DIMENSION OF HOMOGENEOUS ITERATED FUNCTION SYSTEMS WITH ALGEBRAIC TRANSLATIONS

DE-JUN FENG AND ZHOU FENG

ABSTRACT. Let μ be the self-similar measure associated with a homogeneous iterated function system $\Phi = \{\lambda x + t_j\}_{j=1}^m$ on $\mathbb{R}$ and a probability vector $(p_j)_{j=1}^m$, where $0 \neq \lambda \in (-1, 1)$ and $t_j \in \mathbb{R}$. Recently by modifying the arguments of Varjú in [28], Rapaport and Varjú [24] showed that if $t_1, \ldots, t_m$ are rational numbers and $0 < \lambda < 1$, then

$$\dim \mu = \min\left\{1, \ \frac{\sum_{j=1}^m p_j \log p_j}{\log |\lambda|}\right\}$$

unless Φ has exact overlaps. In this paper, we further show that the above result holds in the case when $t_1, \ldots, t_m$ are algebraic numbers and $0 < |\lambda| < 1$. This is done by adapting the ideas and extending the results of Breuillard, Rapaport and Varjú in [8, 9, 24, 28].

1. INTRODUCTION

1.1. **Background and our main result.** In this paper, by an *iterated function system* (IFS) we mean a finite family $\Phi = \{\varphi_j\}_{j=1}^m$ of contracting similarities on $\mathbb{R}$, taking the form

$$\varphi_j(x) = \lambda_j x + t_j,$$

where $0 \neq \lambda_j \in (-1, 1)$ and $t_j \in \mathbb{R}$ for $1 \leq j \leq m$. It is well known [17] that there exists a unique nonempty compact set $K \subset \mathbb{R}$ such that

$$K = \bigcup_{j=1}^m \varphi_j(K).$$

We call K the *self-similar set* generated by Φ. Let $p = (p_j)_{j=1}^m$ be a probability vector with strictly positive entries. By [17] there exists a unique Borel probability measure μ on $\mathbb{R}$ such that

$$\mu = \sum_{j=1}^m p_j \mu \circ \varphi_j^{-1}.$$

This μ is fully supported on K. We call μ the *self-similar measure* associated with Φ and p. It is known [13] that μ is exact dimensional, in the sense that the limit

$$\lim_{r \to 0} \frac{\log \mu([x - r, x + r])}{\log r}$$

2020 *Mathematics Subject Classification.* Primary 28A80, 42A85.

Key words and phrases. Self-similar measures, dimension of measures, exact overlaps conjecture, Mahler measure, entropy.

exists and is equal to a constant for μ-almost every x; we write $\dim \mu$ for this constant and call it the *dimension* of μ.

The dimension theory of self-similar measures is a central topic in fractal geometry. Let $\mu = \mu_{\Phi,p}$ be the self-similar measure associated with an IFS $\Phi = \{\varphi_j(x) = \lambda_j x + t_j\}_{j=1}^m$ on $\mathbb{R}$ and a probability vector p with strictly positive entries. There is a natural upper bound for $\dim \mu$ in terms of the entropy $H(p)$ and the Lyapunov exponent χ. That is, writing

$$H(p) := \sum_{j=1}^m -p_j \log p_j \quad \text{and} \quad \chi := -\sum_{j=1}^m p_j \log|\lambda_j|, \tag{1.1}$$

one has

$$\dim \mu \leq \min\left\{1, \frac{H(p)}{\chi}\right\}; \tag{1.2}$$

see e.g. [12, Corollary 5.2.3]. It turns out that equality in (1.2) holds in many cases, such as when the IFS Φ satisfies the open set condition (see e.g. [12, Theorem 5.2.5]). Here Φ is said to satisfy the *open set condition* if there is a nonempty open set $U \subset \mathbb{R}$ such that $\varphi_j(U)$, $1 \leq j \leq m$, are disjoint subsets of U. On the other hand, there are some circumstances in which dimension drop (that is, strict inequality in (1.2)) can occur. For $n \in \mathbb{N}$ and $J = j_1 \dots j_n \in \{1, \dots, m\}^n$, write for brevity

$$\varphi_J = \varphi_{j_1} \circ \cdots \circ \varphi_{j_n} \quad \text{and} \quad \lambda_J = \lambda_{j_1} \cdots \lambda_{j_n}.$$

The IFS Φ is said to have *exact overlaps* if there exist n and distinct $J_1, J_2 \in \{1, \dots, m\}^n$ such that $\varphi_{J_1} = \varphi_{J_2}$. It is not difficult to see that $\dim \mu < \min\{1, H(p)/\chi\}$ whenever Φ has exact overlaps and $\dim \mu < 1$. The following folklore conjecture, which is also called the *exact overlaps conjecture*, asserts that these are the only circumstances in which dimension drop can occur.

