

THE EXACT OVERLAPS CONJECTURE FOR SELF-SIMILAR MEASURES ON THE REAL LINE

SAMUEL KITTLE AND CONSTANTIN KOGLER

ABSTRACT. A generalization of the exact overlaps conjecture predicts a formula for the dimension of self-similar measures in terms of the random walk entropy and the Lyapunov exponent. We prove the exact overlaps conjecture and its generalized version in dimension one. The proof relies on a new entropy inequality quantifying the information lost under addition of independent random variables in terms of a quantity averaging variance across scales.

1. INTRODUCTION

Determining the dimension of self-similar measures is a fundamental problem in fractal geometry. In this paper, we address this question in dimension one. For $1 \leq i \leq k$ and $x \in \mathbb{R}$ let $g_i(x) = r_i x + b_i$ with $0 < |r_i| < 1$ and $b_i \in \mathbb{R}$ be finitely many distinct contracting similarities on $\mathbb{R}$. We consider a probability measure $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ supported on the given similarities with $p_i > 0$ for all $1 \leq i \leq k$. It is well-established [Hut81] that there exists a unique μ -stationary probability measure ν on $\mathbb{R}$, that is, a measure satisfying

$$\mu * \nu = \sum_{i=1}^k p_i (g_i)_* \nu = \nu.$$

The measure ν is called the self-similar measure of μ . In [FH09] it is shown that ν is exact dimensional, meaning that there exists $\alpha \in [0, 1]$ such that for ν -almost all $x \in \mathbb{R}$ it holds that $\nu([x - r, x + r]) = r^{\alpha + o(1)}$ as $r \rightarrow 0$. We refer to α as the dimension of ν and write $\dim \nu = \alpha$.

The fundamental quantities to calculate the dimension of ν are the Lyapunov exponent χ_μ and the random walk entropy h_μ given by

$$\chi_\mu = \sum_i p_i \log |r_i| \quad \text{and} \quad h_\mu = \lim_{n \rightarrow \infty} \frac{H(\mu^{*n})}{n},$$

where H denotes the Shannon entropy. It is well-known that $\dim \nu \leq \min\{1, \frac{h_\mu}{|\chi_\mu|}\}$ and conjectured for example in [Var, Conjecture 3] that equality always holds in the latter inequality, that is

$$\dim \nu = \min \left\{ 1, \frac{h_\mu}{|\chi_\mu|} \right\}. \tag{1.1}$$

The exact overlaps conjecture, which is mentioned for example in [Bak22], [Rap22], [FF25], and [Wu25], claims for $H(p) = H(\mu) = -\sum_{i=1}^k p_i \log p_i$ that $\dim \nu < \min\{1, \frac{H(p)}{|\chi_\mu|}\}$ only if μ has exact overlaps, meaning that $\text{supp}(\mu)$ does not freely generate a semigroup. As $h_\mu = H(p)$ if there are no exact overlaps, the conjecture

(1.1) is a generalization of the exact overlaps conjecture. A version of this conjecture for self-similar sets dates back to Simon from 1996 [Sim96].

A breakthrough on the exact overlaps conjecture was achieved by Hochman [Hoc14], establishing the conjecture under an exponential separation condition and thereby proving the conjecture for similarities with algebraic parameters. As [Hoc14] leaves the case of transcendental parameters open, there has been substantial effort to establish Hochman's results also for transcendental parameters. In [Hoc18], Hochman asked if failure of exponential separation necessarily implies exact overlaps, which was disproven by Baker [Bak21] and Bárány and Käenmäki [BK21]. Shmerkin [Shm19] extended Hochman's results to L^q dimensions and as a remarkable application proved Furstenberg's intersection conjecture.

An important achievement was Varjú [Var19b] proving the exact overlaps conjecture for transcendental Bernoulli convolutions, relying on delicate results in Diophantine approximation combined with work of Breuillard and Varjú [BV19], [BV20]. Rapaport [Rap22] settled systems with algebraic contraction ratios and arbitrary translations. Rapaport and Varjú [RV24] proved further cases for homogeneous systems of three maps. Indeed, for the maps $x \mapsto \lambda x$, $x \mapsto \lambda x + 1$ and $x \mapsto \lambda x + \tau$ with $\lambda \in (0, 1)$ and $\tau \in \mathbb{R}$ and μ the probability measure on these three maps with equal probability weights, Rapaport and Varjú [RV24] proved, for example, the exact overlaps conjecture if $(\lambda, \tau) \in (2^{-2/3}, 1) \times \mathbb{R}$. Moreover, [RV24] reduced the exact overlaps conjecture for this system of three maps to a criterion involving the Mahler measure of λ and showed that the set of exceptions for $(\lambda, \tau) \in (0, 1) \times \mathbb{R}$ has Hausdorff dimension 0. In the thesis of Streck [Str24], the exact overlaps conjecture was established for the equal weight measure of the maps $x \mapsto \lambda x$, $x \mapsto \lambda x + 1$, $x \mapsto \lambda x + \tau$ and $x \mapsto \lambda x + 1 + \tau$ with $\lambda \in (0, 1)$ and $\tau \in \mathbb{R}$ if $(\lambda, \tau) \in (2^{-3/2}, 1) \times \mathbb{R}$.

Feng and Feng [FF25] treated algebraic translations with an arbitrary nonzero real common contraction. Wu [Wu25] showed under a transversality condition that the set of exceptions for the exact overlaps conjecture is at most countable, without making the set of exceptions explicit. We also mention that the variance summation method, as initiated by the first-named author in [Kit25], was used by the authors [KK25] to generalize Hochman's theorem to contracting on average measures and to show that a weaker condition than exponential separation is sufficient.

In this paper we establish the exact overlaps conjecture in dimension one as well as the more general conjecture (1.1).

Theorem 1.1. *Let μ be a probability measure supported on finitely many contracting similarities of $\mathbb{R}$. Then the self-similar measure ν of μ satisfies*

$$\dim \nu = \min \left\{ 1, \frac{h_\mu}{|\chi_\mu|} \right\}.$$

A proof of Theorem 1.1 was independently announced in [Ope] as paper number 148. We will not discuss the mathematical content of the latter paper in detail, yet remark that Theorem 1.3 stated below is a quantitative version of Lemma 3.2 of [Ope, Paper 148] and therefore a strengthening. We highlight that Theorem 1.3 and Lemma 3.2 of [Ope, Paper 148] are, when applied to self-similar measures, the core novelty of the proofs. In this paper, we used ideas from [Kit25], [KK24] and [KK25] to conclude the proof of Theorem 1.1, whereas [Ope, Paper 148] uses entropy-growth arguments in the spirit of [Hoc14, Var19a].

We refer the reader to our machine contribution statement explaining how the authors found an initial proof of Theorem 1.1 and discussing the authors' contribution. We remark that [Ope] claim that their proof of Theorem 1.1 was found on September 24, 2026, while the authors first discovered a proof of Theorem 1.1 on September 27, 2026. [Ope] became public on October 6, 2026 and the authors made this paper available online on October 7, 2026.

To give a concrete consequence of Theorem 1.1, consider for $x \in \mathbb{R}$ the similarities

$$g_1(x) = \lambda_1 x + 1 \quad \text{and} \quad g_2(x) = \lambda_2 x - 1$$

with $\lambda_1, \lambda_2 \in (0, 1)$ and let $\mu = \frac{1}{2}\delta_{g_1} + \frac{1}{2}\delta_{g_2}$ be the unbiased probability measure supported on these two maps. Suppose that, after interchanging indices if necessary, λ_2 is transcendental over $\mathbb{Q}(\lambda_1)$. Then the maps have no exact overlaps and therefore it follows from Theorem 1.1 that

$$\dim \nu = \min \left\{ 1, \frac{2 \log 2}{\log \lambda_1^{-1} + \log \lambda_2^{-1}} \right\}.$$

Given contracting similarities $g_1, \dots, g_k$ on $\mathbb{R}$, there exists a unique non-empty compact set $K \subset \mathbb{R}$ satisfying $K = \bigcup_{i=1}^k g_i(K)$. The set K is called the self-similar set or the attractor of $g_1, \dots, g_k$. A significant motivation for the exact overlaps conjecture is, as is well-known, that it implies an explicit formula for the Hausdorff dimension of K when there are no exact overlaps. Under exponential separation, this dimension formula follows from [Hoc14, Corollary 1.2]. Theorem 1.1 implies the formula assuming only the absence of exact overlaps.

Corollary 1.2. *Let $g_1, \dots, g_k$ be contracting similarities on $\mathbb{R}$ with contraction rates r_i satisfying $0 < |r_i| < 1$ for $1 \leq i \leq k$. Let $s \geq 0$ be the unique non-negative solution of $\sum_{i=1}^k |r_i|^s = 1$. If $g_1, \dots, g_k$ have no exact overlaps, then the self-similar set K satisfies*

$$\dim K = \min\{1, s\}.$$

The proof of Theorem 1.1 presented in this paper relies on the entropy inequality Theorem 1.3, which is of independent interest, combined with [Hoc14, Theorem 1.4] and a simplified version of the variance summation method from [KK25].

As shown in [KK26S], the method of this paper extends to arbitrary dimensions provided that the rotations of the similarities in $\text{supp } \mu$ generate a virtually solvable group, which, in particular, includes the two dimensional case in full generality. By the Tits alternative [Tit72], the latter condition is equivalent to requiring that the entropy of the linear part of the product of the first n similarities is $o(n)$. However, the application of the entropy inequality Theorem 1.3 in Proposition 3.3 as discussed below fundamentally relies on addition being an abelian operation. Therefore, we do not regard this paper as having made significant progress on the general d -dimensional case.

A new entropy inequality. The main novelty in the proof of Theorem 1.1 is a general inequality estimating how much the entropy of a sum of random variables can differ from the entropy of the summands. We first state the new inequality and then discuss how it can be combined with previous ideas to deduce Theorem 1.1.

