{n}] \xrightarrow{\delta \rightarrow 0} 1.$$

So for any  $\varepsilon > 0$ , there exists an  $\delta > 0$  such that the event

$$A_\delta := \{\#\{x \in R_n(Y_n^\rho) : \text{Loc}(x, n) \geq \delta\sqrt{n}\} \geq (1 - \varepsilon)|R_n(Y_n^\rho)|\}$$

has probability  $\mathbb{P}(A_\delta) \geq 1 - \varepsilon$ . In this set each lamp is refreshed with probability at least  $1 - (1 - \rho)^{\delta\sqrt{n}} \geq 1 - \varepsilon$ , provided  $n$  is large enough, so:

$$\mathbb{P}[|R_n^{\text{ref}}(Y_n^\rho)| \geq (1 - 2\varepsilon)\#\{x \in R_n(Y_n^\rho) : \text{Loc}(x, n) \geq \delta\sqrt{n}\}] \xrightarrow{n \rightarrow \infty} 1.$$

It follows that conditioned on  $A_\delta$  we have for  $n$  large:

$$\mathbb{P}[|R_n^{\text{ref}}(Y_n^\rho)| \geq (1 - 3\varepsilon)|R_n(Y_n^\rho)||A_\delta] \geq 1 - \varepsilon.$$

Denote  $B$  the event that  $|R_n^{\text{ref}}(Y_n^\rho)| \geq (1 - 3\varepsilon)|R_n(Y_n^\rho)|$ , we finally have  $\mathbb{P}(B) \geq 1 - 2\varepsilon$  for  $n$  large. We can compute:

$$\mathbb{E}|R_n^{\text{ref}}| = \mathbb{E}\mathbb{1}_B|R_n^{\text{ref}}| + \mathbb{E}\mathbb{1}_{B^c}|R_n^{\text{ref}}| \geq (1 - 3\varepsilon)\mathbb{E}\mathbb{1}_B|R_n| \geq (1 - 3\varepsilon)(1 - f(2\varepsilon))\mathbb{E}|R_n|.$$

Together with (7.1) and (7.2), this implies entropy noise sensitivity.

The distance between two elements  $(f, x)$  and  $(g, y)$  is the minimal number of steps for the lighter to start from position  $x$ , switch all lamps at positions  $t$  with  $f(t) \neq g(t)$  and go to position  $y$ . In particular,  $\mathbb{E}d(X_n, Y_n^\rho) \geq \mathbb{E}|R_n^{\text{ref}}(Y_n^\rho)| - o(\sqrt{n}) \asymp \sqrt{n}$ . This implies partial noise sensitivity in average word distance.  $\square$

**7.2. A lower bound for permutational wreath products.** The ideas in the previous proof can be used in arbitrary permutational wreath products, but it is usually difficult to obtain informations about local times. We give a weaker statement which will be used in the next section.

Given an action of  $G$  on  $S$ , recall that the inverted orbit of a point  $x$  in  $S$  under a word  $w = s_1 \dots s_n$  is the set

$$\mathcal{O}(w) = \{x, xs_1^{-1}, xs_2^{-1}s_1^{-1}, \dots, xs_n^{-1} \dots s_1^{-1}\}.$$

A switch-walk measure on  $\Lambda \wr_S G$  is a measure of the form  $\mu_\Lambda * \mu_G$  where  $\mu_G$  is an arbitrary measure on  $G$  and  $\mu_\Lambda$  is an arbitrary measure on the copy of  $\Lambda$  sitting over a fixed point  $o$  in  $S$ , namely  $\Lambda = \{\lambda\delta_o : \lambda \in \Lambda\}$ .

**Lemma 7.2.** *Let  $\Lambda$  be a finite group. A switch-walk random walk on  $\Lambda \wr_S G$  satisfies*

$$H(Y_n^\rho | X_n) \geq \rho H(\mu_\Lambda) \mathbb{E} |\mathcal{O}(X_n)|.$$

By slight abuse of notation, we write on the righthand side  $\mathcal{O}(X_n)$  where we mean the inverted orbit of the inverse of a sample path (or word) of the random walk and not only its evaluation in the group. Heuristically this lemma is simply a lower bound on the expected number of lamps refreshed. In its own, it does not provide information on noise sensitivity because the righthand side depends on the noise parameter  $\rho$ .

