ON L^p DENSITY OF SELF SIMILAR MEASURES AND BERNOULLI CONVOLUTIONS

SAMUEL KITTLE AND CONSTANTIN KOGLER

Abstract. We prove that an absolutely continuous self-similar measure has an L^p density for some $p > 1$. Moreover, explicit examples of Bernoulli convolutions with a continuous density are established.

1. Introduction

Self-similar measures are central objects in fractal geometry [Fal14, BSS23]. Understanding when they are absolutely continuous and establishing results about their regularity is intricate. In this paper, we establish the fundamental property that an absolutely continuous self-similar measure has a density in L^p for some $p > 1$. Moreover, we apply the quantitative proof method of the latter result to establish explicit examples of Bernoulli convolutions with a continuous density going beyond the only previously known case of Garsia numbers. All new results of this paper are formalised in Lean.

A contracting similarity on $\mathbb{R}^d$ is a map of the form $x \mapsto \rho U x + b$ with $\rho \in (0,1)$, U a $d \times d$ orthogonal matrix and $b \in \mathbb{R}^d$. Let $\mu = \sum_{i=1}^k p_i \delta_{g_i}$ be a probability measure supported on finitely many contracting similarities on $\mathbb{R}^d$. The self-similar measure of μ is the unique probability measure ν on $\mathbb{R}^d$ satisfying

$$\mu * \nu := \sum_{i=1}^{k} p_i (g_i)_* \nu = \nu. \tag{1.1}$$

Ever since Hutchinson's 1981 paper [Hut81], where it is proved that they exist and have compact support, self-similar measures have been well studied. We establish the following observation about self-similar measures.

Theorem 1.1. *Let μ be a probability measure supported on finitely many contracting similarities of $\mathbb{R}^d$ and denote by ν the self-similar measure of μ. If ν is absolutely continuous with respect to the Lebesgue measure, then the density of ν is in $L^p(\mathbb{R}^d)$ for some $p > 1$.*

The proof of Theorem 1.1 is short and combines self-similarity with a general higher integrability result. The key lemma of the proof and this paper is Lemma 2.1, which is a version of the reverse Hölder criterion of Spadaro [Spa12], as further developed by Kinnunen–Kurki–Mudarra [KKM22]. Lemma 2.1 assumes for an absolutely continuous measure ν on $\mathbb{R}^d$ control over the mass of $\nu(E)$ for sets E of small measure and then establishes a regularity gain for ν. The proof of Theorem 1.1 will then follow by showing that self-similarity prevents the mass of $\nu(E)$ being too large. Similar higher integrability arguments are used by Giaquinta–Modica [GM86, Theorems 3.1 and 4.1] in the regularity theory of elliptic equations and by De Lellis–Spadaro [DLS14, Section 5] in geometric measure theory.

The exponent $p > 1$ in Theorem 1.1 and Lemma 2.1 can be estimated quantitatively from the mass that ν assigns to sets of small Lebesgue measure. This insight allows us to use recent methods for establishing explicit examples of absolutely continuous self-similar measures [Var19, Kit24, KK25] to obtain suitable measure estimates, yielding quantitative L^p results that we use to construct Bernoulli convolutions with continuous densities.

The Bernoulli convolution ν_λ of parameter $\lambda \in (0,1)$ is the self-similar measure of the balanced probability measure on the similarities $x \mapsto \lambda x + 1$ and $x \mapsto \lambda x - 1$ with $x \in \mathbb{R}$. Another way to describe them is as the law of the random variable $\sum_{n=0}^{\infty} \xi_n \lambda^n$, where the ξ_n are independent and satisfy $\mathbb{P}(\xi_n = 1) = \mathbb{P}(\xi_n = -1) = 1/2$ for all $n \geq 0$. The study of Bernoulli convolutions has a rich history and we refer to surveys by Peres–Schlag–Solomyak [PSS00] and Varjú [Var18].

We highlight that Solomyak [Sol95] showed in 1995 that ν_λ is in L^2 for almost all $\lambda \in (1/2, 1)$. Another important result is due to Shmerkin [Shm19, Theorem 1.3], who showed that ν_λ has a density in L^p for every finite $p > 1$ and all parameters $\lambda \in (1/2, 1)$ outside a set of Hausdorff dimension zero.

Splitting the series $\sum_{n\geq 0} \xi_n \lambda^n$ into even and odd n gives $\nu_\lambda = \nu_{\lambda^2} * (D_\lambda)_* \nu_{\lambda^2}$ with $D_\lambda(x) = \lambda x$. Therefore, by Shmerkin [Shm19, Theorem 1.3] applied to L^2, for all $\lambda \in (1/\sqrt{2}, 1)$ except for a set of Hausdorff dimension zero, the factors ν_{λ^2} and $(D_\lambda)_* \nu_{\lambda^2}$ are both in L^2. So it follows that ν_λ has a continuous density for $\lambda \in (1/\sqrt{2}, 1)$ outside a set of Hausdorff dimension zero. By a similar argument using Young's inequality, it follows that if $\nu_\lambda \in L^p$, then $\nu_{\lambda^{1/m}}$ has a continuous density for an integer m satisfying $m \geq p/(p-1)$. It is therefore a direct consequence of Theorem 1.1 that if ν_λ is absolutely continuous, then $\nu_{\lambda^{1/m}}$ has a continuous density for sufficiently large m. As we are moreover able to make p quantitative, we are led to fully explicit examples of Bernoulli convolutions with a continuous density.

Denote by M_λ the Mahler measure of an algebraic number λ. Garsia proved in 1962 that ν_λ has an L^∞ density if $\lambda \in (1/2, 1)$ is a Garsia number, that is, λ satisfies $M_\lambda = 2$. There was no progress on understanding absolute continuity for explicit λ until Varjú [Var19] proved in 2019 that ν_λ is absolutely continuous if $\lambda > 1 - c \min\{\log M_\lambda, (\log M_\lambda)^{-2}\}$ for $c > 0$ an absolute constant. Varjú's result was improved in [KK25], addressing general self-similar measures, by showing that ν_λ is absolutely continuous if $\lambda > 1 - c \min\{\log M_\lambda, (\log \log M_\lambda)^{-2}\}$ for $c > 0$ an absolute constant. Yu [Yu22] proved in 2022 that ν_λ has a continuous density if λ is a Garsia number satisfying $\lambda \in [1/\sqrt{2}, 1)$.

In this paper, we give the first explicit non-Garsia examples of Bernoulli convolutions with a continuous density.

Theorem 1.2. *If $\lambda \in (1/2, 1)$ is algebraic of degree d_λ, has Mahler measure M_λ and satisfies*

$$\lambda \geq 1 - \left(\frac{\min\{1, \log M_\lambda\}}{1 + d_\lambda + \log M_\lambda} \right)^{100}, \tag{1.2}$$

then ν_λ is absolutely continuous with a density in $C_c(\mathbb{R})$.

We further discuss the proof of Theorem 1.2 in Section 3 and note that the exponent 100 is chosen to make the argument elegant. In fact, a refinement available at [KK] of the methods presented in this paper allow to lower the exponent to 3.

Theorem 1.3. *([KK]) If $\lambda \in (1/2, 1)$ is algebraic of degree d_λ, has Mahler measure M_λ and satisfies*

$$\lambda \geq 1 - \left(\frac{\min\{1, \log M_\lambda\}}{1 + d_\lambda + \log M_\lambda} \right)^3, \tag{1.3}$$

then ν_{λ^2} is absolutely continuous with a density in $L^2(\mathbb{R})$ and therefore ν_λ has a density in $C_c(\mathbb{R})$.