To state the entropy inequality, we first introduce a few definitions. For a compactly supported probability measure ρ on $\mathbb{R}$ and a scale $r > 0$ we denote by

$V_\rho(r)$ the normalized local variance defined as follows. Write $I(a, r) = [a, a + r)$ and for $r > 0$ we define

$$V_\rho(r) = \frac{4}{r^3} \int_{\mathbb{R}} F_\rho(I(a, r)) da, \quad (1.2)$$

where

$$F_\rho(I(a, r)) = \rho(I(a, r)) \operatorname{Var}(\rho \mid I(a, r)) = \inf_{c \in \mathbb{R}} \int_{I(a, r)} (x - c)^2 d\rho(x).$$

The factor $4r^{-3}$ consists of r^{-1} for averaging over the size of the interval, r^{-2} for normalizing the squared error and the constant 4 is chosen so that always $0 \leq V_\rho(r) \leq 1$. We furthermore define for $R > 0$

$$\mathcal{E}_R(\rho) = \int_0^R V_\rho(r) \frac{dr}{r} \quad \text{and} \quad \mathcal{E}(\rho) = \mathcal{E}_\infty(\rho). \quad (1.3)$$

We call $\mathcal{E}(\rho)$ the variance energy. The logarithmic measure dr/r assigns the same weight to every interval of scales $[r, 2r]$. We write $V_Y(r)$ for $V_\rho(r)$ and $\mathcal{E}(Y)$ for $\mathcal{E}(\rho)$ when the random variable Y has law ρ .

We now state the entropy inequality. Note that we call a random variable finitely supported if its law has finite support.

Theorem 1.3 (Entropy loss under addition). *For $m \geq 1$ and independent finitely supported real random variables $X_1, \dots, X_m$, it holds that*

$$\sum_{i=1}^m H(X_i) - H\left(\sum_{i=1}^m X_i\right) \leq 30m \sum_{i=1}^m \mathcal{E}(X_i). \quad (1.4)$$

To the authors' knowledge, Theorem 1.3 is novel. The inequality claims that addition of random variables can only conceal a lot of information if the summands have sufficient variance across scales. The key idea of the proof is to reveal more and more information of X_i by partitioning the space with a Poisson point process and to apply Margulis-Russo formulas on the derivative of Poisson processes. A sketch of the proof of Theorem 1.3 together with a discussion of the related literature is given at the beginning of Section 2.

We highlight that Theorem 1.3 is far more powerful than necessary to prove Theorem 1.1 as any bound of the form $\leq C_m \sum_{i=1}^m \mathcal{E}(X_i)$ for some finite constant C_m depending only on m would be sufficient. A bound of this kind with an unspecified constant C_m was established in Lemma 3.2 of [Ope, Paper 148] and we remark that Theorem 1.3 implies a strengthening of Lemma 3.2 of [Ope, Paper 148] with the explicit constant $C_m = 120m^2$. The proof method of Lemma 3.2 of [Ope, Paper 148] leads to a constant of the form $C_m = O(m^3)$ and, while Lemma 3.2 of [Ope, Paper 148] does not imply Theorem 1.3, the proof method of Lemma 3.2 of [Ope, Paper 148] can be adapted to prove Theorem 1.3 with the constant $C_m = O(m^3)$. In [KK26S] it is shown that a sharpening of the proof of Theorem 1.3 leads to the bound $\leq 62 \log m \sum_{i=1}^m \mathcal{E}(X_i)$ in (1.4) and that the logarithmic dependence on m cannot be improved.

We next discuss how to use Theorem 1.3 to prove Theorem 1.1.

Deduction of Theorem 1.1 from Theorem 1.3. For simplicity, we explain the argument for homogeneous self-similar measures. The general case can be deduced with analogous ideas by using appropriate conditioning on the contraction rate. So

consider for $1 \leq i \leq k$ and $x \in \mathbb{R}$ the maps $g_i(x) = \lambda x + b_i$ with $\lambda \in (0, 1)$ and $b_i \in \mathbb{R}$. A sample from the self-similar measure can by stationarity be written as

$$Z = \sum_{j=0}^{\infty} \lambda^j D_j = B_n + \lambda^n Z' \quad \text{with} \quad B_n = \sum_{j=0}^{n-1} \lambda^j D_j,$$

where the random variables D_j are independent and each takes the value b_i with probability p_i and Z' has law ν and is independent of B_n . Thus the measure decomposes into a finite block of independent digits and an independent rescaled copy of itself. Note that in the homogeneous case $\chi_\mu = \log \lambda$.

To prove Theorem 1.1 we argue by contradiction and assume that $\dim \nu < \min\{1, h_\mu / \log \lambda^{-1}\}$. We will show this leads to the contradiction that $\dim \nu = 1$. By the definition of h_μ , the Shannon entropy of B_n is $nh_\mu + o(n)$. On the other hand, since $\dim \nu < 1$ by [Hoc14, Theorem 1.4], at any fixed exponential scale $R = C^{-n}$ with $C > \lambda^{-1}$, it holds that

$$H(B_n; C^{-n}) = n \log \lambda^{-1} \dim \nu + o(n),$$

where $H(B_n; C^{-n})$ is the entropy at scale $R = C^{-n}$ from [Var19a] as recalled in Section 3. Thus it follows for $\delta = h_\mu - \log \lambda^{-1} \dim \nu > 0$ that we have the entropy gap

$$H(B_n) - H(B_n; C^{-n}) = \delta n + o(n). \quad (1.5)$$

In other words, even after identifying the interval of size $R = C^{-n}$ containing B_n , a linear amount of information remains unresolved inside that interval.

Theorem 1.3 converts this unresolved information into variance at scales below $R = C^{-n}$. To see how to apply Theorem 1.3, we introduce m independent copies of B_n that we denote by $B_n^{(1)}, \dots, B_n^{(m)}$. Their sum has the decomposition

$$\sum_{\ell=1}^m B_n^{(\ell)} = \sum_{j=0}^{n-1} \lambda^j \left(\sum_{\ell=1}^m D_j^{(\ell)} \right).$$

If each digit has at most q possible values, this gives

$$H\left(\sum_{\ell=1}^m B_n^{(\ell)}\right) \leq n(q-1) \log(m+1). \quad (1.6)$$

Thus the entropy of the sum, divided by the number of copies, becomes small compared with n when m is large.

Apply Theorem 1.3 after conditioning each copy on its interval of length $R = C^{-n}$. Now take a grid consisting of intervals of length $R = C^{-n}$ and fix the origin of the grid. Let Q_ℓ label the interval containing $B_n^{(\ell)}$, with Q denoting the label for B_n . Averaging the entropy inequality Theorem 1.3 over these labels therefore gives

$$mH(B_n | Q) - H\left(\sum_{\ell=1}^m B_n^{(\ell)} \middle| Q_1, \dots, Q_m\right) \leq 30m^2 \mathbb{E}_Q[\mathcal{E}(B_n | Q)]. \quad (1.7)$$

Moreover, each conditional law is supported on an interval of length $R = C^{-n}$, so a tail bound and concavity of $\mathcal{E}_R$ give

$$\mathbb{E}_Q[\mathcal{E}(B_n | Q)] \leq \mathbb{E}_Q[\mathcal{E}_{C^{-n}}(B_n | Q)] + \frac{1}{2} \leq \mathcal{E}_{C^{-n}}(B_n) + \frac{1}{2}.$$

Finally, averaging over the grid origin turns $H(B_n | Q)$ into $H(B_n) - H(B_n; C^{-n})$. Applying (1.7) and (1.6) and then dividing by m yields

$$H(B_n) - H(B_n; C^{-n}) \leq \frac{n(q-1)\log(m+1)}{m} + 30m\mathcal{E}_{C^{-n}}(B_n) + 15m.$$

Choose m sufficiently large and then let n tend to infinity. The entropy gap (1.5) therefore shows

$$\mathcal{E}_{C^{-n}}(B_n) \geq cn$$

for some $c > 0$ and sufficiently large n . Crucially, we gain a multiscale variance bound. At the beginning of Section 3, we explain how to use this bound to conclude the proof. Indeed, we combine $\mathcal{E}_{C^{-n}}(B_n) \geq cn$ with a simplified version of the variance summation method from [KK25] to deduce $\dim \nu = 1$, which is a contradiction.

Machine Contribution Statement. On September 27, 2026, GPT-6 Astra found the first proof of Theorem 1.1 presented to the second-named author (C.K.). Motivated by a suggestion of Elon Lindenstrauss, GPT-6 Astra was first asked to propose new approaches to simpler proofs of the main results of [Hoc14] and [Var19b], which was successfully achieved after around two hours. Encouraged by the model having apparently rethought the area in two hours, the exact overlaps conjecture was tried and the model succeeded after around two and a half hours. The initial proof included Theorem 1.3 and the deduction of Theorem 1.1 from Theorem 1.3 was essentially self-contained. The conversation transcript and the initial model drafts are available at [KK26S].

The second-named author (C.K.) immediately shared the discovery with the first-named author (S.K.), who found a new proof of the deduction of Theorem 1.1 from Theorem 1.3 using [Hoc14, Theorem 1.4] and the variance summation method from [Kit25] and [KK25] while the second-named author (C.K.) was rewriting the rest of the paper. The authors chose to write this new proof in the paper in order to make the paper more easily accessible to experts and to the authors themselves. Therefore, the introduction and Section 3 are due to the authors, while Section 2 is due to GPT-6 Astra and was rewritten by the authors. The connection of the proof of Theorem 1.3 to Margulis-Russo formulas as discussed in Section 2 was made by the authors. GPT-6 Astra also assisted with proofreading and stylistic suggestions.

The authors take full responsibility for the correctness of this paper and would be delighted to give talks on the results and to answer any questions.

Lean Formalization. GPT-6 Astra formalized all results from this paper and all necessary results from previous work including [Hoc14, Theorem 1.4]. The formalization is around 47,000 lines of code and took around 9.5 hours with four agents working in parallel. The Lean source code is available in the accompanying repository [KK26].

Supplementary Material. We provide supplementary material in [KK26S] containing the following:

- (1) An 8-page document with a self-contained proof of [Hoc14, Theorem 1.1].
- (2) A 10-page document with a self-contained proof of [Var19b, Theorem 3].
- (3) A 16-page document with a proof of Theorem 1.1 that includes Theorem 1.3 and concludes the proof of Theorem 1.1 using ideas similar to (1).

- (4) The conversation transcript leading to initial versions of (1), (2), and (3). Minor changes to (1), (2), and (3) were performed after the initial session on September 27, 2026.
- (5) A document generalizing Theorem 1.1 to arbitrary dimensions assuming that the rotations of μ generate a virtually solvable subgroup. This document was generated on the basis of a draft of this paper and contains no significant new ideas.
- (6) A document sharpening Theorem 1.3 to the bound $\leq 62 \log m \sum_{i=1}^m \mathcal{E}(X_i)$ and the logarithmic dependence on m cannot be improved. This document was generated on the basis of a draft of this paper and contains no significant new ideas.