*Proof:* Let  $X_n^{\text{word}} = \lambda_1 \delta_o s_1 \dots \lambda_n \delta_o s_n$  denote a sample path of the random walk with evaluation  $X_n = (f_n, g_n)$  in the group  $\Lambda \wr_S G$ . The lamp at  $x$  takes value

$$f_n(x) = \lambda_1 \delta_o(x) \lambda_2 \delta_o(x s_1) \dots \lambda_n \delta_o(x s_1 \dots s_n) = \prod_{t \in \text{Loc}(x, n)} \lambda_t,$$

where  $\text{Loc}(x, n) := \{0 \leq t \leq n : x.s_1 \dots s_{t-1} = o\}$  and  $\lambda_t$  are independent of law  $\mu_\Lambda$ . Note that points with positive local time are precisely points on the inverted orbit of  $X_n^{\text{word}}$ .

Now to sample  $Y_n^\rho$ , we first sample the locations of refreshed increments, then resample only the  $\mu_G$  factors of the increments. This gives an intermediate word  $Y_n^{\prime \text{word}}$  whose evaluation  $Y_n'$  in the group individually has the same law as  $X_n$ . Then we refresh the  $\mu_\Lambda$  factors of the increments. We have :

$$H(Y_n^\rho | X_n) \geq H(Y_n^\rho | X_n^{\text{word}}) \geq H(Y_n^\rho | X_n^{\text{word}}, Y_n^{\prime \text{word}}) = H(Y_n^\rho | Y_n^{\prime \text{word}}).$$

The last equality is due to the fact that all the factors in the word  $X_n^{\text{word}}$  agree with the factors of  $Y_n^{\prime \text{word}}$  except those that will be resampled according to  $\mu_\Lambda$  independently to obtain  $Y_n^\rho$ .

For  $x$  in  $S$ , the lamp at  $x$  is refreshed from  $Y_n^{\prime \text{word}}$  to  $Y_n^\rho$  with probability  $1 - (1 - \rho)^{\text{Loc}(x, n)}$ , which is  $\geq \rho$  as soon as  $x$  has positive local time, i.e. is in the inverted orbit of  $Y_n^{\prime \text{word}}$ . It follows that  $H(Y_n^\rho | Y_n^{\prime \text{word}}) \geq \rho H(\mu_\Lambda) \mathbb{E} |\mathcal{O}(Y_n^{\prime \text{word}})| = \rho H(\mu_\Lambda) \mathbb{E} |\mathcal{O}(X_n^{\text{word}})|$ .  $\square$

**7.3. Permutational wreath products of some groups acting on rooted trees.** Let  $T = T_d$  be a rooted tree of degree  $d$ . We view it as the graph with vertex set  $\sqcup_{\ell \geq 0} X^\ell$  where  $X = \{1, \dots, d\}$  and edges between any two vertices of the form  $v$  and  $vx$  for  $x \in X$ . The root is the empty sequence obtained for  $\ell = 0$ , and the set  $X^\ell$  is called the  $\ell$ th level. The tree boundary is the set  $\partial T = X^{\mathbb{Z}_+}$ . The group  $\text{Aut}(T)$  of (rooted) automorphisms of  $T$  is canonically isomorphic to its permutational wreath product with the symmetric group  $\mathcal{S}_X$  over  $X$

$$\text{Aut}(T) \simeq \text{Aut}(T) \wr_X \mathcal{S}_X. \tag{7.3}$$

This is obtained by identification between  $T$  and the subtrees  $xT$  rooted at vertices  $x$  in the first level. We identify an element and its image under this isomorphism to write  $g = \langle g|_1, \dots, g|_d \rangle \sigma$  where  $g|_x \in \text{Aut}(T)$  is called the section at  $x$  and  $\sigma \in \mathcal{S}_X$  is the root permutation. The action of  $g \in \text{Aut}(T)$  on a vertex of the form  $xu$  for  $x \in X$  is given by  $g.xu = \sigma.xg|_x u$ . The product rule is  $gg' = \langle g|_1 g'|_{\sigma^{-1}(1)}, \dots, g|_d g'|_{\sigma^{-1}(d)} \rangle \sigma \sigma'$ . This isomorphism can be iterated to the  $\ell$ th level as