When one applies Theorem 1.3 to rational numbers, one obtains the following.

Corollary 1.4. *If $\lambda = 1 - p/q \in (1/2, 1)$ is in lowest terms and*

$$p \leq \frac{q}{(2 + \log q)^3},$$

then ν_λ is absolutely continuous with a density in $C_c(\mathbb{R})$. In particular, $\nu_{1-\frac{1}{n}}$ is absolutely continuous with a density in $C_c(\mathbb{R})$ for all $n \geq 588$.

Proof. As $d_\lambda = 1$ and $M_\lambda = q \geq 3$ the claimed inequality follows from Theorem 1.3. The desired result for $\nu_{1-\frac{1}{n}}$ follows from a brief calculation. □

Machine Contribution Statement. The proofs of this paper are due to GPT-6 Astra. On September 10th 2026, the second-named author (C.K.) asked GPT-6 Astra to find explicit examples of self-similar measures with an L^p density resulting in the discovery of Theorem 1.1. Various simplifications led to Lemma 2.1. Realising that Theorem 1.1 is quantitative, numerous conversations led to Theorem 1.2. The authors rewrote the paper and take responsibility for its correctness. A talk by the second-named author on the results of this paper is available at [Kog26].

Lean Formalization. GPT-6 Astra formalized all results from this paper and all necessary results from previous work. The Lean source code is available in the accompanying repository [Kog].

Acknowledgments. 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 1.1

Write $|E|$ for the Lebesgue measure of a subset $E \subset \mathbb{R}^d$. All balls considered are open. If $B = B(x, r)$ is an open ball around $x \in \mathbb{R}^d$ and $r > 0$, then we denote for $a > 0$ by

$$aB = B(x, ar).$$

The main idea of the paper is to apply the following lemma to self-similar measures. It is a form of the weak A_∞ criterion of Spadaro [Spa12, Theorem 1.1], [KKM22, Theorem 3.1 and Section 4] or [AHT17, Corollary 5.5].

Lemma 2.1 (Weak reverse Hölder criterion)**.** *For every $d \geq 1$ write $C = 18^d$. Let $f \geq 0$ belong to $L^1(\mathbb{R}^d)$. Suppose that for some $\varepsilon \in (0, 1)$ and $\eta \in (0, C^{-1})$ it holds that*

$$|E| \leq \varepsilon |B| \quad \Longrightarrow \quad \int_E f \, dx \leq \eta \int_{2B} f \, dx \tag{2.1}$$

for every ball B of radius less than one and every measurable $E \subset B$. Then $f \in L^p(\mathbb{R}^d)$ for $p = 1 + \frac{\log(C\eta)^{-1}}{\log(C\varepsilon^{-1})} > 1$.

Proof of Lemma 2.1. Write

$$A = C\varepsilon^{-1}, \qquad \gamma = 2 \cdot 6^d \eta \quad \text{and} \quad F(t) = \int_{\{f>t\}} f\,dx.$$

We claim that, for all sufficiently large t,

$$F(At) \le \gamma F(t). \tag{2.2}$$

Let $E = \{f > At\}$. By the Lebesgue differentiation theorem, almost all $x \in E$ satisfy that $\lim_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} f\,dx = f(x) > At > 2t$. For each such point $x \in E$, choose the ball B_x centered at x with largest radius whose average of f is at least $2t$. Continuity of ball averages gives

$$\frac{1}{|B_x|}\int_{B_x} f\,dx = 2t \quad \text{and} \quad \frac{1}{|aB_x|}\int_{aB_x} f\,dx \le 2t \quad \text{where} \quad a > 1.$$

For sufficiently large t, every ball $3B_x$ has radius less than one, since $2t|B_x| \le \|f\|_1$.

Cover a compact subset $K \subset E$ of these Lebesgue points by finitely many such balls and then choose a selection of disjoint balls B_i with $K \subset \bigcup_i 3B_i$. Since for each i,

$$|E \cap 3B_i| \le \frac{1}{At}\int_{3B_i} f\,dx \le \frac{2}{A}|3B_i| \le \varepsilon|3B_i|,$$

the hypothesis (2.1) yields

$$\int_K f\,dx \le \eta \sum_i \int_{6B_i} f\,dx \le 2 \cdot 6^d \eta\, t \sum_i |B_i| = \gamma t \sum_i |B_i|.$$

On the other hand, $2t|B_i| = \int_{B_i} f\,dx \le t|B_i| + \int_{B_i\cap\{f>t\}} f\,dx$. Disjointness therefore gives $t\sum_i |B_i| \le F(t)$. By inner regularity, $F(At) \le \gamma F(t)$, proving (2.2).

Choose $T > 0$ so that (2.2) holds for $t \ge T$, and write $I_R = \int_T^R t^{p-2}F(t)\,dt$. Note that $\gamma A^{p-1} = \frac{2\cdot 6^d\eta}{C\eta} = 2\left(\frac{1}{3}\right)^d \le \frac{2}{3}$. Therefore, for $R \ge AT$, changing variables in (2.2) gives $I_R \le I_{AT} + \gamma A^{p-1} I_{R/A} \le I_{AT} + \frac{2}{3}I_R$. Thus $I_R \le 3I_{AT}$ and hence I_R is uniformly bounded. Letting $R \to \infty$ and using $F \le \|f\|_1$, Tonelli's theorem gives

$$\frac{1}{p-1}\int_{\mathbb{R}^d} f^p\,dx = \int_{\mathbb{R}^d} f(x)\left(\int_0^{f(x)} t^{p-2}\,dt\right)dx = \int_0^\infty t^{p-2}F(t)\,dt < \infty.$$

This proves that $f \in L^p(\mathbb{R}^d)$ and so concludes the proof. □

Theorem 1.1 follows by applying Lemma 2.1 together with self-similarity.

Proof of Theorem 1.1. Without loss of generality we may normalise ν such that $\operatorname{diam} K = 1$, where $K = \operatorname{supp}\nu$. For a similarity g, we write $\rho(g)$ for the contraction factor of g and denote

$$\rho_* = \min_i \rho(g_i) > 0 \quad \text{and} \quad \omega(t) = \sup\{\nu(E) : |E| \le t\}.$$

Absolute continuity gives $\omega(t) \to 0$ as $t \downarrow 0$. We claim that for every ball $B = B(x,r)$ with $0 < r < 1$ and every measurable $E \subset B$,

$$\nu(E) \le \omega\left(c\frac{|E|}{|B|}\right)\nu(2B) \qquad \text{for} \qquad c = |B(0,1)|\rho_*^{-d}. \tag{2.3}$$

Since $\omega(t) \to 0$, by (2.3) it follows that the density function of ν verifies the hypothesis of Lemma 2.1 for every $\eta > 0$, which proves the theorem.

So it remains to prove (2.3). Let $\gamma_1, \gamma_2, \cdots$ be independent samples from μ. Write $q_n = \gamma_1 \cdots \gamma_n$ and consider the stopping time $\tau_r = \min\{n \geq 1 : \rho(q_n) \leq r\}$ and note $\rho_* r < \rho(q_{\tau_r}) \leq r$. It is well known that $q_{\tau_r}\nu$ has the same distribution as ν, as is shown in [Kit25, Lemma 2.19], in other words $\mathbb{E}\big[(q_{\tau_r})_*\nu\big] = \nu$.