Acknowledgments. We thank Emmanuel Breuillard, Michael Hochman and Péter Varjú for advice. The first-named author (S.K.) gratefully acknowledges support from the Heilbronn Institute for Mathematical Research. The second-named author (C.K.) holds a Postdoc Mobility Fellowship from the Swiss National Science Foundation (grant number 235409) and thanks the Institute for Advanced Study.

2. PROOF OF THEOREM 1.3: INFORMATION LOST UNDER ADDITION

In this section we prove Theorem 1.3 and first give an outline of the proof. Let $X_1, \dots, X_m$ be independent finitely supported real random variables and write $S = \sum_{i=1}^m X_i$. By independence,

$$\sum_{i=1}^m H(X_i) - H(S) = H(X_1, \dots, X_m \mid S).$$

Thus the entropy loss under addition is the uncertainty about the individual summands that remains after their sum is known. We estimate this uncertainty by revealing the summands gradually.

For each summand X_i , we use an independent family of refining interval partitions. Indeed, for each i place an independent Poisson point process on $\mathbb{R} \times (0, \infty)$ with intensity $da dt$, interpreting a point (a, t) as a cut at location a introduced at time t . For a random variable Y and P_t the partition formed by all cuts introduced by time t arising from an independent Poisson process we denote by $P_t(Y)$ the cell containing Y . Note that the partition given by P_t contains infinitely many points, yet they have no finite accumulation point by the Poisson property. At time $t > 0$, there is an average of t cuts per unit length and the gaps between consecutive cuts are exponentially distributed with mean $1/t$. The partitions are independent of the input variables. Expressions such as $\mathbb{E}_{P_t}[f(Y \mid P_t(Y))]$ average over both the partitions and the cell containing Y .

Using independent partitions $P_t = (P_t^{(1)}, \dots, P_t^{(m)})$, define

$$J(t) = \mathbb{E}_{P_t} \left[H(P_t^{(1)}(X_1), \dots, P_t^{(m)}(X_m) \mid S) \right].$$

Initially, the partitions reveal no information, so $J(0) = 0$. Eventually, they distinguish every point in each finite support. Consequently,

$$\lim_{t \rightarrow \infty} J(t) = H(X_1, \dots, X_m \mid S).$$

It therefore suffices to bound the derivative of J and integrate over time. Expressions for the derivative of point processes are well-known and referred to in the literature as

Margulis-Russo formulas or as perturbation formulas. Such differentiation formulas appeared first for Bernoulli product measures in the work of Esary and Proschan [EP63] and were later independently rediscovered by Margulis [Mar74] and Russo [Rus81] in the study of graph connectivity and percolation. The extension to Poisson processes was established by Zuyev [Zuy92] and we also refer to [LP17, Chapter 19] for a systematic exposition. We remark that the remainder of the proof appears to be novel, yet may be considered more natural from the viewpoint of Margulis-Russo formulas.

Indeed, as explained in the proof of Theorem 1.3, a direct consequence of [LP17, Theorem 19.1] is that

$$J'(t) = \mathbb{E}_{P_t} \left[\sum_{i=1}^m \sum_{C \in \mathcal{C}_t} \mathbb{P}[X \in C] \int_{C_i} H(\mathbf{1}_{\{X_i \leq a\}} \mid S, X \in C) da \right], \quad (2.1)$$

where $C = C_1 \times \cdots \times C_m$ ranges over the partition $\mathcal{C}_t = P_t^{(1)} \times \cdots \times P_t^{(m)}$ and $X = (X_1, \dots, X_m)$. Equation (2.1) motivates the conditional cumulative entropy defined for random variables X and Y as

$$\mathcal{A}(X \mid Y) = \int_{\mathbb{R}} H(\mathbf{1}_{\{X \leq a\}} \mid Y) da.$$

It measures the total uncertainty in threshold tests of X after Y is known and will be applied with the random variable S conditioned on C in (2.1). We will show in Lemma 2.2 for X and Y independent finitely supported real random variables, that

$$\mathcal{A}(X \mid X + Y) \leq 9(\mathcal{D}(X) + \mathcal{D}(Y)), \quad (2.2)$$

where the dispersion $\mathcal{D}(X)$ is defined as

$$\mathcal{D}(X) = \mathbb{E}[|X - X'|], \quad \text{for } X' \text{ is an independent copy of } X.$$

A calculation applying (2.2) to (2.1) leads to the inequality

$$J'(t) \leq 9m \sum_i \mathbb{E}_{P_t} [\mathcal{D}(X_i \mid P_t^{(i)}(X_i))]. \quad (2.3)$$

Finally, we establish in Lemma 2.4 for any random variable Y with independent Poisson point process P_t that

$$\int_0^\infty \mathbb{E}_{P_t} [\mathcal{D}(Y \mid P_t(Y))] dt \leq 3\mathcal{E}(Y). \quad (2.4)$$

Integrating (2.3) combined with the bound (2.4) shows

$$\sum_{i=1}^m H(X_i) - H(S) = \int_0^\infty J'(t) dt \leq 27m \sum_{i=1}^m \mathcal{E}(X_i),$$

which implies Theorem 1.3 as $27 \leq 30$.

We give the details after proving Lemmas 2.2 and 2.4. We begin with some basic properties of variance energy.

Lemma 2.1. *The functionals $F_\rho(I)$, $V_\rho(r)$ and $\mathcal{E}_R(\rho)$ are concave in ρ . Moreover, $0 \leq V_\rho(r) \leq 1$ and, for $a, b \in \mathbb{R}$ with $a \neq 0$,*

$$V_{aY+b}(r) = V_Y(r/|a|) \quad \text{and} \quad \mathcal{E}(aY + b) = \mathcal{E}(Y).$$

If ρ is finitely supported and its support has diameter at most D , then $\mathcal{E}(\rho) < \infty$ and

$$\mathcal{E}(\rho) - \mathcal{E}_R(\rho) \leq \frac{D^2}{2R^2}. \quad (2.5)$$

Proof. $F_\rho(I)$ is defined by an infimum over integrals over ρ and is therefore concave. The same follows therefore for $V_\rho(r)$ and $\mathcal{E}_R$. Taking the midpoint $c = a + r/2$ in the definition of $F_\rho(I(a, r))$ gives

$$0 \leq F_\rho(I(a, r)) \leq \frac{r^2}{4} \rho(I(a, r)).$$

Since $\int_{\mathbb{R}} \rho(I(a, r)) da = r$, the factor $4/r^3$ in (1.2) gives $0 \leq V_\rho(r) \leq 1$. Substitution gives the affine change of variables formula.

If the support of ρ has diameter at most D , then every conditional law $\rho | I$ also has support of diameter at most D , so $\text{Var}(\rho | I) \leq D^2/4$. Since $\int_{\mathbb{R}} \rho(I(a, r)) da = r$, we have $V_\rho(r) \leq D^2/r^2$. Hence

$$\mathcal{E}(\rho) - \mathcal{E}_R(\rho) \leq D^2 \int_R^\infty \frac{dr}{r^3} = \frac{D^2}{2R^2}.$$

For a finitely supported law, $V_\rho(r)$ also vanishes below the smallest spacing between atoms. Thus the full energy $\mathcal{E}(\rho)$ is finite. $\square$

Now we prove Lemma 2.2.

Lemma 2.2 (Conditional cumulative entropy). *If X, Y are independent finitely supported real random variables, then*

$$\mathcal{A}(X | X + Y) \leq 9(\mathcal{D}(X) + \mathcal{D}(Y)). \quad (2.6)$$

Proof of Lemma 2.2. We begin the proof with the following inequality. Suppose that a random variable Z satisfying $Z \in [0, 1]$ almost surely dominates a uniform variable. Write $T = \mathbb{E}[1 - Z]$ and we then claim that

$$\mathbb{E} \left[\sqrt{\frac{1-Z}{Z}} \right] \leq 3\sqrt{2T}. \quad (2.7)$$

We prove this by splitting into the two cases $\{Z \geq 1/2\}$ and $\{Z < 1/2\}$. On $\{Z \geq 1/2\}$, the integrand is at most $\sqrt{2(1-Z)}$, whose expectation is at most $\sqrt{2T}$ by Cauchy–Schwarz.

Now we deal with $\{Z < 1/2\}$. Write $q = \mathbb{P}[Z < 1/2]$ and note $q \leq 2T$. For $Q(u) = \inf\{z : \mathbb{P}[Z \leq z] \geq u\}$ with $u \in (0, 1)$ note that stochastic domination implies $Q(u) \geq u$. Moreover, $Q(U)$ has the same distribution as Z when U is uniform on $(0, 1)$. Therefore

$$\mathbb{E} \left[\mathbf{1}_{\{Z < 1/2\}} \sqrt{\frac{1-Z}{Z}} \right] \leq \mathbb{E}[Z^{-1/2} \mathbf{1}_{\{Z < 1/2\}}] \leq \int_0^q u^{-1/2} du = 2\sqrt{q} \leq 2\sqrt{2T}.$$

The first inequality is trivial and the second inequality holds since $Q(U)$ has the same distribution as Z and therefore $\mathbb{E}[Z^{-1/2} \mathbf{1}_{\{Z < 1/2\}}] = \int_0^1 Q(u)^{-1/2} \mathbf{1}_{\{Q(u) < 1/2\}} du$ and then apply $Q(u) \geq u$ and $\mathbf{1}_{\{Q(u) < 1/2\}} \leq \mathbf{1}_{\{u \leq q\}}$.

We now write $F_X(a) = \mathbb{P}[X \leq a]$ and $T_X(u) = \mathbb{P}[X' > X + u]$ for a fixed $u > 0$ and X' an independent copy of X . Note that $F_X(X)$ stochastically dominates a

uniform variable and therefore the same holds for the larger variable $F_X(X + u)$. Its expected complement is $\mathbb{P}(X' > X + u) = T_X(u)$, and thus by (2.7),

$$\mathbb{E} \left[\sqrt{\frac{1 - F_X(X + u)}{F_X(X + u)}} \right] \leq 3\sqrt{2T_X(u)}. \quad (2.8)$$

The reflected estimate follows from $1 - F_Y(Y - u) \geq \mathbb{P}(Y' \geq Y \mid Y)$ for $u > 0$; the latter variable is the distribution function of $-Y$ evaluated at $-Y$ and therefore also dominates uniform. The expected complement of $1 - F_Y(Y - u)$ is $\mathbb{P}(Y' \leq Y - u) = T_Y(u)$ outside the finite set of positive differences between atoms. Thus the analogous estimate holds for $F_Y(Y - u)/(1 - F_Y(Y - u))$ for almost every $u > 0$.