Conjecture 1.1. *Let μ be the self-similar measure associated with the IFS Φ on $\mathbb{R}$ and the probability vector p. If Φ has no exact overlaps, then*

$$\dim \mu = \min\left\{1, \frac{H(p)}{\chi}\right\}.$$

A version of this conjecture for the dimension of self-similar sets was first stated by Simon [26].

In recent years, significant progress has been made towards Conjecture 1.1. The first breakthrough was achieved by Hochman [14], who proved Conjecture 1.1 if the IFS Φ satisfies the exponential separation condition. To introduce this separation condition, for $n \in \mathbb{N}$ define

$$\Delta_n = \min\left\{|\varphi_{w_1}(0) - \varphi_{w_2}(0)| : \ w_1, w_2 \in \{1, \dots, m\}^n, w_1 \neq w_2 \text{ and } \lambda_{w_1} = \lambda_{w_2}\right\}.$$

The IFS Φ is said to satisfy the *exponential separation condition* if there exists $c > 0$ such that $\Delta_n \geq c^n$ for infinitely many n. The exponential separation condition is satisfied by broad families of ovelapping IFSs. For instance, it is satisfied by every

algebraic IFS Φ on $\mathbb{R}$ (i.e., all the parameters λ_j and t_j are algebraic numbers) that does not allow exact overlaps; see [14, Theorem 1.5].

As a far-reaching extension of Hochman's result, Rapaport [22] recently established the conjecture when only the contraction parameters λ_j are assumed to be algebraic. Another important breakthrough was made by Varjú [28], who showed that if μ is a Bernoulli convolution (that is, μ is the self-similar measure associated with the IFS $\Phi_\lambda = \{\lambda x, \lambda x + 1\}$ and the probability vector $p = (1/2, 1/2)$, where $1/2 < \lambda < 1$), then $\dim \mu = 1$ if λ is transcendental. Combined with Hochman's result regarding IFSs with algebraic parameters, this verifies Conjecture 1.1 for the family of Bernoulli convolutions.

Very recently, Rapaport and Varjú [24] investigated homogeneous IFSs of three maps. Among other things, they proved that for each $(\lambda, \tau) \in (2^{2/3}, 1) \times \mathbb{R}$, Conjecture 1.1 holds for the IFS $\Phi_{\lambda,\tau} = \{\lambda x,\ \lambda x + 1,\ \lambda x + \tau\}$ with equal probability weights. Moreover, by modifying the arguments of Varjú in [28], Rapaport and Varjú showed that Conjecture 1.1 holds for every homogeneous IFS $\Phi = \{\lambda x + t_j\}_{j=1}^m$ with contraction ratio λ and rational translations t_j (see [24, Theorem A.8]). More recently, Rapaport [23] proved a version of Conjecture 1.1 for self-conformal measures associated with exponentially separated real analytic IFSs on $\mathbb{R}$.

We remark that a stronger version of Conjecture 1.1 involving the L^q dimension (for $q \geq 1$) instead of dimension of measures was established by Shmerkin [25] under the exponential separation condition. Subsequently, the L^q dimension (for $0 < q < 1$) of self-similar measures on $\mathbb{R}$ was also determined by Barral and the first author [5] under the same separation condition. It is worth pointing out that there are some IFSs without exact overlaps for which the exponential separation condition fails; see [2, 4] and also [3, 10].

The main result of this paper is the following theorem, which verifies Conjecture 1.1 for homogeneous IFSs with algebraic translations. It is a natural generalization of [24, Theorem A.8].

Theorem 1.2. *Let μ be the self-similar measure associated with a homogeneous IFS $\Phi = \{\lambda x + t_j\}_{j=1}^m$ on $\mathbb{R}$ and a probability vector p, where $0 \neq \lambda \in (-1, 1)$. Assume that all the translations t_j are algebraic numbers. If Φ has no exact overlaps, then*

$$\dim \mu = \min\left\{1,\ \frac{H(p)}{-\log|\lambda|}\right\}.$$

1.2. **About the proof.** Our proof of Theorem 1.2 is adapted from Rapaport and Varjú [24, Appendix A] and Varjú [28]. Below we present two of the main ingredients used in the proof, which are the variants of the corresponding results in [8, 9, 24, 28] in our setting.