Consider a set $E \subset B$. Observe that $|q_{\tau_r}^{-1}E| = \rho(q_{\tau_r})^{-d}|E| \leq c\frac{|E|}{|B|}$. Moreover, for $\nu(q_{\tau_r}^{-1}E) \neq 0$ it must hold that $q_{\tau_r}K \cap B \neq \emptyset$. So consider the event $\mathcal{H} = \{q_{\tau_r}K \cap B \neq \varnothing\}$. On $\mathcal{H}$, the set $q_{\tau_r}K$ has diameter at most r, so it is contained in $2B$ and therefore $\mathbb{P}(\mathcal{H}) \leq \mathbb{E}\big[\nu(q_{\tau_r}^{-1}(2B))\big] = \nu(2B)$. Thus, it follows by the definition of ω,

$$\nu(E) = \mathbb{E}\big[\nu(q_{\tau_r}^{-1}E)\big] \leq \omega\bigg(c\frac{|E|}{|B|}\bigg)\,\mathbb{P}(\mathcal{H}) \leq \omega\bigg(c\frac{|E|}{|B|}\bigg)\,\nu(2B),$$

which proves (2.3). □

3. Bernoulli convolutions

Throughout this section, write d_λ for the degree of an algebraic number $\lambda \in (0,1)$ and

$$\chi_\lambda = -\log\lambda \qquad \text{and} \qquad B_\lambda = \frac{1 + d_\lambda + \log M_\lambda}{\min\{1, \log M_\lambda\}}.$$

Thus (1.2) is equivalent to $(1-\lambda)B_\lambda^{100} \leq 1$. All logarithms are natural. The main result of this section is the following proposition.

Proposition 3.1 (A quantitative gain over L^1)**.** *If $\lambda \in (1/2, 1)$ is algebraic and $\chi_\lambda B_\lambda^{10} \leq 10^{-100}$, then ν_λ is absolutely continuous with density f_λ and*

$$f_\lambda \in L^{1+\delta_\lambda}(\mathbb{R}) \quad \text{for} \quad \delta_\lambda = \frac{1}{10(1 - \log(1-\lambda))}. \tag{3.1}$$

In Section 3.1 we deduce Theorem 1.2 from Proposition 3.1 by applying Young's inequality and then prove Proposition 3.1 in Section 3.2 and Section 3.3. The proof of Proposition 3.1 uses the detail of a measure, which is a concept introduced by the first-named author [Kit24], and ideas from Varjú [Var19] and the authors' paper [KK25]. As the proof of Proposition 3.1 is an illuminating series of ideas, we will not further outline it and simply let it unfold as presented below.

3.1. Deduction of Theorem 1.2.

We will use Young's inequality in the following form.

Lemma 3.2 (Young's inequality for scaled convolutions)**.** *Let $k \geq 2$ be an integer, let $a_1, \ldots, a_k > 0$, and let ν be a compactly supported absolutely continuous probability measure on $\mathbb{R}$ with a density in $L^p(\mathbb{R})$ for some $p \geq k/(k-1)$. Then*

$$\mathop{*}_{j=1}^{k} (D_{a_j})_*\nu$$

has a density in $C_c(\mathbb{R})$, where $D_a(x) = ax$.

Proof. For $1 \leq j \leq k$, let f_j be the density of $(D_{a_j})_*\nu$. Compact support and $p \geq k/(k-1)$ imply $f_j \in L^{k/(k-1)}$. By iterating Young's convolution inequality, the function $g = f_1 * \cdots * f_{k-1}$ is in L^k since $\frac{k-1}{k/(k-1)} - (k-2) = \frac{1}{k}$. Write $h = g * f_k$. Hölder's inequality gives $\|h\|_\infty \leq \|g\|_k \|f_k\|_{k/(k-1)}$ and, for every $x, t \in \mathbb{R}$,

$$|h(x+t) - h(x)| \leq \|g\|_k\, \|f_k(\cdot + t) - f_k\|_{k/(k-1)}.$$

The right-hand side tends to zero as $t \to 0$, by continuity of translations in $L^{k/(k-1)}$. Hence h is uniformly continuous. Moreover, its support is contained in the sum of the compact supports of the factors, so h is compactly supported. □

We will now deduce Theorem 1.2 from Proposition 3.1 and Lemma 3.2.

Proof of Theorem 1.2. We let $\theta = \lambda^k$ for some positive integer k we will choose later. We wish to show that θ satisfies the conditions of Proposition 3.1 and for Lemma 3.2 that

$$\delta_\theta \geq \frac{1}{k-1}. \tag{3.2}$$

First note that since $\theta \leq \lambda$ we have $\delta_\theta \geq \delta_\lambda$. Therefore for (3.2) to hold it is sufficient that $k \geq \frac{1}{\delta_\lambda} + 1 = 11 - 10\log(1-\lambda)$. Thus, we take

$$k = \lfloor 20 - 20\log(1-\lambda) \rfloor$$

and claim that

$$\chi_\theta B_\theta^{10} \leq 10^{-100}. \tag{3.3}$$

As $B_\theta \geq 1$, (3.3) implies $\theta \in (1/2, 1)$ and so Proposition 3.1 gives $f_\theta \in L^{1+\delta_\theta}$. Splitting the defining series modulo k gives $\nu_\lambda = *_{j=0}^{k-1}(D_{\lambda^j})_*\nu_\theta$, so (3.2) and Lemma 3.2 give a continuous density for ν_λ, proving the theorem.

To prove (3.3), note that standard height identities [BG06, Section 1.5 and Section 1.6] give

$$\frac{\log M_\theta}{d_\theta} = k\frac{\log M_\lambda}{d_\lambda} \quad \text{and} \quad \frac{d_\lambda}{k} \leq d_\theta \leq d_\lambda.$$

Hence $\log M_\lambda \leq \log M_\theta \leq k \log M_\lambda$ and $B_\theta \leq kB_\lambda$. Write $t = -\log(1-\lambda)$. Assumption (1.2) gives $e^{-t} = 1 - \lambda \leq B_\lambda^{-100}$ and therefore

$$t \geq 100 \log B_\lambda \quad \text{and} \quad B_\lambda^{10} \leq e^{t/10}.$$

The general estimate Lemma 3.12 below also gives $t \leq B_\lambda$. Since $B_\lambda \geq 3$, we have $t > 100$, hence $B_\lambda > 100$ and $t > 100 \log 100 > 400$. Using $k \leq 20(1+t)$ and $\chi_\lambda \leq 2(1-\lambda) = 2e^{-t}$, we obtain

$$\begin{aligned}
\chi_\theta B_\theta^{10} &\leq k^{11}\chi_\lambda B_\lambda^{10} \\
&\leq 2 \cdot 20^{11}(1+t)^{11}e^{-9t/10} \\
&\leq 2 \cdot 20^{11} \cdot 401^{11} e^{-360} < 10^{-100},
\end{aligned}$$

where we used that $(1+t)^{11}e^{-9t/10}$ is decreasing for $t \geq 400$. This concludes the proof. □

3.2. **The reduction to bounded approximation.** We now show how Lemma 2.1 is used. Indeed, Lemma 2.1 reduces Proposition 3.1 to finding a density function satisfying the property of the following lemma.