Let $p_s(a) := \mathbb{P}[X \leq a \mid X + Y = s]$ and let

$$q_s(a) := \frac{F_X(a)(1 - F_Y(s - a))}{F_X(a)(1 - F_Y(s - a)) + (1 - F_X(a))F_Y(s - a)}.$$

For almost every a we have

$$\begin{aligned} H(\mathbf{1}_{\{X \leq a\}} \mid S) &= \mathbb{E}[-\mathbf{1}_{\{X \leq a\}} \log p_S(a) - \mathbf{1}_{\{X > a\}} \log(1 - p_S(a))] \\ &= \mathbb{E}[-\mathbf{1}_{\{X \leq a\}} \log q_S(a) - \mathbf{1}_{\{X > a\}} \log(1 - q_S(a))] \\ &\quad - \mathbb{E}[\mathcal{D}_{KL}(p_S(a) \parallel q_S(a))] \\ &\leq \mathbb{E}[-\mathbf{1}_{\{X \leq a\}} \log q_S(a) - \mathbf{1}_{\{X > a\}} \log(1 - q_S(a))] \end{aligned}$$

where

$$\mathcal{D}_{KL}(p \parallel q) := p \log \frac{p}{q} + (1 - p) \log \frac{1 - p}{1 - q} \geq 0.$$

Write

$$L_a(X, S) := -\mathbf{1}_{\{X \leq a\}} \log q_S(a) - \mathbf{1}_{\{X > a\}} \log(1 - q_S(a)).$$

By the above and Tonelli's theorem we have

$$\mathcal{A}(X \mid X + Y) \leq \mathbb{E} \left[\int_{\mathbb{R}} L_{X+u}(X, X + Y) du \right].$$

We therefore wish to estimate the integrals

$$\int_{\mathbb{R}} L_a(x, x + y) da.$$

On $a > x$, write $a = x + u$ with $u > 0$. We have

$$L_{x+u}(x, x + y) = \log \left(1 + \frac{1 - F_X(x + u)}{F_X(x + u)} \frac{F_Y(y - u)}{1 - F_Y(y - u)} \right).$$

Using $\log(1 + z) \leq \sqrt{z}$, independence, and (2.8), its expectation $\mathbb{E}[L_{X+u}(X, X + Y)]$ is at most

$$18\sqrt{T_X(u)T_Y(u)} \leq 9(T_X(u) + T_Y(u)).$$

Similarly on $a < x$, write $a = x - u$. We have

$$L_{x-u}(x, x + y) = \log \left(1 + \frac{F_X(x - u)}{1 - F_X(x - u)} \frac{1 - F_Y(y + u)}{F_Y(y + u)} \right).$$

The reflected estimates give the same bound. Since

$$\int_0^\infty T_X(u) du = \frac{1}{2}\mathcal{D}(X),$$

integration over both halves proves (2.6). $\square$

Before proving Lemma 2.4 we need the following estimate for variance in terms of $\mathcal{D}$. Whilst it is straightforward to find a linear upper bound for the variance of the measure restricted to intervals of a given scale in terms of $\mathcal{D}$, it is necessary to look at many scales simultaneously in order to find a linear lower bound for the variance in terms of $\mathcal{D}$.

Lemma 2.3. *Suppose that Y is a random variable. Define*

$$v_Y(t) := \mathbb{E}[\text{Var}(Y|P_t(Y))].$$

Then

$$\mathcal{D}(Y) \leq 2 \int_0^\infty v_Y(t) dt.$$

Proof. Write $\rho = \mathcal{L}(Y)$ and let Y' be an independent copy of Y also independent of P_t . Then we have

$$\mathbb{E}[(Y - Y')^2 \mathbf{1}_{P_t(Y)=P_t(Y')} | P_t] = 2 \sum_{C \in P_t} \rho(C)^2 \text{Var}(\rho|C) \leq 2 \sum_{C \in P_t} \rho(C) \text{Var}(\rho|C).$$

Two points at distance d lie in the same cell of P_t with probability e^{-dt} . Therefore averaging over P_t we get

$$\mathbb{E}[(Y - Y')^2 e^{-t|Y-Y'|}] \leq 2v_Y(t).$$

Integrating over t and applying Tonelli's theorem gives the required inequality. $\square$

Next we prove Lemma 2.4.

Lemma 2.4 (Poisson dispersion). *For every finitely supported real random variable Y ,*

$$\begin{aligned} \int_0^\infty \mathbb{E}_P[\mathcal{D}(Y | P_t(Y))] dt &\leq 2 \int_0^\infty t \mathbb{E}_P[\text{Var}(Y | P_t(Y))] dt \\ &= 3 \int_0^\infty V_Y(r) \frac{dr}{r} = 3\mathcal{E}(Y). \end{aligned} \quad (2.9)$$

Proof. Let P'_t be an independent process of Poisson cuts. By Lemma 2.3 we have

$$\mathbb{E}_P[\mathcal{D}(Y|P_t(Y))] \leq 2 \int_{s=0}^\infty \mathbb{E}_{P,P'}[\text{Var}(Y|P_t(Y), P'_s(Y))] ds.$$

Since the refinement of two independent Poisson partitions with intensities t and s is a Poisson partition with intensity $t + s$ the right hand side is $2 \int_{s=0}^\infty v_Y(s + t) ds$. Therefore

$$\int_0^\infty \mathbb{E}_P[\mathcal{D}(Y | P_t(Y))] dt \leq 2 \int_0^\infty t \mathbb{E}_P[\text{Var}(Y | P_t(Y))] dt.$$

At intensity t , cells with left endpoint a and length r have intensity $t^2 e^{-tr} da dr$. This is because left endpoints have intensity $t da$, and the distance to the next cut is exponential of rate t . Tonelli's theorem and $\int_0^\infty t^3 e^{-tr} dt = 6r^{-4}$ therefore give

$$\begin{aligned} \int_0^\infty t \mathbb{E}_P \text{Var}(Y | P_t(Y)) dt &= \int_0^\infty \int_0^\infty \int_{\mathbb{R}} t^3 e^{-tr} F_\rho(I(a, r)) da dr dt \\ &= \frac{3}{2} \int_0^\infty V_\rho(r) \frac{dr}{r}. \end{aligned}$$

This proves (2.9). $\square$

We finish this section by proving Theorem 1.3.

Proof of Theorem 1.3. Use the notation X , S , P_t and $J(t)$ introduced at the beginning of this section. We have $J(0) = 0$. The partitions eventually distinguish all support points almost surely, so bounded convergence gives

$$\lim_{t \rightarrow \infty} J(t) = H(X | S) = \sum_{i=1}^m H(X_i) - H(S).$$

To justify (2.1), restrict cuts to a compact interval K containing all supports. This does not change the information carried by the cell labels. The restricted cuts form a Poisson process η_t on $E = \{1, \dots, m\} \times K$ with intensity $t\nu$, where $\nu = \sum_i \delta_i \otimes (da|_K)$ is finite. For a fixed configuration η , let $L_\eta(X)$ denote its tuple of cell labels and set $f(\eta) = H(L_\eta(X) | S)$. Then f is measurable, $0 \leq f \leq H(X)$, and $J(t) = \mathbb{E}[f(\eta_t)]$. Thus [LP17, Theorem 19.1], applied with base intensity $\lambda = 0$, perturbation measure ν and $n = 1$, gives continuous differentiability and

$$J'(t) = \int_E \mathbb{E}[f(\eta_t + \delta_z) - f(\eta_t)] \nu(dz).$$

At $t = 0$ the derivative is understood from the right. For $z = (i, a)$, adding the cut reveals the threshold test $\mathbf{1}_{\{X_i \leq a\}}$, so the entropy chain rule gives, for almost every a ,

$$f(\eta + \delta_{(i,a)}) - f(\eta) = H(\mathbf{1}_{\{X_i \leq a\}} | S, L_\eta(X)).$$

Expanding over the cell labels and integrating proves (2.1), since the threshold test is constant on $X \in C$ whenever $a \notin C_i$.

Fix a product cell $C = C_1 \times \dots \times C_m$ of positive probability. Conditional on $X \in C$, the coordinates remain independent. Applying Lemma 2.2 to X_i and $\sum_{j \neq i} X_j$ under this conditional law, and using the triangle inequality for independent copies, gives

$$\int_{C_i} H(\mathbf{1}_{\{X_i \leq a\}} | S, X \in C) da \leq 9 \sum_{j=1}^m \mathcal{D}(X_j | X_j \in C_j).$$

Substitute this into (2.1), sum over i and average over the cells. The marginal cell weights give

$$J'(t) \leq 9m \sum_{i=1}^m \mathbb{E}_{P_t} [\mathcal{D}(X_i | P_t^{(i)}(X_i))].$$

Integrating and applying Lemma 2.4, we conclude

$$\sum_{i=1}^m H(X_i) - H(S) = \int_0^\infty J'(t) dt \leq 27m \sum_{i=1}^m \mathcal{E}(X_i),$$

which proves the theorem since $27 \leq 30$. $\square$

3. PROOF OF THE MAIN THEOREM

Throughout this section, as in the introduction, let $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ be supported on distinct contracting similarities $g_i(x) = r_i x + b_i$ of $\mathbb{R}$, with $0 < |r_i| < 1$, $b_i \in \mathbb{R}$ and $p_i > 0$. For a similarity $g(x) = rx + b$, write $\rho(g) = r$ and $b(g) = b$ and denote $\rho_{\min} = \min_{1 \leq i \leq k} |r_i|$. Let $X_1, X_2, \dots$ be independent with law μ and write

$$G_n = X_1 \cdots X_n, \quad R_n = \rho(G_n), \quad B_n = b(G_n).$$

We use the same notation at stopping times for the natural filtration of (X_n) .