$$\text{Aut}(T) \simeq \text{Aut}(T) \wr_X \mathcal{S}_X \wr_X \mathcal{S}_X \cdots \wr_X \mathcal{S}_X \hookrightarrow \text{Aut}(T) \wr_{X^\ell} \mathcal{S}_{X^\ell},$$

and in particular, the section of  $g$  at any vertex is defined by induction as  $g|_{vx} = (g|_v)|_x$ . We refer to Grigorchuk (2000) or Nekrashevych (2005) for details on the structure of  $\text{Aut}(T)$  and wreath product isomorphisms.

We consider a finitely generated subgroup  $G = G(A, H)$  of  $\text{Aut}(T)$  with a generating set of the form  $A \cup H$  where  $A$  and  $H$  are two subgroups such that



- all the elements of  $A$  are rooted, i.e. have all their sections of the first level trivial, and  $A$  acts transitively on the first level,
- all the elements of  $H$  have the form  $h = \langle h|_1, \dots, h|_d \rangle \sigma$  with sections  $h|_1 \in H$  and  $h|_x \in A$  for all  $x \in X \setminus \{1\}$  and permutation fixing 1, i.e.  $\sigma 1 = 1$ .
- let  $\mathbf{u}_H$  and  $\mathbf{u}_A$  denote respectively the uniform measure on  $H$  and  $A$ , we assume moreover that the first section map  $h \mapsto h|_1$  pushes  $\mathbf{u}_H$  forward to  $\mathbf{u}_H$  itself, and that the other section maps  $h \mapsto h|_x$  for  $x \in X \setminus \{1\}$  push  $\mathbf{u}_H$  forward to  $\mathbf{u}_A$  (this together with transitivity of the  $A$ -action on  $X$  implies in particular that the image of  $G$  under the section map  $g \mapsto g|_v$  is  $G$  itself, and this for any vertex  $v$ ).

Recall that the first Grigorchuk group is the group acting on a binary rooted tree generated by four elements recursively defined by  $a = \langle e, e \rangle (12)$ ,  $b = \langle c, a \rangle$ ,  $c = \langle d, a \rangle$  and  $d = \langle b, e \rangle$ . It satisfies this set of hypotheses by putting  $A = \{e, a\}$  and  $H = \{e, b, c, d\}$ . Other classical examples covered by our hypotheses are the various Grigorchuk groups [Grigorchuk \(1985, 1986\)](#), as well as generalisations in [Bartholdi and Šuník \(2001\)](#), and mother automata groups of degree 0 introduced in [Bartholdi et al. \(2010\)](#).

We now consider some specific permutational wreath products  $\Lambda \wr_S G$  of a group  $G = G(A, H)$  with a finite group  $\Lambda$  over the set  $S = G.1^\infty \subset \partial T$ , which is the orbit of the boundary point  $1^\infty$  under the action of  $G$ . This group is generated by  $A \cup H \cup \Lambda$  where we identify  $\Lambda$  with the subgroup  $\{\lambda \delta_{1^\infty} : \lambda \in \Lambda\}$  where  $\lambda \delta_{1^\infty}(x)$  takes value  $\lambda$  at  $x = 1^\infty$  and  $e$  elsewhere. Under identifications of the tree boundaries  $\partial T = \sqcup_{x \in X} \partial xT$ , the isomorphism (7.3) extends to the action on the boundary and we obtain an embedding

$$\Lambda \wr_S G(A, H) \hookrightarrow (\Lambda \wr_S G(A, H)) \wr_X A.$$

Notice that under this embedding  $\lambda \delta_{1^\infty} = \langle \lambda \delta_{1^\infty}, e, \dots, e \rangle$ . Moreover it is immediately checked that  $H$  and  $\Lambda$  commute. For such groups, we have:

**Theorem 7.3.** *Let  $G = G(A, H)$  be as above and let  $\Lambda$  be a finite group. The random walk  $(\Lambda \wr_S G, \mathbf{u}_{\Lambda H} * \mathbf{u}_A)$  is partially entropy noise sensitive.*