Lemma 3.3. *Let ν_λ be absolutely continuous with density function f_λ. Assume that there is a function g_λ satisfying*

$$\|f_\lambda - g_\lambda\|_1 \leq \frac{1}{2 \cdot 21^2} \quad \textit{and} \quad \|g_\lambda\|_\infty \leq \tfrac{1}{2}. \tag{3.4}$$

Then

$$f_\lambda \in L^{1+\delta_\lambda}(\mathbb{R}) \quad \textit{for} \quad \delta_\lambda = \frac{1}{10(1-\log(1-\lambda))}. \tag{3.5}$$

Proof. Normalize the support of ν_λ to $I = [0, 1]$ and denote the resulting measure and functions by $\sigma = F_\lambda\, dx$ and G_λ. Since the original support has length $2/(1-\lambda)$, we have

$$\|F_\lambda - G_\lambda\|_1 \le \frac{1}{2 \cdot 21^2} \quad \text{and} \quad \|G_\lambda\|_\infty \le K := (1-\lambda)^{-1}.$$

Thus $\omega(t) := \sup_{|E|\le t} \sigma(E) \le (2\cdot 21^2)^{-1} + Kt$. By (2.3), applied to σ with $c = 2/\lambda$, the hypothesis of Lemma 2.1 holds with

$$\eta = 21^{-2} \quad \text{and} \quad \varepsilon = \frac{\lambda}{4 \cdot 21^2 K}.$$

Consequently, $F_\lambda \in L^p$ with

$$p - 1 = \frac{\log 21}{2\log(4 \cdot 21^3 K/\lambda)} \ge \frac{1}{10(1 + \log K)} = \delta_\lambda,$$

where we used $\lambda > 1/2$ and $K \ge 1$. Compact support and affine rescaling give (3.1). □

Next we recall the definitions of detail and higher order detail from [Kit24, Kit25]. We will use detail to establish the assumptions of Lemma 3.3 and then deduce Proposition 3.1. Let η_y be the centred Gaussian density of variance y. We will write $g_\lambda = \nu_\lambda * \eta_1$, whose supremum is at most $(2\pi)^{-1/2} < 1/2$. Write $\eta_y^{(k)} = \partial_y^k \eta_y$ and set

$$Q = \|\eta_1^{(1)}\|_1^{-1} = \sqrt{\pi e/2} \in (2, 21/10).$$

For a probability measure σ on $\mathbb{R}$, $r > 0$ and an integer $k \ge 1$, define

$$s_r^{(k)}(\sigma) = r^{2k} Q^k \|\sigma * \eta_{kr^2}^{(k)}\|_1 \quad \text{and} \quad s_r(\sigma) = s_r^{(1)}(\sigma).$$

We refer the reader to [Kit24, Kit25] and [KK25] for a discussion of detail and its properties. We just highlight that for $0 < a < b$,

$$\|\sigma * \eta_{a^2} - \sigma * \eta_{b^2}\|_1 \le \frac{2}{Q} \int_a^b s_t(\sigma)\, \frac{dt}{t}. \tag{3.6}$$

We first establish the following inequality, which is a more efficient version of [Kit25, Lemma 1.14] and of [KK25, Proposition 4.5].

Lemma 3.4. *Suppose that σ is a probability measure on $\mathbb{R}$, that k is a positive integer and that $r, b > 0$ with $b \ge r/\sqrt{k}$. Then*

$$s_r(\sigma) \le 10^k \left(\sup_{r/\sqrt{k} \le t \le b} s_t^{(k)}(\sigma) + \frac{r^2}{b^2} \right).$$

Proof. The case $k = 1$ is immediate so assume $k \ge 2$. Repeated integration by parts gives

$$\begin{aligned} \frac{(-1)^{k-1}}{(k-2)!} \int_{r^2}^\infty (y - r^2)^{k-2} \sigma * \eta_y^{(k)}\, dy &= \frac{(-1)^{k-2}}{(k-3)!} \int_{r^2}^\infty (y - r^2)^{k-3} \sigma * \eta_y^{(k-1)}\, dy \\ &= \dots \\ &= -\int_{r^2}^\infty \sigma * \eta_y^{(2)}\, dy \\ &= \sigma * \eta_{r^2}^{(1)}. \end{aligned}$$

Since $\frac{d}{dy}(1-\frac{r^2}{y})^{k-1} = (k-1)r^2(y-r^2)^{k-2}y^{-k}$, the kernel $(k-1)r^2(y-r^2)^{k-2}y^{-k}\,dy$ is a probability measure on $[r^2,\infty)$. For $b \geq r/\sqrt{k}$, its mass above kb^2 is, by Bernoulli's inequality,

$$1-\left(1-\frac{r^2}{kb^2}\right)^{k-1} \leq \frac{k-1}{k}\frac{r^2}{b^2} \leq \frac{r^2}{b^2}.$$

Since $Q^{1-k}k^k/(k-1)! < 10^k$, splitting at kb^2 and using that $s_r^{(k)}(\sigma) \leq 1$ yields

$$s_r(\sigma) \leq 10^k \left(\sup_{r/\sqrt{k}\leq t\leq b} s_t^{(k)}(\sigma) + \frac{r^2}{b^2}\right),$$

concluding the proof. □

Towards applying Lemma 3.3, we note that the following estimates of detail are sufficient.

Lemma 3.5 (Integrating detail)**.** *Let $R_0 \geq 100$ and $\delta > 0$. Suppose that a probability measure σ satisfies for $R \geq R_0$, $1 \leq k \leq 10\log R$ and $r \in (0,1)$ that*

$$s_{e^{-R}}^{(k)}(\sigma) \leq 100^{-k} \quad \text{and} \quad s_r(\sigma) \leq \delta. \tag{3.7}$$

Then σ is absolutely continuous with density f and

$$\|f-\sigma*\eta_1\|_1 \leq 4\delta R_0 + \frac{1}{4R_0}. \tag{3.8}$$

Proof. We will let $R \geq 4R_0$ and apply Lemma 3.4 with $r = e^{-R}$, $k = \lceil 2\log R\rceil$ and $b = e^{-R/2}$. For $t \in [r/\sqrt{k}, b]$, we have $\log(1/t) \geq R/2 \geq R_0$ and $k \leq 10\log(R/2)$. Thus (3.7) applies uniformly on this interval, bounding the supremum by 100^{-k}. Therefore, for $R \geq 4R_0$,

$$s_{e^{-R}}(\sigma) \leq 10^k(100^{-k}+e^{-R}) \leq R^{-2}. \tag{3.9}$$

By [Kit24, Lemma 1.18], σ is absolutely continuous and denote by $f \in L^1(\mathbb{R})$ the density of σ. Since $\sigma*\eta_{a^2} \to f$ in L^1 as $a \to 0$ and letting $a \to 0$ in (3.6) with $b = 1$, gives

$$\begin{aligned}\|f-\sigma*\eta_1\|_1 &\leq \frac{2}{Q}\int_0^1 s_t(\sigma)\,\frac{dt}{t} \\ &\leq \frac{2}{Q}\left(\int_0^{4R_0}\delta\,dR + \int_{4R_0}^\infty \frac{dR}{R^2}\right) \leq 4\delta R_0 + \frac{1}{4R_0},\end{aligned}$$

where we substituted $t = e^{-R}$, used $2/Q < 1$ and the inequality (3.9) and the assumption. This proves (3.8). □

In the next subsection, we will establish the estimates necessary for Lemma 3.5 with suitable constants as stated in the following proposition.

Proposition 3.6 (Effective detail)**.** *If $\lambda \in (1/2,1)$ is algebraic and $\chi_\lambda B_\lambda^{10} \leq 10^{-100}$, then with $R_0 = 10^6 B_\lambda^3$ we have for $R \geq R_0$ and $1 \leq k \leq 10\log R$ that*

$$s_{e^{-R}}^{(k)}(\nu_\lambda) \leq 100^{-k}. \tag{3.10}$$

Moreover, for every $\lambda \in (1/2,1)$ and $r \in (0,1)$,

$$s_r(\nu_\lambda) \leq 40\chi_\lambda^{1/3}. \tag{3.11}$$

Having established all the above and in particular assuming Proposition 3.6, we prove Proposition 3.1.