An important fact used throughout this section, as proved in [Kit25, Lemma 2.19], is that if τ is a bounded stopping time and Z' has law ν independently of the stopped walk, then

$$B_\tau + R_\tau Z' \quad \text{has law } \nu. \quad (3.1)$$

We now give an outline of the proof of Theorem 1.1 and consider for simplicity the homogeneous case $g_i(x) = \lambda x + b_i$ with $\lambda \in (0, 1)$. It is assumed for a contradiction that $\dim \nu < \min\{1, \frac{h_\mu}{|\chi_\mu|}\}$ and we aim to show that $\dim \nu = 1$. Recall that we discussed in the introduction how Theorem 1.3 together with [Hoc14, Theorem 1.4] implies that for $B_n = \sum_{\ell=0}^{n-1} \lambda^\ell D_\ell$ and $C > \lambda^{-1}$ that there is $c > 0$ such that for sufficiently large n

$$\mathcal{E}_{C^{-n}}(B_n) = \int_0^{C^{-n}} V_{B_n}(s) \frac{ds}{s} \geq cn.$$

Denote $d = (n+1) \log \lambda^{-1}$. For $u \in [0, d)$, consider the grid of scales e^{u+kd} with $k \in \mathbb{Z}$. Averaging over u gives

$$\frac{1}{d} \int_0^d \sum_{\substack{k \in \mathbb{Z} \\ e^{u+kd} < C^{-n}}} V_{B_n}(e^{u+kd}) du = \frac{1}{d} \int_0^{C^{-n}} V_{B_n}(s) \frac{ds}{s} \geq \frac{cn}{d}.$$

Thus some grid has total variance contribution at least cn/d . Since $V_{B_n}(s) = 0$ for sufficiently small s , retaining the nonzero contributions gives finitely many scales $s_1 < \dots < s_m < C^{-n}$ satisfying

$$s_{j+1} \geq \lambda^{-(n+1)} s_j \quad \text{and} \quad \sum_{j=1}^m V_{B_n}(s_j) \geq \frac{cn}{(n+1) \log \lambda^{-1}} \geq \frac{c}{2 \log \lambda^{-1}}.$$

Each $V_{B_n}(s_j)$ is the average conditional variance obtained by conditioning B_n on its cell in a randomly shifted grid of mesh s_j , multiplied by $4/s_j^2$.

To bring these contributions to a common scale $r \leq s_1$, choose integers $k_j \geq 0$ satisfying

$$\lambda \frac{r}{s_j} < \lambda^{k_j} \leq \frac{r}{s_j}.$$

The separation of the scales ensures that the digit blocks $[k_j, k_j + n)$ are disjoint. Thus the random-series representation of ν contains the independent summands

$$S = \sum_{j=1}^m \lambda^{k_j} B_n^{(j)},$$

where the $B_n^{(j)}$ are independent copies of B_n . S is therefore part of a sample of ν and the unused digits form an independent remainder.

Condition each $B_n^{(j)}$ on its cell in an independently shifted grid of mesh s_j , and let $\mathcal{F}$ record the shifts and cell labels. This fixes the coarse location of each block, leaving its fluctuation inside an interval of length s_j . The blocks remain conditionally independent. After centering and rescaling, the variables

$$X_j = \lambda^{k_j} (B_n^{(j)} - \mathbb{E}[B_n^{(j)} \mid \mathcal{F}])$$

satisfy $|X_j| \leq \lambda^{k_j} s_j \leq r$. Thus conditioning turns the local variance into the variance of small independent summands. Moreover,

$$\begin{aligned} \frac{4}{r^2} \mathbb{E}[\text{Var}(S \mid \mathcal{F})] &= \frac{4}{r^2} \sum_{j=1}^m \mathbb{E}[\text{Var}(X_j \mid \mathcal{F})] \\ &= \sum_{j=1}^m \frac{\lambda^{2k_j} s_j^2}{r^2} V_{B_n}(s_j) \geq \lambda^2 \sum_{j=1}^m V_{B_n}(s_j) \geq \frac{c\lambda^2}{2 \log \lambda^{-1}}. \end{aligned}$$

No mass is discarded as the original law is a mixture of these conditional laws and centering only translates each component.

For a fixed conditional component, denote $\sigma^2 = \sum_j \text{Var}(X_j \mid \mathcal{F})$. Since the X_j are centered, conditionally independent and satisfy $|X_j| \leq r$, a suitable version of the Berry–Esseen theorem approximates $\sigma^{-1} \sum_j X_j$ by a standard Gaussian with Wasserstein error bounded by

$$C \frac{\sum_j \mathbb{E}[|X_j|^3 \mid \mathcal{F}]}{\sigma^3} \leq C \frac{r}{\sigma},$$

for C an absolute constant. Thus a good approximation requires the total variance σ^2 to be large compared with r^2 . Taking $r \rightarrow 0$ alone does not guarantee this, since σ may decrease at the same rate.

The key point allowing us to repeat the construction is that $C > \lambda^{-1}$. Indeed, if $a > 0$ is the smallest scale already selected, we can choose a new block length N sufficiently large satisfying $C^{-N} < \lambda^{N+1} a$. Applying the energy estimate to B_N then supplies a new group of scales below C^{-N} , with the same uniform positive lower bound on their total variance contribution. This separates the new scales sufficiently from the previous ones that, upon placing all groups at a sufficiently small common scale r , their associated digit blocks remain disjoint. Iterating gives arbitrarily many independent groups, each contributing a fixed positive amount to the expected variance normalized by r^2 . Using the latter, we will conclude $\dim \nu = 1$, which is a contradiction and therefore shows that $\dim \nu = \min\{1, \frac{h_\mu}{|\chi_\mu|}\}$.

To make these arguments also work in the non-homogeneous case, we work with conditioning on the rate of contraction at step n . To do so, we use conditional laws. Indeed, for random variables Y and T , we write $\mathcal{L}(Y \mid T)$ for its conditional law, that is, the a.s.-unique $\sigma(T)$ -measurable random probability measure satisfying

$$\mathbb{E}[\mathbb{I}_E \mathcal{L}(Y \mid T)(A)] = \mathbb{P}(Y \in A, E)$$

for every Borel set $A \subset \mathbb{R}$ and every $E \in \sigma(T)$.

In Section 3.1, we use Theorem 1.3 and [Hoc14, Theorem 1.4] to deduce the suitably conditional version of $\mathcal{E}_{C^{-n}}(B_n) \geq cn$ as outlined in the introduction. We highlight that Section 3.1 is the core novelty of the proof in the deduction of Theorem 1.1 from Theorem 1.3. In Section 3.2 we introduce the quantity $W(\lambda; r)$ tracking the amount of variance at scale r the measure λ has and show that $W(\lambda; r) \rightarrow 0$ as $r \rightarrow 0$ implies that $\dim \lambda = 1$. The simplified version of the variance summation method of [KK25] is discussed in Section 3.3. Finally, we conclude the proof of Theorem 1.1 in Section 3.4.

3.1. Entropy and local variance. We use the same notion of entropy at a scale as Varjú [Var19a, Sections 1.2 and 2.2], which measures entropy by averaging over

translations. The concept originates in Wang [Wan11, Section 4.1]. For an integrable real-valued random variable Y , set

$$H(Y; r) = \frac{1}{r} \int_0^r H\left(\left\lfloor \frac{Y+t}{r} \right\rfloor\right) dt. \quad (3.2)$$

This is concave in the law of Y . As will be important later, if Y is finitely supported and $Z_t = \lfloor (Y+t)/r \rfloor$, the chain rule for entropy gives $H(Y) = H(Y, Z_t) = H(Z_t) + H(Y|Z_t)$ and therefore

$$H(Y) - H(Y; r) = \frac{1}{r} \int_0^r H(Y | Z_t) dt. \quad (3.3)$$

We also use Varjú's notion of entropy between two scales for some $C > 0$,

$$H(Y; r | Cr) := H(Y; r) - H(Y; Cr).$$

The following is stated in [Var19a].

Lemma 3.1. ([Var19a, Lemma 10]) *Suppose that $C > 1$ is an integer. Then $H(Y; r | Cr)$ is concave in the law of Y .*

We use the following estimate from Hochman.

Lemma 3.2 (Hochman's entropy estimate). *If $\dim \nu < 1$, then, for every $C > e^{\lfloor \chi_\mu \rfloor}$,*

$$H(B_n; C^{-n}) = n|\chi_\mu| \dim \nu + o(n).$$

Proof. Let $n' = n|\chi_\mu|/\log 2$. Hochman's theorem [Hoc14, Theorem 1.4 and the discussion following it], after converting to natural logarithms, gives

$$H(\lfloor 2^{qn'} B_n \rfloor, R_n) = n|\chi_\mu| \dim \nu + o(n) \quad (q > 1).$$

Integer parts of the partition levels change the entropy by $O(1)$. Clearly $H(R_n) \leq k \log(n+1) = o(n)$. We may therefore discard R_n . Any two translated grids of the same mesh have entropies differing by at most $\log 2$, since each cell of either grid meets at most two cells of the other. Averaging over the translation and taking $q = \log C/|\chi_\mu|$ proves the claim. $\square$

We next use Theorem 1.3 to prove a general inequality for finitely supported random variables.

Proposition 3.3. *Suppose $Y = \sum_{j=1}^L a_j D_j$, where the a_j are real constants and the D_j are independent real random variables, each taking at most K values. For every $R > 0$ and positive integer m ,*

$$\mathcal{E}_R(Y) \geq \frac{1}{30m} \left(H(Y) - H(Y; R) - \frac{L(K-1) \log(m+1)}{m} \right) - \frac{1}{2}. \quad (3.4)$$

Proof. Take independent copies of the digit vector and put $Y_i = \sum_{j=1}^L a_j D_j^{(i)}$ and $S = \sum_{i=1}^m Y_i$. For each j , the sum $\sum_i D_j^{(i)}$ is determined by the multiplicities of its possible values. The first $K-1$ multiplicities determine the last, so

$$H(S) \leq L(K-1) \log(m+1).$$

Fix $t \in [0, R)$ and write $Q_i = \lfloor (Y_i + t)/R \rfloor$ and $Q = \lfloor (Y + t)/R \rfloor$. Conditioned on $(Q_1, \dots, Q_m)$, the Y_i remain independent. Apply Theorem 1.3, average over the Q_i , and use $H(S | Q_1, \dots, Q_m) \leq H(S)$ to obtain

$$mH(Y | Q) \leq H(S) + 30m^2 \mathbb{E}[\mathcal{E}(\mathcal{L}(Y | Q))].$$

Each conditional law is supported on an interval of length R . The tail bound and concavity from Lemma 2.1 with $D = R$ therefore give

$$\mathbb{E}[\mathcal{E}(\mathcal{L}(Y | Q))] \leq \mathbb{E}[\mathcal{E}_R(\mathcal{L}(Y | Q))] + \frac{1}{2} \leq \mathcal{E}_R(Y) + \frac{1}{2}.$$

Average over t , use (3.3), and divide by m . Rearranging gives (3.4). $\square$

We now convert Lemma 3.2 and Proposition 3.3 to an estimate of the conditional variance energy after conditioning on R_n .