Throughout this subsection, let $\Phi = \{\lambda x + t_j\}_{j=1}^m$ be a fixed homogeneous IFS with $0 \neq \lambda \in (-1, 1)$ and t_j being algebraic numbers, and let $p = (p_j)_{j=1}^m$ be a fixed

probability vector. Write

$$\mathcal{D} = \{t_i - t_j \colon 1 \le i, j \le m\}. \tag{1.3}$$

For $n \in \mathbb{N}$, let $\mathcal{P}^n$ denote the set of polynomials in one variable of degree not exceeding n and with coefficients in $\mathcal{D}$. That is,

$$\mathcal{P}^n = \left\{ \sum_{i=0}^{n} a_i X^i \colon a_i \in \mathcal{D} \text{ for all } 0 \le i \le n \right\}. \tag{1.4}$$

For $\alpha > 0$ and $n \in \mathbb{N}$, write

$$E_\alpha^n = \{\eta \in (-1,1) \setminus \{0\} : \dim \mu_\eta < \alpha \text{ and } P(\eta) = 0 \text{ for some } 0 \neq P \in \mathcal{P}^n\},$$

where μ_η denotes the self-similar measure associated with the IFS $\{\eta x + t_j\}_{j=1}^m$ and the probability vector p.

The first main ingredient of our proof is the following theorem, which is a variant of [8, Theorem 1] and [24, Theorem A.2] for homogeneous IFSs with algebraic translations. It states roughly that if there is a dimension drop for μ_λ, then there are fast algebraic approximations (η_n) to λ for infinitely many n such that $\dim \mu_{\eta_n}$ also drops.

Theorem 1.3. *Suppose that* $\dim \mu_\lambda < \min\{1, -H(p)/\log|\lambda|\}$. *Then for every* $\varepsilon > 0$ *and* $n_0 \ge 1$, *there exist* $n \ge n_0$ *and* $\eta \in E^n_{\dim \mu_\lambda + \varepsilon}$ *such that*

$$|\lambda - \eta| \le \exp(-n^{1/\varepsilon}).$$

To present another ingredient of our proof, let us first introduce some notation and definitions. Let ν be an atomic measure defined by $\nu = \sum_{j=1}^m p_j \delta_{t_j}$. For $\eta \in \mathbb{C}$ with $0 < |\eta| < 1$, the *Garsia entropy* $h_{\eta,\nu}$ of η and ν is defined by

$$h_{\eta,\nu} = \lim_{n\to\infty} \frac{H(\sum_{j=0}^{n-1} \xi_j \eta^j)}{n}, \tag{1.5}$$

where $(\xi_n)_{n=0}^\infty$ is a sequence of independent random variables with common law ν, and $H(\cdot)$ stands for the Shannon entropy of a discrete random variable (see (3.1)). The above limit always exists, following from a subadditive argument.

Recall that if $f(X) = a_n \prod_{i=1}^n (X - \beta_i) = \sum_{i=0}^n a_i X^i \in \mathbb{Z}[X]$ is the minimal polynomial of an algebraic number β, then the *Mahler measure* of β is defined as

$$M(\beta) = |a_n| \prod_{i=1}^n \max\{1, |\beta_i|\}. \tag{1.6}$$

The following theorem is a variant of [28, Theorem 9] for homogeneous IFSs with algebraic translations, which gives a natural relation between the Garsia entropy and the Mahler measure. We remark that [28, Theorem 9] directly follows from [9, Proposition 13].

Theorem 1.4. *For any* $h \in (0, H(p))$, *there is a positive number* M *depending on* h, $t_1, \dots, t_m$ *and* $p_1, \dots, p_m$, *such that* $h_{\eta,\nu} \ge h$ *for every algebraic number* $\eta \in (-1,1)$ *with* $M(\eta) \ge M$.

By assuming Theorems 1.3 and 1.4, we can prove Theorem 1.2 by simply following the arguments in [24, Theorem A.8] with slight modifications. Below we briefly introduce the ideas for the proofs of Theorems 1.3 and 1.4. The details will be given in Sections 5 and 6.