Proof of Proposition 3.1. Set $R_0 = 10^6 B_\lambda^3$. Note $B_\lambda \geq 2 + d_\lambda \geq 3$ and hence $R_0 \geq 100$. Proposition 3.6 shows that ν_λ satisfies the hypotheses of Lemma 3.5 with this value of R_0 and $\delta = 40\chi_\lambda^{1/3}$. Thus ν_λ is absolutely continuous with density f_λ. Writing $g_\lambda = \nu_\lambda * \eta_1$, (3.8) gives

$$\begin{aligned}\|f_\lambda - g_\lambda\|_1 &\leq 160\chi_\lambda^{1/3} R_0 + \frac{1}{4R_0} \\ &\leq 160 \cdot 10^6 \cdot 10^{-100/3} + \frac{1}{108 \cdot 10^6} < \frac{1}{2 \cdot 21^2},\end{aligned}$$

where we used $B_\lambda \geq 3$ and $\chi_\lambda^{1/3} B_\lambda^3 = (\chi_\lambda B_\lambda^{10})^{1/3} B_\lambda^{-1/3} \leq 10^{-100/3}$ by the assumption of Proposition 3.1 that $\chi_\lambda B_\lambda^{10} \leq 10^{-100}$. Moreover, since ν_λ is a probability measure, $\|g_\lambda\|_\infty \leq \|\eta_1\|_\infty = (2\pi)^{-1/2} < \frac{1}{2}$. Thus f_λ and g_λ satisfy (3.4) and Lemma 3.3 gives $f_\lambda \in L^{1+\delta_\lambda}(\mathbb{R})$, as required. □

3.3. **Effective detail from entropy and variance.** We now prove Proposition 3.6. We first bound detail in terms of conditional variance. We then use entropy and algebraic separation to find enough conditional variance in disjoint blocks of the Bernoulli series. We will also use for probability measures $\sigma_1, \ldots, \sigma_k$ that, as shown in [KK25, Lemma 4.3],

$$s_r^{(k)}(\sigma_1 * \cdots * \sigma_k) \leq \prod_{i=1}^k s_r(\sigma_i) \quad \text{and} \quad s_r(\mathcal{N}(0,v)) = \frac{r^2}{r^2+v}, \tag{3.12}$$

where $\mathcal{N}(0,v)$ is the standard Gaussian measure and $v \geq 0$. The following is an explicit one-dimensional version of [KK25, Lemma 4.11], whose proof uses [KK25, Lemmas 4.7 and 4.10] and thereby also the Berry-Esseen theorem.

Lemma 3.7 (Berry-Essen Estimate)**.** *Let $\varepsilon > 0$ and let $U_1, \ldots, U_n$ be independent real random variables satisfying $|U_i| \leq \varepsilon$. Let $\mu_i = \mathbb{E}[U_i]$ and let $\omega_i^2 = \mathrm{Var}[U_i]$. Let $\mu = \sum_{i=1}^n \mu_i$ and let $\omega^2 = \sum_{i=1}^n \omega_i^2$. Then*

$$\mathcal{W}_1(\sum_{i=1}^n U_i, N(\mu, \omega^2)) < 2\varepsilon.$$

TODO / NOTE - we need to check the constant 2 is good. Also find a good reference. I think we have this in our paper on the overlaps conjecture.

Lemma 3.8. *Suppose that λ_1 and λ_2 are probability measures on $\mathbb{R}$ and $r > 0$. Then*

$$|s_r(\lambda_1) - s_r(\lambda_2)| \leq e r^{-1} \mathcal{W}_1(\lambda_1, \lambda_2)$$

where e is Euler's number.

Proof. This is a special case of [KK25, Lemma 4.7]. Note that the proof of [KK25, Lemma 4.7] allows us to replace e with 2 though for simplicity we have not done that here. □

Lemma 3.9 (Small summands)**.** *Let $r > 0$, $0 < \varepsilon \leq 1$, and let $U_1, \ldots, U_n$ be independent real random variables satisfying $|U_i| \leq \varepsilon r$. Write $v = \sum_i \mathrm{Var}(U_i)$.*

Then

$$s_r\left(\sum_i U_i\right) \le \frac{r^2}{r^2+v} + 30\varepsilon.$$

Proof. Let $\mu = \sum_i \mathbb{E}[U_i]$. Let λ be the law of $N(\mu, v)$. Then

$$\begin{aligned} s_r\left(\sum_i U_i\right) &\le s_r(\lambda) + \left| s_r\left(\sum_i U_i\right) - s_r(\lambda)\right| \\ &\le \frac{r^2}{r^2+v} + er^{-1}\mathcal{W}_1(\lambda, \sum_i U_i) \\ &\le \frac{r^2}{r^2+v} + 6\varepsilon \end{aligned}$$

□

We can already prove (3.11). Let $\lambda \in (1/2, 1)$, $r \in (0,1)$ and $\epsilon = \chi_\lambda^{1/3} < 1$. Choose $n \ge 0$ with $\lambda\epsilon r < \lambda^n \le \epsilon r$ and define

$$Y_N = \sum_{j=n}^{N} \xi_j \lambda^j.$$

Since $1 - \lambda^2 \le 2\chi_\lambda$ and $\lambda^2 > 1/4$,

$$v_N := \mathrm{Var}(Y_N) = \frac{\lambda^{2n} - \lambda^{2N+2}}{1-\lambda^2} \to \frac{\lambda^{2n}}{1-\lambda^2} > \frac{\epsilon^2 r^2}{8\chi_\lambda} = \frac{r^2}{8\chi_\lambda^{1/3}}.$$

Applying Lemma 3.9 for all sufficiently large N we have

$$s_r(\nu_\lambda) \le s_r(Y_N) \le \frac{r^2}{r^2+v_N} + 30\chi_\lambda^{1/3} \le 38\chi_\lambda^{1/3} < 40\chi_\lambda^{1/3}.$$

When our summands are large we need to condition on small intervals in order to apply these tools. The next lemma is a conditional version of [KK25, Proposition 4.8], playing the role of [KK25, Proposition 6.6]. It is an easy consequence of (3.12) and Lemma 3.9.

Lemma 3.10 (Conditional variance)**.** *Put $\epsilon_0 = 10^{-4}$, let $r > 0$ and let $k \ge 1$ be an integer. Let $X_1, \dots, X_n$ be real random variables with finite second moments on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{A}_1, \dots, \mathcal{A}_n$ be sub-σ-algebras of $\mathcal{F}$. Suppose that the σ-algebras $\sigma(X_i) \vee \mathcal{A}_i$ are independent. Assume that there are $\mathcal{A}_i$-measurable real random variables C_i such that*

$$|X_i - C_i| \le \epsilon_0 r \quad \text{and} \quad \sum_i \mathbb{E}[\mathrm{Var}(X_i \mid \mathcal{A}_i)] \ge 1000kr^2.$$

Then

$$s_r^{(k)}\left(\sum_i X_i\right) \le 100^{-k}.$$

Proof. Put $W_i = \mathrm{Var}(X_i \mid \mathcal{A}_i)$. Note that the W_i are $\mathcal{A}_i$-measurable random variables. The W_i are independent and $0 \le W_i \le \mathbb{E}[(X_i - C_i)^2 \mid \mathcal{A}_i] \le \epsilon_0^2 r^2$. In particular

$$\mathbb{E}[W_i] \in [0, r^2] \quad \text{and} \quad \mathrm{Var}[W_i] \le \epsilon_0^2 r^2 \mathbb{E}[W_i]. \tag{3.13}$$