Corollary 3.4. *Suppose that $\dim \nu < \min\{1, h_\mu/|\chi_\mu|\}$. Then there is $\eta > 0$, depending only on μ , such that, for every $C > e^{|\chi_\mu|}$ and all sufficiently large (in terms of C) n ,*

$$\mathbb{E}[\mathcal{E}_{C^{-n}}(\mathcal{L}(B_n | R_n))] \geq \eta n. \quad (3.5)$$

Proof. Put $\delta = h_\mu - |\chi_\mu| \dim \nu > 0$. Choose a block length N such that $k \log(N+1)/N < \delta/4$. To reveal only little entropy, split the first n increments into $L = \lceil n/N \rceil$ consecutive blocks of length N , except possibly the last, and write their products as $x \mapsto A_j x + D_j$. Set $T = (A_1, \dots, A_L)$. Each A_j takes at most $(N+1)^k$ values, and R_n is determined by T . Hence

$$\begin{aligned} H(B_n | T) &= H(G_n | T) \geq H(G_n) - H(T) \\ &\geq nh_\mu - Lk \log(N+1) \geq n(h_\mu - \delta/4) - O(1), \end{aligned}$$

where $H(G_n) \geq nh_\mu$ follows from subadditivity of random-walk entropy. Concavity of entropy at a scale and Lemma 3.2 give

$$\mathbb{E}[H(\mathcal{L}(B_n | T); C^{-n})] \leq H(B_n; C^{-n}) = n|\chi_\mu| \dim \nu + o(n).$$

Thus the average conditional entropy gap is at least $\delta n/2$ for large n .

Given T , the D_j remain independent and

$$B_n = D_1 + A_1 D_2 + \dots + (A_1 \dots A_{L-1}) D_L$$

has fixed coefficients. Each digit has at most $K = k^N$ values. Choose m so large that $(K-1) \log(m+1)/m < \delta N/8$. Apply Proposition 3.3 conditionally and average. For large n , this gives

$$\mathbb{E}[\mathcal{E}_{C^{-n}}(\mathcal{L}(B_n | T))] \geq \frac{\delta n}{120m} - \frac{1}{2}.$$

Finally, R_n is a function of T , so concavity of $\mathcal{E}_{C^{-n}}$ allows us to replace T by R_n . Decreasing the constant proves (3.5), for example with $\eta = \delta/(240m)$. $\square$

3.2. Convolution disintegrations. We introduce the following functional to record how much variance can be placed in independent summands at a common scale.

Definition 3.5. *For a compactly supported probability measure λ and $r > 0$, a scale- r convolution disintegration is a representation*

$$\lambda = \int_I \lambda_i^{(1)} * \dots * \lambda_i^{(m_i)} d\theta(i),$$

where (I, θ) is a probability space, m_i is a finite positive integer, and each factor is supported on an interval of length at most r . All families are required to be

measurable. Define

$$W(\lambda; r) = \inf \int_I \exp \left(-\frac{4}{r^2} \sum_{j=1}^{m_i} \text{Var}(\lambda_i^{(j)}) \right) d\theta(i), \quad (3.6)$$

with the infimum over these disintegrations.

We use W as it has many suitable properties, as shown in Lemma 3.7 and in particular it is well behaved under convolution. We now provide some further motivation for this definition. At every sufficiently small scale, we aim to decompose the self-similar measure into components that are approximately normal, together with a small remainder. The standard deviations of these normal distributions should be comparable to the given scale. Standard Berry–Esseen estimates (Lemma 3.8) give the required normal approximations. Lemma 3.9 then provides entropy bounds between two scales for measures close to normal distributions. By concavity of entropy between two scales, these bounds imply that the self-similar measure is full dimensional.

We use $W(\cdot; \cdot)$ to keep track of this approximation. Suppose that $W(\nu; r)$ is sufficiently small. A simple application of Markov’s inequality then gives the desired decomposition after discarding an arbitrarily small fraction of the mass of ν . Each remaining component is the law of a sum of many independent random variables. These laws approximate normal distributions with variance Cr^2 arbitrarily well. More precisely, we prove the following proposition.

Proposition 3.6. *Suppose that ν is a self-similar measure on $\mathbb{R}$ and that there is a sequence $r_n \rightarrow 0$ such that $W(\nu; r_n) \rightarrow 0$. Then $\dim \nu = 1$.*

In [KK25] we do not use $W(\cdot; r)$ but instead use effective Cramér-type bounds to achieve the same goal of showing that all but an arbitrarily small portion of the measure can be approximated by the law of the sum of many independent random variables. This has the advantage of being more efficient but leads to a more complicated argument. However, both approaches are closely related.

The remaining aim of this section is to prove Proposition 3.6, which we establish by showing that $H(\nu; r) = \log r^{-1} + o(\log r^{-1})$ as $r \rightarrow 0$, which implies $\dim \nu = 1$ by exact dimensionality. We first prove some basic properties of $W(\cdot; r)$. We will then state and prove the lemmas necessary to prove Proposition 3.6 and prove Proposition 3.6.

Lemma 3.7. *The functional W takes values in $[0, 1]$ and is convex in its measure argument. Moreover,*

$$W(aY + b; r) = W(Y; r/|a|) \quad (a \neq 0), \quad (3.7)$$

$$W(\lambda * \theta; r) \leq W(\lambda; r)W(\theta; r), \quad (3.8)$$

$$W(\lambda; s) \leq W(\lambda; r)^{r^2/s^2} \quad (s \geq r), \quad (3.9)$$

$$W(\lambda; r) \leq 1 - (1 - e^{-1})V_\lambda(r). \quad (3.10)$$

This lemma is important as we will use (3.10) to estimate W of the translation part of the product of n similarities after conditioning on the linear part in terms of the variance energy. We then use the other properties to put together estimates for many such products into an improved estimate.

Proof of Lemma 3.7. Clearly $W \leq 1$. Convexity follows by combining disintegrations, affine changes of variables give (3.7), and concatenating two independent disintegrations gives (3.8). A scale- r convolution disintegration is also a scale s convolution disintegration for $s \geq r$. From the definition it is immediate that if its integrand at scale r is $Z \in [0, 1]$, its integrand at scale s is Z^{r^2/s^2} . Jensen's inequality followed by the infimum gives (3.9).

For (3.10), condition λ on a shifted grid of mesh r , taking one factor in each component, and average over a uniform shift. If Q denotes the resulting cell label, every normalized component variance $u = 4 \text{Var}(\lambda | Q)/r^2$ lies in $[0, 1]$. The inequality $e^{-u} \leq 1 - (1 - e^{-1})u$ and the definition of $V_\lambda(r)$ give the result. $\square$

We need the following Lemma for the proof of Proposition 3.6.

Lemma 3.8. *There is an absolute constant $C > 0$ such that the following holds. Suppose that $Y_1, \dots, Y_n$ are independent real-valued random variables with finite third moments and mean zero. Let*

$$\omega_i^2 = \mathbb{E}[Y_i^2], \quad \gamma_i^3 = \mathbb{E}[|Y_i|^3], \quad r^2 = \sum_{i=1}^n \omega_i^2 > 0.$$

Then

$$\mathcal{W}_1 \left(\sum_{i=1}^n Y_i, \mathcal{N}(0, r^2) \right) \leq C \frac{\sum_{i=1}^n \gamma_i^3}{r^2}.$$

In particular, if $|Y_i| \leq \varepsilon$ almost surely for every i , then

$$\mathcal{W}_1 \left(\sum_{i=1}^n Y_i, \mathcal{N}(0, r^2) \right) \leq C\varepsilon.$$

Proof. Many estimates of this form are available in the literature and we cite [Gol10, Theorem 1.1] as one convenient version. In one dimension, the Wasserstein distance is the L^1 distance from [Gol10, Theorem 1.1]. So [Gol10, Theorem 1.1], after rescaling by r , gives the first assertion with $C = 1$. For the second assertion simply note that $\gamma_i^3 \leq \varepsilon \omega_i^2$ and sum over i . $\square$

The following lemma allows us to estimate entropy at a given scale from a normal approximation in Wasserstein distance. We will apply it at two scales to estimate the entropy between them.

Lemma 3.9. *For every $0 < a < b < \infty$ and $\varepsilon > 0$, there is $\delta > 0$ such that the following holds. If $r > 0$, $\sigma \in [ar, br]$, and Y is a mean-zero random variable satisfying*

$$\mathbb{E}[Y^2] \leq b^2 r^2, \quad \mathcal{W}_1(Y, \mathcal{N}(0, \sigma^2)) < \delta r,$$

then

$$|H(Y; r) - H(\mathcal{N}(0, \sigma^2); r)| < \varepsilon.$$

Proof. By scaling, it suffices to take $r = 1$. If the assertion failed, there would be mean-zero random variables Y_n and numbers $\sigma_n \in [a, b]$ such that

$$\mathbb{E}[Y_n^2] \leq b^2, \quad \mathcal{W}_1(Y_n, \mathcal{N}(0, \sigma_n^2)) \rightarrow 0,$$

but the asserted entropy convergence fails. Passing to a subsequence, we may assume that $\sigma_n \rightarrow \sigma \in [a, b]$. It follows that Y_n converges in distribution to $\mathcal{N}(0, \sigma^2)$.

For $k \in \mathbb{Z}$ and $t \in [0, 1]$, write

$$A_{k,t}^{(n)} = \mathbb{P}[Y_n \in [k-t, k+1-t)], \quad \text{and} \quad B_{k,t} = \mathbb{P}[\mathcal{N}(0, \sigma^2) \in [k-t, k+1-t)].$$

Then $A_{k,t}^{(n)} \rightarrow B_{k,t}$ for every k and almost every t . Moreover, Chebyshev's inequality gives, uniformly in n and t , $A_{k,t}^{(n)} \ll_b (1 + |k|)^{-2}$. Consequently the terms $-A_{k,t}^{(n)} \log A_{k,t}^{(n)}$ admit a summable dominating function. Dominated convergence gives $H(Y_n; 1) \rightarrow H(\mathcal{N}(0, \sigma^2); 1)$ as $n \rightarrow \infty$, which is a contradiction. $\square$

We need the following to estimate the entropy between two scales.