To prove Theorem 1.3, we first establish a separation result (see Proposition 2.9), stating that if λ_1, λ_2 are distinct algebraic numbers each of which is a root of a polynomial in $\mathcal{P}^n$, then

$$|\lambda_1 - \lambda_2| \geq 2n^{-Mn} \quad \text{if } n \geq N,$$

where M, N are two positive constants depending on $t_1, \ldots, t_m$. This is a slight generalization of [24, Lemma 4.1] (which is due to Mahler [19], giving a similar lower bound $2n^{-5n}$ for the distance between roots of integer polynomials with bounded coefficients). Thanks to this separation result, Theorem 1.3 can be proved by following the exact proof of [24, Theorem A.2] except for some minor modifications of the involved auxiliary lemmas and propositions.

The proof of Theorem 1.4 is based on a lower bound estimate of the involved Garsia entropy. For a probability measure ν on $\mathbb{C}$ with bounded support, following [9] we define

$$\Phi_\nu(a) = \sup_{t>0} \{H(ta\xi + G) - H(t\xi + G)\} \quad \text{for } a \in \mathbb{R}, \tag{1.7}$$

where ξ is a complex random variable with law ν, G is a complex Gaussian random variable independent of ξ, with density $\exp(-|z|^2/2)/(2\pi)$ for $z \in \mathbb{C}$, and $H(\cdot)$ stands for differential entropy (see (3.2)). For an algebraic number η over a subfield $\mathbb{K}$ of $\mathbb{C}$, define

$$\widetilde{M}_{\mathbb{K}}(\eta) = \prod_{\widetilde{\eta}} \frac{1}{\min\{1, |\widetilde{\eta}|\}}, \tag{1.8}$$

where the product is taken over all the conjugates $\widetilde{\eta}$ (including η itself) of η over $\mathbb{K}$. Let $\mathbb{K}(\eta)$ denote the smallest field containing $\mathbb{K}$ and η, and $[\mathbb{K}_1 : \mathbb{K}_2]$ the degree of a field extension ([21]). The following proposition gives a lower bound of the Garsia entropy, which extends a result of Breuillard and Varjú [9, Proposition 13] from $\mathbb{Q}$ to $\mathbb{K}$.

Proposition 1.5. *Let $\mathbb{K}$ be a subfield of $\mathbb{C}$. Let $\eta \in \mathbb{C}$ such that $0 < |\eta| < 1$ and $[\mathbb{K}(\eta) : \mathbb{K}] < \infty$. Let $\{u_j\}_{j=1}^m \subset \mathbb{K}$ and $p = (p_j)_{j=1}^m$ be a probability vector. Set $\nu = \sum_{j=1}^m p_j \delta_{u_j}$. Then $h_{\eta,\nu} \geq \Phi_\nu\left(\widetilde{M}_{\mathbb{K}}(\eta)\right)$.*

The proof of the above proposition is adapted from the arguments in [9]. While in the rational case the proof of Theorem 1.4 can be completed from here, a further argument in needed in the algebraic case because the relationship of $\widetilde{M}_{\mathbb{K}}(\eta)$ and $M(\eta)$ is not a priori clear. For instance, if we let η be an algebraic integer with $0 < |\eta| < 1$ and take $\mathbb{K} = \mathbb{Q}(\eta)$, then $\widetilde{M}_{\mathbb{K}}(\eta) = 1/|\eta|$, which can be significantly smaller than $M(\eta)$.

To illustrate our argument, we may assume that η is a root of a polynomial in $\bigcup_{n=1}^{\infty}\mathcal{P}^n$; otherwise, by definition $h_{\eta,\nu}=H(p)$ and there is nothing to prove. Next we choose a real algebraic integer θ such that

$$\mathbb{Q}(\theta)=\mathbb{Q}(\mathcal{D}),$$

where for $Z\subset\mathbb{C}$, $\mathbb{Q}(Z)$ stands for the smallest subfield of $\mathbb{C}$ containing both $\mathbb{Q}$ and Z. The existence of such θ follows from Lemma 2.6. Set $d=\deg(\theta)$ and let θ_i, $i=1,\dots,d$, be the algebraic conjugates of θ over $\mathbb{Q}$, including θ itself. We manage to show that there is a constant C (depending only on $\mathcal{D}$ and θ) such that