By (3.13) and a greedy algorithm we can partition $\{1,\dots,n\}$ into k sets $J_1,\dots,J_k$ such that

$$\sum_{j\in J_i}\mathbb{E}[W_j]\geq 999r^2$$

For one such J_i, put $m=\sum_{j\in J_i}\mathbb{E}[W_j]\geq 999r^2$. Since by (3.13) we have $\mathrm{Var}(\sum_{j\in J_i}W_j)\leq\epsilon_0^2r^2m$, Chebyshev's inequality gives

$$\mathbb{P}\left(\sum_{j\in J_i}W_j<m/2\right)\leq\frac{4\epsilon_0^2r^2}{m}\leq\frac{4\epsilon_0^2}{999}.$$

Condition on $\bigvee_{j\in J_i}\mathcal{A}_j$. The X_j remain conditionally independent, with conditional variances W_j, and the centres C_j are known. On the event $\sum_{j\in J_i}W_j\geq m/2$, Lemma 3.9 bounds the conditional detail by $2/999+10\epsilon_0$; on its complement use the bound 1. Convexity and translation invariance therefore give

$$s_r\left(\sum_{j\in J_i}X_j\right)\leq\frac{2}{999}+10\epsilon_0+\frac{4\epsilon_0^2}{999}<\frac{1}{100}.$$

The k group sums are independent. Apply (3.12) and convolve with the unused variables to obtain the claimed bound. □

We use Varjú's averaged entropy at scale [Var19, Section 2.2], though unlike Varjú we use natural logarithms instead of base 2. For a finitely supported real random variable X and $s>0$, define

$$H(X;s)=\int_0^s H(\lfloor\frac{X+t}{s}\rfloor)\,dt,$$

where $H(X)=-\sum_x\mathbb{P}(X=x)\log\mathbb{P}(X=x)$ is Shannon entropy. If T is independent of X and uniform on $[0,1]$, put

$$\mathcal{A}_s=\sigma(T,\lfloor X/s+T\rfloor),\qquad V(X;s)=s^{-2}\mathbb{E}[\mathrm{Var}(X\mid\mathcal{A}_s)].$$

The $\mathcal{A}_s$-measurable midpoint $C_s=s(\lfloor X/s+T\rfloor-T+1/2)$ satisfies $|X-C_s|\leq s/2$. Thus $V(X;s)$ is an additive analogue, with a fixed conditioning, of the trace in [KK25, Section 5.1].

An entropy increment forces conditional variance, as in [KK25, (5.2)]. Here the shifted grids give the following elementary bound.

Lemma 3.11 (Entropy gives variance)**.** *For every finitely supported X and $s>0$,*

$$H(X;s/4)-H(X;s)\leq 10V(X;s)^{1/4}.$$

This should be compared with [Kit26, Theorem 1.4]. In there the authors show that, with their notion of entropy at a scale $H_a(\cdot;r)$ which behaves in qualitatively a similar way to Varjú's entropy at a scale, we have

$$V(X;r)\gg a^{-2}(H_a(X;r/4a)-H_a(X;r/a)-O(e^{-a^2/4}+a^3r)).$$

Provided that we can substitute in $a=C\sqrt{\log(H_a(X;r/4a)-H_a(X;r/a))^{-1}}$ without effecting the entropy gap too badly this gives

$$V(X;r)\gg\frac{\tilde{\alpha}}{\log\tilde{\alpha}^{-1}}$$

where $\tilde{\alpha} = H_a(X; r/4a) - H_a(X; r/a)$ which, at least when the gap is small, is stronger than the estimate of

$$V(X; r) \geq \alpha^4/10$$

for $\alpha = H(X; r/4) - H(X; r)$ coming from Lemma 3.11. The estimate from Lemma 3.11 is still useful for our current purposes as it is simple and gives reasonable bounds at all scales.

Proof of Lemma 3.11. First let Y be finitely supported in an interval of length s. A shifted grid of mesh $s/4$ has at most five occupied cells. For any probability vector p on at most five points, the inequality $-\log u \leq 2(u^{-1/2} - 1)$ and Cauchy–Schwarz give

$$H(p) \leq 2 \sum_j \sqrt{p_j}(1 - p_j) \leq 4 \sqrt{1 - \sum_j p_j^2}.$$

Let Y' be an independent copy of Y. For a fixed grid, $1 - \sum_j p_j^2$ is the probability that Y, Y' occupy different cells. Averaging over a uniform shift, this probability becomes $\mathbb{E}[\min\{1, 4|Y - Y'|/s\}]$. Jensen's and Cauchy–Schwarz inequalities yield

$$H(Y; s/4) \leq 4 \left(\frac{4}{s} \mathbb{E}[|Y - Y'|] \right)^{1/2} \leq 8 \cdot 2^{1/4} \left(\frac{\mathrm{Var}(Y)}{s^2} \right)^{1/4} \leq 10 \left(\frac{\mathrm{Var}(Y)}{s^2} \right)^{1/4}.$$

Now take uniform shifts T, U, independent of each other and of X, for the coarse and fine grids. Writing $A = \lfloor X/s + T \rfloor$ and $Z = \lfloor 4X/s + U \rfloor$, the chain rule gives

$$H(X; s/4) - H(X; s) = H(Z \mid T, U) - H(A \mid T, U) \leq H(Z \mid A, T, U).$$

Conditionally on (A, T), the variable X lies in an interval of length s and U remains independent. Applying the preceding bound to this conditional law and then Jensen's inequality gives

$$H(Z \mid A, T, U) \leq 10 \mathbb{E} \left[\left(\frac{\mathrm{Var}(X \mid \mathcal{A}_s)}{s^2} \right)^{1/4} \right] \leq 10 V(X; s)^{1/4}.$$

This concludes the proof. □

We next record the arithmetic estimates needed to produce an entropy gap.

Lemma 3.12 (Arithmetic input)**.** *Let $\lambda \in (1/2, 1)$ be algebraic and let $\ell \geq 1$ be an integer. Write $X_\ell = \sum_{j=0}^{\ell-1} \xi_j \lambda^j$, where the ξ_j are independent balanced signs. Then*

$$B_\lambda \geq 3, \qquad \Delta_\ell \geq (\ell+1)^{1-d} M_\lambda^{-\ell} \geq e^{-B_\lambda \ell}, \qquad H(X_\ell) \geq \frac{\ell}{8B_\lambda}, \qquad \log \frac{1}{1-\lambda} \leq B_\lambda, \tag{3.14}$$

where Δ_ℓ is the minimum distance between distinct atoms.

This should be compared with Garsia's separation estimate (TODO) and Breuillard–Varjú [BV20, Theorem 5] give the stronger, degree-independent estimate $H(X_\ell) \geq c\ell \min\{1, \log M_\lambda\}$ for an absolute $c > 0$

Unlike for example [KK25], [Var19], [Kit24], [Kit25] (TODO maybe check that there aren't any other important ones) for the techniques of this paper we care about the strength of separation at small scales instead of just the asymptopic rate of exponential separation so to use (TODO eqref Garsia) we would need to estimate c_λ. Garsia's separation estimate has the disadvantage that the constant c_λ cannot be estimated just in terms of M_λ, d_λ and χ_λ (TODO - check this) meaning that we

would need to insert more quantities into Theorem 1.2. The estimates of Lemma 3.12 have the same asymptopic rate of exponential separation as Garsia's estimate though the exact separation predicted is slightly weaker for large ℓ.

The entropy estimate is weaker than that of Breuillard-Varjú due to the dependence on the degree d_λ.