Lemma 3.10. *For every integer $C > 1$ and $\varepsilon > 0$, there are $\delta > 0$ and $A > 0$ such that the following holds. If $r > 0$ and $Y_1, \dots, Y_n$ are independent mean-zero real-valued random variables satisfying*

$$|Y_i| \leq \delta r \quad \text{almost surely,} \quad \sum_{i=1}^n \text{Var}(Y_i) \geq Ar^2,$$

then

$$H\left(\sum_{i=1}^n Y_i; r \mid Cr\right) > \log C - \varepsilon.$$

Proof. By scale invariance and the smoothness of the Gaussian density, we may choose A sufficiently large such that whenever $Ar^2 \leq \sigma^2 \leq (A+1)r^2$ it holds that

$$H(\mathcal{N}(0, \sigma^2); r \mid Cr) > \log C - \frac{\varepsilon}{2}.$$

Choose $\delta < 1$ sufficiently small for Lemmas 3.8 and 3.9, applied at both scales r and Cr , to give a total entropy error of less than $\varepsilon/2$.

Take the shortest initial subsum $S = Y_1 + \dots + Y_\ell$ whose variance is at least Ar^2 . Writing $\sigma^2 = \text{Var}(S)$, minimality and $\text{Var}(Y_i) \leq \delta^2 r^2$ give

$$Ar^2 \leq \sigma^2 \leq (A + \delta^2)r^2 \leq (A+1)r^2.$$

Lemma 3.8 gives

$$\mathcal{W}_1(S, \mathcal{N}(0, \sigma^2)) \ll \delta r.$$

It follows from Lemma 3.9, used at the two scales, that

$$H(S; r \mid Cr) > \log C - \varepsilon.$$

Finally, adding the remaining independent summands cannot decrease entropy between these scales by, for example, Lemma 3.1 and the translation invariance of entropy between two scales. $\square$

We now prove Proposition 3.6.

Proof of Proposition 3.6. We first show $W(\nu; s) \rightarrow 0$ as $s \rightarrow 0$. For $0 < s \leq t$, let $\sigma = \min\{n \geq 0 : |R_n| \leq s/t\}$. This stopping time is bounded, and for $\rho_{\min} = \min_{1 \leq i \leq k} |r_i|$ we have

$$\rho_{\min} s/t < |R_\sigma| \leq s/t.$$

By (3.1), convexity, affine invariance and (3.9),

$$W(\nu; s) \leq \mathbb{E}[W(\nu; s/|R_\sigma|)] \leq W(\nu; t) \rho_{\min}^2. \quad (3.11)$$

Passing to a decreasing subsequence of (r_n) and applying (3.11) with $t = r_n$ shows that $W(\nu; s) \rightarrow 0$ as $s \rightarrow 0$.

Next we show that $H(\nu; r \mid 2r) \rightarrow \log 2$ as $r \rightarrow 0$. Let $\varepsilon > 0$, and choose δ and A from Lemma 3.10, with $C = 2$. Since $W(\nu; \delta r) \rightarrow 0$, for all sufficiently small r there is a scale- δr convolution disintegration

$$\nu = \int_I \lambda_i^{(1)} * \dots * \lambda_i^{(m_i)} d\theta(i)$$

such that, writing $S_i = \sum_{j=1}^{m_i} \text{Var}(\lambda_i^{(j)})$,

$$\int_I \exp\left(-\frac{4S_i}{\delta^2 r^2}\right) d\theta(i) < \varepsilon \exp\left(-\frac{4A}{\delta^2}\right).$$

It follows that $\theta(\{i : S_i < Ar^2\}) < \varepsilon$. For every other i , Lemma 3.10 gives

$$H\left(\lambda_i^{(1)} * \dots * \lambda_i^{(m_i)}; r \mid 2r\right) \geq \log 2 - \varepsilon.$$

The concavity of entropy between two scales therefore gives

$$H(\nu; r \mid 2r) \geq \log 2 - O(\varepsilon).$$

Consequently, as $r \rightarrow 0$,

$$H(\nu; r) = \log r^{-1} + o(\log r^{-1}),$$

and hence $\dim \nu = 1$. $\square$

3.3. Variance Summation. Throughout this section write $\rho_{\min} = \min_{1 \leq i \leq k} |r_i|$. The following lemma allows us to combine variance found in different stopped random walks and is a simple one-dimensional analogue of [KK24, Proposition 6.4]. As outlined in the sketch of the argument, we will apply Lemma 3.11 twice, namely in the proof of Proposition 3.12 and of Theorem 1.1.

Lemma 3.11 (Concatenation). *Let $0 < s_1 < \dots < s_m$ and let $\tau_1, \dots, \tau_m$ be bounded stopping times for independent walks with independent increments of law μ . Write (C_j, D_j) for their signed contraction ratio and translation. Suppose that, for $j < m$,*

$$s_j \leq \rho_{\min} s_{j+1} |C_j| \quad \text{almost surely.} \quad (3.12)$$

For $0 < r \leq s_1$, these blocks can be concatenated to one walk ending at a bounded stopping time T , such that

$$\mathbb{E}[W(\mathcal{L}(B_T \mid R_T); r)] \leq \prod_{j=1}^m \mathbb{E}[W(\mathcal{L}(D_j \mid C_j); s_j)]^{\rho_{\min}^2}, \quad (3.13)$$

and consequently

$$W(\nu; r) \leq \prod_{j=1}^m \mathbb{E}[W(\mathcal{L}(D_j \mid C_j); s_j)]^{\rho_{\min}^2}. \quad (3.14)$$

Moreover, $|R_T| > \rho_{\min} r |C_m| / s_m$ almost surely.

Proof. For $1 \leq j \leq m$ define

$$S_j = \min\{n \geq 0 : |R_n| \leq r/s_j\} \quad \text{and} \quad T_j = S_j + \tau_j^{(S_j)},$$

where $\tau_j^{(S_j)}$ applies the j th stopping rule to the increments after S_j . In other words the law of $(X_{S_j+1}, X_{S_j+2}, \dots, X_{S_j+\tau_j^{(S_j)}})$ given $X_1, \dots, X_{S_j}$ is $\mu^{*\tau_j}$. Denote $A_j = R_{S_j}$. By construction

$$\rho_{\min} r / s_j < |A_j| \leq r / s_j. \quad (3.15)$$

Thus (3.12) implies

$$|R_{T_j}| = |A_j C_j| > r/s_{j+1},$$

so $T_j < S_{j+1}$. Therefore, the selected blocks are disjoint and their stopped products are independent. As $\max_i |r_i| < 1$ all stopping times are uniformly bounded. We take $T = T_m$ and the final assertion that $|R_T| > \rho_{\min} r |C_m|/s_m$ follows from (3.15).

Let $\mathcal{A}$ include the ratios $C_1, \dots, C_m$ and the complete words in all intervening gaps including the initial gap. Then the A_j and R_T are $\mathcal{A}$ -measurable, and

$$B_T = d_{\mathcal{A}} + \sum_{j=1}^m A_j D_j \quad (3.16)$$

for an $\mathcal{A}$ -measurable translation $d_{\mathcal{A}}$. Given $\mathcal{A}$, the D_j are independent and have laws $\mathcal{L}(D_j | C_j)$.

Put $v_j(c) = W(\mathcal{L}(D_j | C_j = c); s_j)$. From (3.16), (3.8) and (3.15),

$$W(\mathcal{L}(B_T | \mathcal{A}); r) \leq \prod_{j=1}^m W(\mathcal{L}(D_j | C_j); r/|A_j|) \leq \prod_{j=1}^m v_j(C_j)^{\rho_{\min}^2}.$$

Average and use convexity to pass from $\mathcal{A}$ to R_T . The C_j are independent, so Jensen's inequality yields

$$\mathbb{E} \left[\prod_j v_j(C_j)^{\rho_{\min}^2} \right] = \prod_j \mathbb{E}[v_j(C_j)^{\rho_{\min}^2}] \leq \prod_j \mathbb{E}[v_j(C_j)]^{\rho_{\min}^2}.$$

This proves (3.13). Adding the tail $R_T Z'$ with $Z' \sim \nu$ independent of the random walk proves (3.14). $\square$

We next state the following proposition, which is similar to [KK24, Proposition 6.7] or [KK25, Proposition 5.8] but importantly does not rely on any separation assumption as we use Corollary 3.4.

Proposition 3.12. *Suppose $\dim \nu < \min\{1, h_\mu/|\chi_\mu|\}$. Then there is $\kappa > 0$ such that, for every $\varepsilon > 0$, there is a bounded stopping time τ and some $r > 0$ satisfying*

$$\mathbb{E}[W(\mathcal{L}(B_\tau | R_\tau); r)] \leq e^{-\kappa}, \quad r < \varepsilon |R_\tau| \quad \text{almost surely.} \quad (3.17)$$

Proof. Fix $C > \rho_{\min}^{-1}$ with $\rho_{\min} = \min_{1 \leq i \leq k} |r_i|$ and let η be as in Corollary 3.4. For sufficiently large N , write

$$w_N(s) = \mathbb{E}[V_{\mathcal{L}(B_N | R_N)}(s)] \quad \text{and} \quad d_N = (N+1) \log \rho_{\min}^{-1}.$$

Then $\int_0^{C^{-N}} w_N(s) ds/s \geq \eta N$ for a suitable $\eta > 0$. The conditional laws involved are finitely supported, so $w_N(s) = 0$ for all sufficiently small s . Averaging the logarithmic grids $\{e^{u+jd_N} : j \in \mathbb{Z}\}$ over $u \in [0, d_N)$ gives

$$\frac{1}{d_N} \int_0^{d_N} \sum_{e^{u+jd_N} < C^{-N}} w_N(e^{u+jd_N}) du = \frac{1}{d_N} \int_0^{C^{-N}} w_N(s) \frac{ds}{s} \geq \frac{\eta}{2 \log \rho_{\min}^{-1}}.$$

Consequently, some finite collection $0 < s_1 < \dots < s_m < C^{-N}$ satisfies

$$s_{j+1} \geq \rho_{\min}^{-(N+1)} s_j \quad \text{and} \quad \sum_{j=1}^m w_N(s_j) \geq \eta_0 := \frac{\eta}{2 \log \rho_{\min}^{-1}} > 0. \quad (3.18)$$