The measure  $\mathbf{u}_{\Lambda H} * \mathbf{u}_A = \mathbf{u}_\Lambda * \mathbf{u}_{HA}$  was called “switch and walk” in the statement of Theorem 1.3. This theorem applies for instance to extensions of the Grigorchuk groups introduced by Bartholdi and Erschler, and which can have prescribed growth, see [Bartholdi and Erschler \(2012\)](#); [Brieussel \(2014\)](#); [Bartholdi and Erschler \(2014\)](#). For simplicity we have decided to restrict ourselves to the case of tree of constant degree, but this hypothesis is unnecessary and the arguments of the proof below apply to rooted trees with bounded degree. Such generalisations provide interesting groups with prescribed entropy when  $\Lambda$  is finite and prescribed speed when  $\Lambda = \mathbb{Z}$ , see [Brieussel \(2013\)](#); [Amir and Virág \(2017\)](#).

*Proof:* For these groups, [Amir and Virág \(2017, Inequality \(9\) and Corollary 24\)](#) have shown that the entropy of the random walk satisfies

$$c\mathbb{E} |\mathcal{O}(X_n)| \leq H(X_n) \leq C\mathbb{E} |\mathcal{O}(X_n)| \tag{7.4}$$

where  $\mathcal{O}(X_n)$  is the inverted orbit of  $1^\infty$  and the constants  $c, C$  depend only on the degree of the tree.

We will also use [Brieussel \(2013, Proposition 3.8\)](#) that the size of the inverted orbit is the sum of that of the sections  $|\mathcal{O}(X_n)| = \sum_{x=1}^d |\mathcal{O}(X_n|_x)|$ . By induction, we deduce that for any level  $\ell$

$$|\mathcal{O}(X_n)| = \sum_{v \in X^\ell} |\mathcal{O}(X_n|_v)|. \tag{7.5}$$

Let us first understand a section  $X_n|_x$  of the first level. By [Brieussel \(2013, Lemma 4.1\)](#), for each  $1 \leq x \leq d$ , the random word on the section at  $x$  has the form  $X_n|_x = b_1^x a_1^x b_2^x \dots b_s^x a_s^x$  where the factors  $a_r^x$  are uniform in  $A$  and the factors  $b_r^x$  are uniform in  $\Lambda H$ . These factors are all independent



and their number  $s$  follows a binomial law  $\mathcal{B}(n, \frac{d-1}{d^2})$ . In particular, the  $\mathbf{u}_{\Lambda H} * \mathbf{u}_A$  random walk induces (slowed down)  $\mathbf{u}_{\Lambda H} * \mathbf{u}_A$  random walks on the sections.

More precisely, if we denote an increment of  $X_n$  by  $s_i = b_i a_i$  with  $a_i$  uniform in  $A$  and  $b_i = \langle \lambda_i \delta_{1^\infty} h_i, a_{i2}, \dots, a_{id_0} \rangle \pi_i$  uniform in  $\Lambda H$ , then each factor  $b_j^x$  of the section  $X_n|_x$  is a product of exactly  $k$  factors of the form  $\lambda_i \delta_{1^\infty} h_i$  with probability  $(\frac{1}{d})^k \frac{d-1}{d}$ . Similarly each factor  $a_j^x$  is a product of exactly  $k$  factors  $a_{ij}$  with probability  $(\frac{d-1}{d})^k \frac{1}{d}$ . This is so because for each  $i$  the independent factor  $a_i$  resamples which section will receive the increment in  $\Lambda H$  (the unique section at  $1.X_{i-1}^{-1}$ ), while the  $d - 1$  other sections will receive an increment in  $A$ . (Note that this is not true for the last factor of  $X_n|_x$  because of time truncation.)

We will now consider the refreshed sample  $Y_n^\rho$ . As in the proof of Lemma 7.2, we consider the word  $Y_n^{\text{word}}$  obtained by resampling the locations of the refreshments and the  $\mu_G = \mathbf{u}_H * \mathbf{u}_A$  part of the refreshed samples (but not yet the  $\mathbf{u}_\Lambda$  part). The conditioned entropy  $H(Y_n^\rho | Y_n^{\text{word}})$  is bounded below by the expected number of sites in  $S$  where the lamps will be refreshed. The key point is that the refreshing parameter  $\rho$  increases under taking sections.