Proof. The first assertion is immediate from the definition of B_λ. Write $d = d_\lambda$, $M = M_\lambda$ and let $P_\lambda(T) = a\prod_{i=1}^{d}(T-\lambda_i)$ be the primitive integer minimal polynomial, with $a > 0$ and $\lambda_1 = \lambda$. Suppose that P has coefficients in $\{-1, 0, 1\}$, degree $n < \ell$ and $P(\lambda) \neq 0$. Noting that the resultant of P and P_λ is a non zero integer we get

$$1 \leq a^n \prod_{i=1}^{d} |P(\lambda_i)|$$

and so from $|P(\lambda_i)| \leq (n+1)\max\{|\lambda_i|, 1\}^n$ we get

$$|P(\lambda)| \geq (n+1)^{1-d} M_\lambda^{-n}.$$

Differences of partial sums appearing in Bernoulli convolutions are twice such values, so $\Delta_\ell \geq 2\ell^{1-d} M^{-(\ell-1)} \geq e^{-B_\lambda \ell}$, where we have used $\log \ell \leq \ell$ and $B_\lambda \geq 1 + d + \log M$.

We prove the entropy bound by instead showing that if λ is an algebraic unit we can find some conjugate z and some sufficiently small k such that $|z|^k \geq 2$. This implies that $H(X_\ell) \geq \lceil \frac{\ell}{k} \rceil \log 2$.

If λ is not an algebraic unit then it is not a root of a polynomial with coefficients in $\{0, \pm 1\}$ and so each choice of signs gives a different partial sum and $H(X_\ell) = \ell \log 2 \geq$ TODO put bound here so we may assume that λ is an algebraic unit.

In this case we have

$$\log M = \sum_{i=1}^{d} \log \max\{1, |\lambda_i|\}$$

and in particular there is some conjugate z of λ with

$$\log |z| \geq \frac{\log M}{d}.$$

Let $k = \lceil \frac{d \log 2}{\log M} \rceil$. Then $k \log |z| \geq \log 2$ and so in particular $|z|^k \geq 2$. This gives

$$H(X_\ell) \geq \log 2 / \lceil \frac{d \log 2}{\log M} \rceil.$$

Finally, $P_\lambda(1)$ is a nonzero integer and $|P_\lambda(1)| \leq 2^{d-1} M(1-\lambda)$. Hence

$$\log \frac{1}{1-\lambda} \leq (d-1)\log 2 + \log M \leq B_\lambda,$$

which proves the last bound in (3.14). □

We now turn the entropy gap into variance over a range of scales, the counterpart for a fixed block of [KK25, Propositions 5.1 and 5.4].

Lemma 3.13 (Variance in a block)**.** *In the notation above, every integer $\ell \geq 100 B_\lambda^2$ satisfies*

$$\int_{e^{-2B_\lambda \ell}}^{e^{-\ell/(32B_\lambda)}} V(X_\ell; s)\, \frac{ds}{s} \geq \frac{\ell}{10^{10} B_\lambda^7}. \tag{3.15}$$

Proof. Put $a = e^{-2B_\lambda \ell}$ and $b = e^{-\ell/(32B_\lambda)}$. At scales $t \le a$, the atoms are resolved, so $H(X_\ell; t) = H(X_\ell) \ge \ell/(8B_\lambda)$. Since $|X_\ell| \le (1-\lambda)^{-1}$, a grid of mesh t meets its support in at most $2(1-\lambda)^{-1}t^{-1} + 2$ cells. Thus (3.14) gives, for $b/4 \le t \le b$,

$$H(X_\ell; t) \le \log\bigl(2(1-\lambda)^{-1}t^{-1} + 2\bigr) \le B_\lambda + \frac{\ell}{32B_\lambda} + \log 10 \le \frac{\ell}{16B_\lambda}.$$

The last inequality follows from $\ell \ge 100B_\lambda^2$ and $B_\lambda \ge 3$. Telescoping yields

$$\int_a^b \bigl(H(X_\ell; s/4) - H(X_\ell; s)\bigr)\frac{ds}{s} = \int_{a/4}^{a} H(X_\ell; t)\frac{dt}{t} - \int_{b/4}^{b} H(X_\ell; t)\frac{dt}{t} \ge \frac{(\log 4)\ell}{16B_\lambda}.$$

The interval has logarithmic length at most $2B_\lambda\ell$. Lemma 3.11 and Hölder's inequality therefore give

$$\int_a^b V(X_\ell; s)\,\frac{ds}{s} \ge \frac{((\log 4)\ell/(16B_\lambda))^4}{10^4(2B_\lambda\ell)^3} > \frac{\ell}{10^{10}B_\lambda^7}. \qquad \square$$

It remains to combine disjoint blocks of geometrically increasing lengths, as in [KK25, Proposition 6.7].

Proof of Proposition 3.6. We have already proved (3.11). For (3.10), write $B = B_\lambda$, $\chi = \chi_\lambda$, and fix $r = e^{-R}$ with $R \ge R_0 = 10^6 B^3$. Put $\epsilon = 10^{-4}$ and $N = \lceil 1000B^2 \rceil$. We will use block lengths $\ell = N, N^2, \dots$ not exceeding $R/(4B)$. Each length satisfies $\ell \ge 100B^2$, as required in Lemma 3.13.

First fix one such length ℓ, and set $I_\ell = [\ell/(32B), 2B\ell]$ and $h = \ell\chi$. Averaging arithmetic grids of spacing h gives

$$\frac{1}{h}\int_0^h \sum_{v \in (u+h\mathbb{Z}) \cap I_\ell} V(X_\ell; e^{-v})\,du = \frac{1}{h}\int_{I_\ell} V(X_\ell; e^{-v})\,dv \ge \frac{1}{10^{10}\chi B^7},$$

by Lemma 3.13. Choose a shift u_ℓ for which the sum is at least this average, and write $\mathcal{V}_\ell = (u_\ell + h\mathbb{Z}) \cap I_\ell$. For each $v \in \mathcal{V}_\ell$, set $s = e^{-v}$ and

$$n(v) = \left\lceil \frac{R + \log \epsilon^{-1} - v}{\chi} \right\rceil.$$

Since $v \le 2B\ell \le R/2$, we have $n(v) \ge 0$, and rounding gives

$$\lambda\epsilon r \le \lambda^{n(v)} s \le \epsilon r. \tag{3.16}$$

Use the digits starting at $n(v)$ to define

$$X_{\ell,v} = \sum_{j=0}^{\ell-1} \xi_{n(v)+j}\lambda^j, \qquad Y_{\ell,v} = \lambda^{n(v)} X_{\ell,v}.$$

The variable $X_{\ell,v}$ has the same law as X_ℓ. Consecutive points of $\mathcal{V}_\ell$ differ by $\ell\chi$, so their starting indices differ by exactly ℓ. Thus the blocks of this length use disjoint digits.

For each block, take a uniform shift $T_{\ell,v}$, with all shifts independent of one another and of the digits, and condition on

$$\mathcal{A}_{\ell,v} = \sigma(T_{\ell,v}, \lfloor X_{\ell,v}/s + T_{\ell,v} \rfloor).$$

The midpoint of the scaled interval gives an $\mathcal{A}_{\ell,v}$-measurable centre $C_{\ell,v}$ with $|Y_{\ell,v} - C_{\ell,v}| \le \lambda^{n(v)} s/2 \le \epsilon r$. Moreover, the definition of V and (3.16) give

$$\begin{aligned}\sum_{v\in\mathcal{V}_\ell} \mathbb{E}[\mathrm{Var}(Y_{\ell,v} \mid \mathcal{A}_{\ell,v})] &= \sum_{v\in\mathcal{V}_\ell} \lambda^{2n(v)} e^{-2v} V(X_\ell; e^{-v}) \\ &\ge \frac{\epsilon^2 r^2}{4\cdot 10^{10}\chi B^7},\end{aligned}$$

where we used $\lambda^2 > 1/4$.