Apply Lemma 3.11 with $r = s_1$ and with every block having the fixed length N . Condition (3.12) holds because $|R_N| \geq \rho_{\min}^N$. Using (3.10) and $1 - x \leq e^{-x}$ gives

$$\begin{aligned} \mathbb{E}[W(\mathcal{L}(B_T | R_T); r)] &\leq \prod_{j=1}^m (1 - (1 - e^{-1})w_N(s_j))^{\rho_{\min}^2} \\ &\leq \exp(-\rho_{\min}^2(1 - e^{-1})\eta_0). \end{aligned}$$

Moreover, the last assertion of Lemma 3.11 gives

$$\frac{r}{|R_T|} < \frac{s_m}{\rho_{\min}^{N+1}} < \rho_{\min}^{-1}(C\rho_{\min})^{-N}.$$

Since $C\rho_{\min} > 1$, choose N sufficiently large that the last expression is at most ε . Taking $\tau = T$ and $\kappa = \rho_{\min}^2(1 - e^{-1})\eta_0$ proves the proposition. $\square$

3.4. Conclusion.

Proof of Theorem 1.1. The upper bound

$$\dim \nu \leq \min \left\{ 1, \frac{h_\mu}{|\chi_\mu|} \right\}$$

is well known. In particular it follows from Feng–Hu [FH09]. Indeed, apply their projection-entropy bound [FH09, Theorem 2.2(i)] to the system given by the distinct similarities in $\text{supp}(\mu^{*n})$. Its symbolic entropy is $H(\mu^{*n})$ and its Lyapunov exponent is $n\chi_\mu$, so $\dim \nu \leq \min\{1, \frac{H(\mu^{*n})}{n|\chi_\mu|}\}$. Letting $n \rightarrow \infty$ gives the claimed upper bound.

It remains to show that $\dim \nu \geq \min\{1, \frac{h_\mu}{|\chi_\mu|}\}$. Assume for contradiction that this does not hold.

Choose (τ_1, r_1) from Proposition 3.12 with 1 in the role of ε . Inductively choose (τ_{j+1}, r_{j+1}) with $\rho_{\min}r_j$ in the role of ε . Thus

$$r_{j+1} < \rho_{\min}r_j|R_{\tau_{j+1}}| \quad \text{and} \quad \mathbb{E}[W(\mathcal{L}(B_{\tau_j} | R_{\tau_j}); r_j)] \leq e^{-\kappa}. \quad (3.19)$$

For each n , apply Lemma 3.11 at the scale r_n with the blocks ordered as

$$(\tau_n, r_n), (\tau_{n-1}, r_{n-1}), \dots, (\tau_1, r_1).$$

The separation hypothesis is exactly (3.19). Consequently

$$W(\nu; r_n) \leq e^{-n\kappa\rho_{\min}^2} \rightarrow 0.$$

Proposition 3.6 gives $\dim \nu = 1$, a contradiction. This proves the theorem. $\square$

For completeness we record that Corollary 1.2 follows from Theorem 1.1 exactly as [Hoc14, Corollary 1.2] from [Hoc14, Theorem 1.1].

REFERENCES

- [Bak21] S. Baker, *Iterated function systems with super-exponentially close cylinders*, Adv. Math. **379** (2021), 107548, DOI 10.1016/j.aim.2020.107548.
- [Bak22] ———, *Iterated function systems with super-exponentially close cylinders II*, Proc. Amer. Math. Soc. **150** (2022), no. 1, 245–256, DOI 10.1090/proc/15644.
- [BK21] B. Bárány and A. Käenmäki, *Super-exponential condensation without exact overlaps*, Adv. Math. **379** (2021), 107549, DOI 10.1016/j.aim.2020.107549.
- [BV20] E. Breuillard and P. P. Varjú, *Entropy of Bernoulli convolutions and uniform exponential growth for linear groups*, J. Anal. Math. **140** (2020), no. 2, 443–481, DOI 10.1007/s11854-020-0100-0.
- [BV19] ———, *On the dimension of Bernoulli convolutions*, Ann. Probab. **47** (2019), no. 4, 2582–2617, DOI 10.1214/18-AOP1324.

- [EP63] J. D. Esary and F. Proschan, *Coherent structures of non-identical components*, Technometrics **5** (1963), no. 2, 191–209, DOI 10.1080/00401706.1963.10490075.
- [FF25] D.-J. Feng and Z. Feng, *Dimension of homogeneous iterated function systems with algebraic translations*, J. Lond. Math. Soc. (2) **112** (2025), no. 1, e70222, DOI 10.1112/jlms.70222.
- [FH09] D.-J. Feng and H. Hu, *Dimension theory of iterated function systems*, Comm. Pure Appl. Math. **62** (2009), no. 11, 1435–1500.
- [Gol10] L. Goldstein, *Bounds on the constant in the mean central limit theorem*, Ann. Probab. **38** (2010), no. 4, 1672–1689, DOI 10.1214/10-AOP527.
- [Hoc18] M. Hochman, *Dimension theory of self-similar sets and measures*, Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018 (B. Sirakov, P. Ney de Souza, and M. Viana, eds.), Vol. III, World Sci. Publ., Hackensack, NJ, 2018, pp. 1949–1972, DOI 10.1142/9789813272880.0125.
- [Hoc14] ———, *On self-similar sets with overlaps and inverse theorems for entropy*, Ann. of Math. (2) **180** (2014), no. 2, 773–822, DOI 10.4007/annals.2014.180.2.7.
- [Hut81] J. E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30** (1981), no. 5, 713–747, DOI 10.1512/iumj.1981.30.30055.
- [Kit25] S. Kittle, *Absolutely continuous Furstenberg measures*, Proc. Lond. Math. Soc. (3) **131** (2025), no. 1, e70072, DOI 10.1112/plms.70072.
- [KK25] S. Kittle and C. Kogler, *Dimension of contracting on average self-similar measures* (2025), available at [arXiv:2501.17795v1](https://arxiv.org/abs/2501.17795).
- [KK24] ———, *On absolute continuity of inhomogeneous and contracting on average self-similar measures* (2024), available at [arXiv:2409.18936](https://arxiv.org/abs/2409.18936).
- [KK26S] ———, *Supplementary material for the exact overlaps conjecture for self-similar measures on the real line* (2026), <https://github.com/ckkogler/kk26-supplementary-material>. Supplementary material repository. All documents written by GPT-6 Astra.
- [KMD16] I. Klein, B. Mangold, and M. Doll, *Cumulative paired ϕ -entropy*, Entropy **18** (2016), no. 7, 248, DOI 10.3390/e18070248.
- [KK26] C. Kogler and S. Kittle, *A Lean formalization of the exact overlaps conjecture in dimension one* (2026), <https://github.com/ckkogler/exact-overlaps-one-dim-lean>. Lean source code.
- [Ope] OpenAI, *Math*, <https://github.com/openai/math>. GitHub repository.
- [LP17] G. Last and M. Penrose, *Lectures on the Poisson process*, Institute of Mathematical Statistics Textbooks, vol. 7, Cambridge University Press, Cambridge, 2017.
- [Mar74] G. A. Margulis, *Probabilistic characteristics of graphs with large connectivity*, Probl. Peredachi Inf. **10** (1974), no. 2, 101–108. English translation: Problems Inform. Transmission 10 (1974), no. 2, 174–179.
- [Rap22] A. Rapaport, *Proof of the exact overlaps conjecture for systems with algebraic contractions*, Ann. Sci. Éc. Norm. Supér. (4) **55** (2022), no. 5, 1357–1377, DOI 10.24033/asens.2518.
- [RV24] A. Rapaport and P. P. Varjú, *Self-similar measures associated to a homogeneous system of three maps*, Duke Math. J. **173** (2024), no. 3, 513–602, DOI 10.1215/00127094-2023-0019.
- [Rus81] L. Russo, *On the critical percolation probabilities*, Z. Wahrsch. Verw. Gebiete **56** (1981), no. 2, 229–237, DOI 10.1007/BF00535742.
- [Shm19] P. Shmerkin, *On Furstenberg’s intersection conjecture, self-similar measures, and the L^q norms of convolutions*, Ann. of Math. (2) **189** (2019), no. 2, 319–391, DOI 10.4007/annals.2019.189.2.1.
- [Sim96] K. Simon, *Overlapping cylinders: the size of a dynamically defined Cantor-set*, Ergodic Theory and $\mathbb{Z}^d$ Actions (M. Pollicott and K. Schmidt, eds.), London Math. Soc. Lecture Note Ser., vol. 228, Cambridge Univ. Press, Cambridge, 1996, pp. 259–272, DOI 10.1017/CBO9780511662812.009.
- [Str24] L. Streck, *On the Fourier decay and the dimension of self-similar measures*, Ph.D. thesis, University of Cambridge, 2024. DOI 10.17863/CAM.116133.
- [Tit72] J. Tits, *Free subgroups in linear groups*, J. Algebra **20** (1972), no. 2, 250–270, DOI 10.1016/0021-8693(72)90058-0.

- [Var19a] P. P. Varjú, *Absolute continuity of Bernoulli convolutions for algebraic parameters*, J. Amer. Math. Soc. **32** (2019), no. 2, 351–397, DOI 10.1090/jams/916.
- [Var] ———, *Entropy rates in the dimension theory of self-similar measures*, available at [arXiv:2509.22042](#). Version 2, 22 March 2026.
- [Var19b] ———, *On the dimension of Bernoulli convolutions for all transcendental parameters*, Ann. of Math. (2) **189** (2019), no. 3, 1001–1011.
- [Wan11] Z. Wang, *Quantitative density under higher rank abelian algebraic toral actions*, Int. Math. Res. Not. IMRN **2011** (2011), no. 16, 3744–3821.
- [Wu25] M. Wu, *Projection theorems with countably many exceptions and applications to the exact overlaps conjecture* (2025), available at [arXiv:2503.21923](#).
- [Zuy92] S. A. Zuyev, *Russo’s formula for Poisson point fields and its applications*, Diskretnaya Matematika **4** (1992), no. 3, 149–160.

Email address: s.kittle@ucl.ac.uk, kogler@ias.edu

SAMUEL KITTLE, DEPARTMENT OF MATHEMATICS, UNIVERSITY COLLEGE LONDON, 25 GORDON STREET, LONDON WC1H 0AY, UNITED KINGDOM

CONSTANTIN KOGLER, INSTITUTE FOR ADVANCED STUDY, 1 EINSTEIN DR, PRINCETON, NJ 08540, UNITED STATES OF AMERICA