Let us denote an increment of  $Y_n^{\text{word}}$  by  $s_i = b_i a_i$ , where the factor  $a_i$  is uniform in  $A$  and  $b_i = \langle \lambda_i \delta_{1^\infty} h_i, a_{i2}, \dots, a_{id_0} \rangle \pi_i$  is uniform in  $\Lambda H$ . By the above discussion, each factor  $\lambda_j^x$  of the section  $Y_n^{\text{word}}|_x$  is a product of exactly  $k$  factors of the form  $\lambda_i$  with probability  $(\frac{1}{d})^k \frac{d-1}{d}$ . Such a product is refreshed with probability  $1 - (1 - \rho)^k$  (always, except if no factor is refreshed). It follows that in the word  $Y_n^{\text{word}}|_x$ , the probability to refresh a  $\mathbf{u}_\Lambda$ -factor is

$$\rho_1 = \sum_{k=1}^{\infty} (1 - (1 - \rho)^k) \left(\frac{1}{d}\right)^k \frac{d-1}{d} = 1 - \frac{1 - \rho}{1 + \frac{\rho}{d-1}} > \rho.$$

By induction, we deduce that for any initial  $\rho > 0$ , there exists a level  $\ell$  at which the refreshing parameter of  $\mathbf{u}_\Lambda$ -factors is  $\rho_\ell \geq \frac{1}{2}$ . Let  $v$  belong to level  $\ell$ , the average number of points of  $\mathcal{O}(Y_n^{\text{word}}|_v) \subset \partial vT$  where the lamps are refreshed is, as in the proof of Lemma 7.2, at least  $\rho_\ell \mathbb{E}|\mathcal{O}(Y_n^{\text{word}}|_v)| = \rho_\ell \mathbb{E}|\mathcal{O}(X_n|_v)|$ .

Using the partition (7.5) and inequalities (7.4), we conclude that

$$H(Y_n^\rho | X_n) \geq H(Y_n^\rho | Y_n^{\text{word}}) \geq \sum_{v \in X^\ell} \rho_\ell H(\mathbf{u}_\Lambda) \mathbb{E}|\mathcal{O}(X_n|_v)| \geq \frac{1}{2} H(\mathbf{u}_\Lambda) \mathbb{E}|\mathcal{O}(X_n)| \geq \frac{c}{2} H(\mathbf{u}_\Lambda) H(X_n).$$

□

The choice of measure is heavily used in the proof to get similar random walks at the sections.

*Remark 7.4.* It would be interesting to have a similar result for the groups  $G(A, H)$  themselves rather than such permutational product extensions, for instance for the first Grigorchuk group. However it is difficult to describe precisely the effect of noise on the  $\mu_{HA}$  part. Indeed, not only does this resample the increments in one section but it also shuffles the increments between different sections of a given level. Heuristically, this means that the effect of noise in the sections is even stronger than simply “increasing the parameter”, as was used in the proof above. We should also mention that the precise asymptotic behaviour of the entropy of (even simple) random walks on the first Grigorchuk group is still unknown.

### 8. Perspectives and questions

As mentioned in the introduction, our observations on noise sensitivity lead us to believe that the only obstructions to noise sensitivity are homomorphisms onto  $\mathbb{Z}$  and non-Liouville property, whence Question 1.5 and Conjectures 1.6 and 1.7. We record here some questions and tasks for further study of noise sensitivity of groups.