$$\prod_{i=1}^{d}\widetilde{M}_{\mathbb{Q}(\theta_i)}(\eta_i)\geq CM(\eta), \tag{1.9}$$

where η_i $(i=1,\dots,d)$ are some suitably chosen algebraic numbers with

$$[\mathbb{Q}(\theta_i,\eta_i):\mathbb{Q}(\theta_i)]<\infty,$$

and $\widetilde{M}_{\mathbb{Q}(\theta_i)}(\eta_i)$ are defined as

$$\widetilde{M}_{\mathbb{Q}(\theta_i)}(\eta_i)=\prod_{j}\frac{1}{\min\{1,|\beta_j|\}},$$

where the product is taken over all algebraic conjugates β_j of η_i over the field $\mathbb{Q}(\theta_i)$, including η_i itself. Due to (1.9), we may choose $k_0\in\{1,\dots,d\}$ such that

$$\widetilde{M}_{\mathbb{Q}(\theta_{k_0})}(\eta_{k_0})\geq C^{1/d}M(\eta)^{1/d}. \tag{1.10}$$

Assume that $M(\eta)$ is large enough (saying, greater than $1/C$). Then the left hand side of (1.10) is larger than 1, so replacing η_{k_0} by one of its algebraic conjugates over $\mathbb{Q}(\theta_{k_0})$ if necessary, we may assume that

$$|\eta_{k_0}|<1.$$

Notice that η_{k_0} might take value in $\mathbb{C}\setminus\mathbb{R}$.

To conclude Theorem 1.4, we apply Proposition 1.5 to derive that

$$h_{\eta,\nu}\geq\Phi_{\nu'}\left(\widetilde{M}_{\mathbb{Q}(\theta_{k_0})}(\eta_{k_0})\right)$$

for a certain atomic probability measure ν' on $\mathbb{C}$, where $\Phi_{\nu'}$ is defined as in (1.7) in which ν is replaced by ν'. Since $\Phi_{\nu'}(a)$ tends to $H(p)$ as $a\to\infty$, the above inequality and (1.10) yield Theorem 1.4.

We actually prove a more general version of Theorem 1.4 in which $t_1,\dots,t_m$ and η are assumed to be complex algebraic numbers (see Theorem 6.1), which might be helpful for the study of homogeneous self-similar measures on $\mathbb{R}^2$ or higher dimensional spaces.

1.3. **Structure of the article.** In Section 2 we collect some results about roots of polynomials with integer or algebraic coefficients, which we will use at various places in the paper, in particular in the proofs of Theorems 1.2 and 1.3. In Section 3, we review two notions of entropy and present their basic properties. In Section 4 we prove Theorem 1.2 by assuming Theorems 1.3 and 1.4. The proofs of Theorems 1.3 and 1.4 are respectively given in Sections 5 and 6.

2. Distribution of roots of polynomials

In this section, we collect some properties about roots of polynomials with coefficients in a given finite set of algebraic numbers, which we will mainly use in the proofs of Theorems 1.2 and 1.3. Many of them are standard, and the others are the extensions of the corresponding results in [8, 24] about integer polynomials.

Throughout this section, let $\mathcal{D}$ be a fixed finite set of algebraic numbers in $\mathbb{C}$. For $n \geq 1$, let $\mathcal{P}^n$ denote the collection of polynomials in X of degree not exceeding n and with coefficients in $\mathcal{D}$. That is,

$$\mathcal{P}^n = \left\{ \sum_{i=0}^{n} a_i X^i \colon a_i \in \mathcal{D} \text{ for all } 0 \leq i \leq n \right\}. \tag{2.1}$$

2.1. **Polynomials, Mahler measure and height.** We first introduce some standard notation and definitions. For $\mathcal{S} \subset \mathbb{C}$, by $\mathcal{S}[X]$ and $\mathcal{S}[X_1, \ldots, X_k]$ we respectively denote the sets of the univariate polynomials in variable X and the multivariate polynomials in variables $X_1, \ldots, X_k$ with coefficients in $\mathcal{S}$.