We next check that blocks of different lengths also use disjoint digits. The assumption $\chi B^{10} \le 10^{-100}$ implies $\chi \le 1/(64B)$. If j belongs to the block starting at $n(v)$, the definition of $n(v)$ gives

$$v - \ell\chi < R + \log\epsilon^{-1} - j\chi \le v.$$

Consequently every digit in a block of length ℓ satisfies

$$R + \log\epsilon^{-1} - j\chi \in [\ell/(64B), 2B\ell].$$

These intervals are disjoint for distinct powers of N, since $N > 128B^2$. Thus all the blocks use disjoint digits, and the σ-algebras $\sigma(Y_{\ell,v}) \vee \mathcal{A}_{\ell,v}$ are independent.

Finally, $R \ge 10^6 B^3$ and $B \ge 3$ imply $R/(4B) \ge N$ and $\log(R/(4B)) \ge \frac{1}{2}\log R$. Using $\lfloor x \rfloor \ge x/2$ for $x \ge 1$, the number of admissible lengths is therefore

$$\left\lfloor \frac{\log(R/(4B))}{\log N} \right\rfloor \ge \frac{\log R}{4\log N}.$$

Since $N \le 1001B^2$ and $B \ge 3$, we also have $\log N \le B^3$. Summing the expected conditional variances over all blocks yields

$$\begin{aligned}r^{-2}\sum_{\ell}\sum_{v\in\mathcal{V}_\ell} \mathbb{E}[\mathrm{Var}(Y_{\ell,v} \mid \mathcal{A}_{\ell,v})] &\ge \frac{\epsilon^2 \log R}{16\cdot 10^{10}\chi B^7 \log N} \\ &\ge \frac{10^{82}}{16}\log R > 10^5 \log R,\end{aligned}$$

using $\epsilon = 10^{-4}$ and $\chi B^{10} \le 10^{-100}$. This exceeds $1001k$ whenever $1 \le k \le 10\log R$. Lemma 3.10 therefore bounds the order k detail of the sum of the selected blocks by 100^{-k}. The unused digits are independent of these blocks, so convolution with their sum proves (3.10). □

References

[DLS14] C. De Lellis and E. Spadaro, *Regularity of area minimizing currents I: gradient L^p estimates*, Geom. Funct. Anal. **24** (2014), no. 6, 1831–1884, DOI 10.1007/s00039-014-0306-3.

[DX10] A. Dubickas and Z. Xu, *Refinement equations and spline functions*, Adv. Comput. Math. **32** (2010), no. 1, 1–23, DOI 10.1007/s10444-008-9079-2.

[Gra14] L. Grafakos, *Classical Fourier analysis*, 3rd ed., Graduate Texts in Mathematics, vol. 249, Springer, New York, 2014. DOI 10.1007/978-1-4939-1194-3.

[GM86] M. Giaquinta and G. Modica, *Partial regularity of minimizers of quasiconvex integrals*, Ann. Inst. H. Poincaré Anal. Non Linéaire **3** (1986), no. 3, 185–208.

[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.

[KKM22] J. Kinnunen, E.-K. Kurki, and C. Mudarra, *Characterizations of weak reverse Hölder inequalities on metric measure spaces*, Math. Z. **301** (2022), 2269–2290, DOI 10.1007/s00209-022-02976-y.

[AHT17] T. C. Anderson, T. Hytönen, and O. Tapiola, *Weak A_∞ weights and weak reverse Hölder property in a space of homogeneous type*, J. Geom. Anal. **27** (2017), no. 1, 95–119, DOI 10.1007/s12220-015-9675-6.

[Mey63] N. G. Meyers, *An L^p-estimate for the gradient of solutions of second order elliptic divergence equations*, Ann. Scuola Norm. Sup. Pisa (3) **17** (1963), no. 3, 189–206.

[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.

[Sol95] B. Solomyak, *On the random series $\sum \pm\lambda^n$ (an Erdős problem)*, Ann. of Math. (2) **142** (1995), no. 3, 611–625, DOI 10.2307/2118556.

[SS16] P. Shmerkin and B. Solomyak, *Absolute continuity of self-similar measures, their projections and convolutions*, Trans. Amer. Math. Soc. **368** (2016), no. 7, 5125–5151, DOI 10.1090/tran6696.

[Spa12] E. Spadaro, *Nondoubling A_∞ weights*, Adv. Calc. Var. **5** (2012), no. 3, 345–354, DOI 10.1515/acv.2011.019.

[Tao11] T. Tao, *An introduction to measure theory*, Graduate Studies in Mathematics, vol. 126, American Mathematical Society, Providence, RI, 2011. DOI 10.1090/gsm/126.

[Var19] 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.

[Yu22] H. Yu, *Bernoulli convolutions with Garsia parameters in $(1, \sqrt{2}]$ have continuous density functions*, 2022. arXiv:2108.01008v2.

[BG06] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs, vol. 4, Cambridge University Press, Cambridge, 2006.

[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, available at arXiv:1510.04043.

[Kit24] S. Kittle, *Absolutely continuous self-similar measures with exponential separation*, Ann. Sci. Éc. Norm. Supér. (4) **57** (2024), no. 4, 1191–1231, DOI 10.24033/asens.2589, available at arXiv:2103.12684.

[Kit25] ———, *Absolutely continuous Furstenberg measures*, Proc. Lond. Math. Soc. (3) **131** (2025), no. 1, e70072, DOI 10.1112/plms.70072, available at arXiv:2305.05757.

[KK25] S. Kittle and C. Kogler, *On absolute continuity of inhomogeneous and contracting on average self-similar measures* (2025), available at arXiv:2409.18936v4.

[Kit26] S. a. K. Kittle Constantin, *Entropy Theory for Random Walks on Lie Groups*, Journal of Theoretical Probability **39** (2026), no. 13, DOI 10.1007/s10959-025-01468-z, available at arXiv:2501.05395.

[Kog26] C. Kogler, *On L^p density of self-similar measures and continuous Bernoulli convolutions*, Sep. 18, 2026. Recorded talk, Special Dynamics Learning Seminar, Institute for Advanced Study.

[KK] S. Kittle and C. Kogler, *Supplementary material for On L^p density of self-similar measures and Bernoulli convolutions*. GitHub repository, github.com/ckkogler/l1-to-lp-supplementary-material.

[Kog] C. Kogler, *L^p densities of self-similar measures and Bernoulli convolutions*. Lean formalization, GitHub repository, github.com/ckkogler/lpselfsimilar-lean.

[PSS00] Y. Peres, W. Schlag, and B. Solomyak, *Sixty years of Bernoulli convolutions*, Fractal geometry and stochastics II, Progress in Probability, vol. 46, Birkhäuser, Basel, 2000, pp. 39–65, DOI 10.1007/978-3-0348-8380-1_2.

[Var18] P. P. Varjú, *Recent progress on Bernoulli convolutions*, European Congress of Mathematics, European Mathematical Society, Zürich, 2018, pp. 847–867, DOI 10.4171/176-1/38, available at arXiv:1608.04210.

[Fal14] K. Falconer, *Fractal geometry: Mathematical foundations and applications*, 3rd ed., John Wiley & Sons, 2014.

[BSS23] B. Bárány, K. Simon, and B. Solomyak, *Self-similar and self-affine sets and measures*, Mathematical Surveys and Monographs, vol. 276, American Mathematical Society, Providence, RI, 2023. DOI 10.1090/surv/276.

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