1. Find more examples of  $\ell^1$ -noise sensitive groups.

- Clarify which virtually abelian (or more generally nilpotent) groups are  $\ell^1$ -noise sensitive, and for which measures.
  - Is the wreath product of a finite group with the dihedral group  $\ell^1$ -noise sensitive for some measure? If yes, it would provide an example with exponential growth.
  - Study noise sensitivity phenomena in other Liouville groups, such as degree 0 automata groups [Amir et al. \(2016\)](#), degree 1 mother automata groups [Amir et al. \(2013\)](#) or simple groups [Matte Bon \(2014\)](#); [Nekrashevych \(2018\)](#).
  - Find examples of *strongly*  $\ell^1$ -noise sensitive groups, i.e. noise sensitive with respect to any (finitely supported generating) probability measure. Possibly, this would be the case for any torsion or simple Liouville group. We expect this property to hold for the first Grigorchuk group.
2. Noise sensitivity could also be studied *quantitatively*.
- The choice of refreshing parameter  $\rho$  to be constant is arbitrary and it is natural to consider a parameter  $\rho(n)$  depending on the length  $n$ . There should be a threshold over which the noised random walk resembles an independent sample. Our definitions of noise sensitivity simply require that this threshold is tending to 0.  
A lower bound is given by entropy consideration: the entropy of the noise should be no less than that of the independent sample:  $\rho(n)nH(\mu) \geq H(\mu_n)$ . However this is not enough in general (e.g.  $\mathbb{Z}$  is not  $\ell^1$ -noise sensitive).
  - Does partial  $\ell^1$ -noise sensitivity ( $\exists c > 0, \forall \rho \in (0, 1), \liminf \|\pi_n^\rho - \mu_n^2\|_1 \leq 2 - c$ ) imply Liouville property?
3. About entropy noise sensitivity:
- it is likely that the proof of [Theorem 7.1](#) could be improved to show that when  $G$  is entropy noise sensitive, then  $G \wr \mathbb{Z}$  and  $G \wr \mathbb{Z}^2$  are entropy noise sensitive as well. It would be a consequence of [Conjecture 1.7](#). By [Proposition 4.3](#), this is no longer true for wreath products with  $\mathbb{Z}^d$  for  $d \geq 3$  as they are non-Liouville.
  - Is it true that partial entropy noise sensitivity is equivalent to entropy noise sensitivity?
4. About the relationship between notions of noise sensitivity:
- Is it true that  $\ell^1$ -noise sensitivity implies entropy noise sensitivity?
  - Is it true that large scale noise sensitivity implies noise sensitivity in average word distance?
  - Are there random walks on groups not satisfying the homogeneity assumptions of [Definitions 3.1](#) and [3.4](#)?
5. About average distance noise sensitivity.
- Abelian groups are not (even partially) noise sensitive in average distance for any measure. Are there other groups with this property ?
  - Is it possible that  $\liminf \frac{\mathbb{E}d_G(X_n, Y_n^\rho)}{\mathbb{E}d_G(X_n, X_n')} > 1$ ?

## Acknowledgements

We thank anonymous referees for useful comments. We also thank Gidi Amir and Ryokichi Tanaka for interesting discussions regarding noise sensitivity. We are especially grateful to Gidi Amir and an anonymous referee for pointing out an error in a previous version, and contribution to repair it.

## References

- Alėšin, S. V. Finite automata and the Burnside problem for periodic groups. *Mat. Zametki*, **11**, 319–328 (1972). [MR301107](#).

- Amir, G., Angel, O., Matte Bon, N., and Virág, B. The Liouville property for groups acting on rooted trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, **52** (4), 1763–1783 (2016). [MR3573294](#).
- Amir, G., Angel, O., and Virág, B. Amenability of linear-activity automaton groups. *J. Eur. Math. Soc. (JEMS)*, **15** (3), 705–730 (2013). [MR3085088](#).
- Amir, G. and Virág, B. Speed exponents of random walks on groups. *Int. Math. Res. Not. IMRN*, **2017** (9), 2567–2598 (2017). [MR3658209](#).
- Bartholdi, L. and Erschler, A. Growth of permutational extensions. *Invent. Math.*, **189** (2), 431–455 (2012). [MR2947548](#).
- Bartholdi, L. and Erschler, A. Groups of given intermediate word growth. *Ann. Inst. Fourier (Grenoble)*, **64** (5), 2003–2036 (2014). [MR3330929](#).
- Bartholdi, L., Kaimanovich, V. A., and Nekrashevych, V. V. On amenability of automata groups. *Duke Math. J.*, **154** (3), 575–598 (2010). [MR2730578](#).
- Bartholdi, L. and Šuník, Z. On the word and period growth of some groups of tree automorphisms. *Comm. Algebra*, **29** (11), 4923–4964 (2001). [MR1856923](#).
- Benjamini, I., Kalai, G., and Schramm, O. Noise sensitivity of Boolean functions and applications to percolation. *Inst. Hautes Études Sci. Publ. Math.*, **90**, 5–43 (2001) (1999). [MR1813223](#).
- Brieussel, J. Behaviors of entropy on finitely generated groups. *Ann. Probab.*, **41** (6), 4116–4161 (2013). [MR3161471](#).
- Brieussel, J. Growth behaviors in the range  $e^{r^\alpha}$ . *Afr. Mat.*, **25** (4), 1143–1163 (2014). [MR3277875](#).
- Brieussel, J. and Zheng, T. Speed of random walks, isoperimetry and compression of finitely generated groups. *Ann. of Math. (2)*, **193** (1), 1–105 (2021). [MR4199729](#).
- Davis, B. and McDonald, D. An elementary proof of the local central limit theorem. *J. Theoret. Probab.*, **8** (3), 693–701 (1995). [MR1340834](#).
- Frisch, J., Hartman, Y., Tamuz, O., and Vahidi Ferdowsi, P. Choquet-Deny groups and the infinite conjugacy class property. *Ann. of Math. (2)*, **190** (1), 307–320 (2019). [MR3990605](#).
- Furstenberg, H. Random walks and discrete subgroups of Lie groups. In *Advances in Probability and Related Topics, Vol. 1*, pp. 1–63. Dekker, New York (1971). [MR0284569](#).
- Garban, C., Pete, G., and Schramm, O. The Fourier spectrum of critical percolation. *Acta Math.*, **205** (1), 19–104 (2010). [MR2736153](#).
- Garban, C. and Steif, J. E. *Noise sensitivity of Boolean functions and percolation*, volume 5 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, New York (2015). ISBN 978-1-107-43255-0; 978-1-107-07643-3. [MR3468568](#).
- Grigorchuk, R. I. Degrees of growth of finitely generated groups and the theory of invariant means. *Math. USSR Izv.*, **25** (2), 259–300 (1985). DOI: [10.1070/IM1985v025n02ABEH001281](#).
- Grigorchuk, R. I. On the growth degrees of  $p$ -groups and torsion-free groups. *Math. USSR Sb.*, **54** (1), 185–205 (1986). DOI: [10.1070/SM1986v054n01ABEH002967](#).
- Grigorchuk, R. I. Just infinite branch groups. In *New horizons in pro- $p$  groups*, volume 184 of *Progr. Math.*, pp. 121–179. Birkhäuser Boston, Boston, MA (2000). [MR1765119](#).
- Kalai, G. Three puzzles on mathematics, computation, and games. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. I. Plenary lectures*, pp. 551–606. World Sci. Publ., Hackensack, NJ (2018). [MR3966739](#).
- Kaĭmanovich, V. A. and Vershik, A. M. Random walks on discrete groups: boundary and entropy. *Ann. Probab.*, **11** (3), 457–490 (1983). [MR704539](#).
- Lawler, G. F. and Limic, V. *Random walk: a modern introduction*, volume 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (2010). ISBN 978-0-521-51918-2. [MR2677157](#).
- Matte Bon, N. Subshifts with slow complexity and simple groups with the Liouville property. *Geom. Funct. Anal.*, **24** (5), 1637–1659 (2014). [MR3261637](#).
- Nekrashevych, V. *Self-similar groups*, volume 117 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (2005). ISBN 0-8218-3831-8. [MR2162164](#).

- Nekrashevych, V. Palindromic subshifts and simple periodic groups of intermediate growth. *Ann. of Math. (2)*, **187** (3), 667–719 (2018). [MR3779956](#).
- Révész, P. *Random walk in random and non-random environments*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, third edition (2013). [MR3060348](#).
- Saloff-Coste, L. Random walks on finite groups. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pp. 263–346. Springer, Berlin (2004). [MR2023654](#).
- Shannon, C. E. A mathematical theory of communication. *Bell System Tech. J.*, **27**, 379–423, 623–656 (1948). [MR26286](#).