For $\alpha, \alpha_1, \ldots, \alpha_k \in \mathbb{C}$, we write

$$\mathcal{S}[\alpha] = \{P(\alpha) \colon P \in \mathcal{S}[X]\}$$

and

$$\mathcal{S}[\alpha_1, \ldots, \alpha_k] = \{P(\alpha_1, \ldots, \alpha_k) \colon P \in \mathcal{S}[X_1, \ldots, X_k]\},$$

where $P(\alpha)$ or $P(\alpha_1, \ldots, \alpha_k)$ denotes the evaluation of a polynomial P at α or $(\alpha_1, \ldots, \alpha_k)$. Write $\mathcal{S}[\mathcal{A}] = \mathcal{S}[\alpha_1, \ldots, \alpha_k]$ if $\mathcal{A} = \{\alpha_1, \ldots, \alpha_k\}$. We adapt the notation with $\mathcal{S}$ replaced by $\mathbb{Q}, \mathbb{Z}$, $\mathbb{C}$ or any field $\mathbb{K} \subset \mathbb{C}$. For $\mathcal{A} \subset \mathbb{C}$, let $\mathbb{Q}(\mathcal{A})$ denote the smallest field containing $\mathbb{Q}$ and $\mathcal{A}$, and let $\mathbb{Q}(\alpha)$ be the smallest field containing $\mathbb{Q}$ and $\alpha \in \mathbb{C}$.

For $f = \sum_{i=0}^{n} a_i X^i \in \mathbb{C}[X]$, the degree of f is denoted as $\deg f$, and we write

$$\ell_q(f) = \left(\sum_{i=0}^{n} |a_i|^q \right)^{1/q} \quad \text{for } 1 \leq q < \infty, \qquad \text{and} \quad \ell_\infty(f) = \max_{0 \leq i \leq n} |a_i|.$$

Similar notations are used for multivariate polynomials.

Definition 2.1 (Mahler measure)**.** The *Mahler measure* of $f = \sum_{i=0}^{n} a_i X^i \in \mathbb{C}[X]$ is defined as

$$M(f) := |a_n| \prod_{i=1}^{n} \max\{1, |\alpha_i|\}, \tag{2.2}$$

where $\alpha_1, \ldots, \alpha_n$ are the roots of f. For an algebraic number $\alpha \in \mathbb{C}$, the Mahler measure of α is defined by $M(\alpha) := M(g)$, where g is the minimal polynomial of α in $\mathbb{Z}[X]$ with coprime coefficients.

By the above definition, $M(\alpha) \geq 1$ for every algebraic number α. Following [20, Chapter 14], we define the *height* of an algebraic number $\alpha \in \mathbb{C}$ as

$$H(\alpha) = M(\alpha)^{1/\deg \alpha}, \tag{2.3}$$

where $\deg \alpha$ stands for the degree of α over $\mathbb{Q}$. Clearly $H(\alpha) \geq 1$.

It is immediate from (2.2) that

$$M(fg) = M(f)M(g) \quad \text{for } f, g \in \mathbb{C}[X].$$

Below we list more properties about the Mahler measure and height.

Lemma 2.2 ([7, Lemma 1.6.7])**.** *Let $f \in \mathbb{C}[X]$ with $\deg f = n$. Then $M(f) \leq \ell_1(f)$. Moreover,*

$$\binom{n}{\lfloor n/2 \rfloor}^{-1} \ell_\infty(f) \leq M(f) \leq \ell_2(f) \leq \sqrt{n+1}\, \ell_\infty(f),$$

where $\lfloor x \rfloor$ stands for the integer part of x.

Lemma 2.3 ([18, Equation 4])**.** *Let $f \in \mathbb{C}[X]$ with $\deg f = n$. Then*

$$\ell_1(f) \leq 2^n M(f).$$

Lemma 2.4 ([20, Proposition 14.7])**.** *Let $f \in \mathbb{Z}[X_1, \ldots, X_k]$ such that the degree of f in X_j is at most L_j for $1 \leq j \leq k$. Let $\alpha_1, \ldots, \alpha_k$ be algebraic numbers. Then*

$$H(f(\alpha_1, \ldots, \alpha_k)) \leq \ell_1(f) \prod_{j=1}^{k} H(\alpha_j)^{L_j}.$$

Let $[\mathbb{K}_1 : \mathbb{K}_2]$ denote the degree of a field extension $\mathbb{K}_1/\mathbb{K}_2$, that is, the dimension of $\mathbb{K}_1$ as a vector space over $\mathbb{K}_2$ ([21]).

Lemma 2.5 ([20, Proposition 14.13])**.** *Let $\mathbb{K} \subset \mathbb{C}$ be an extension field of $\mathbb{Q}$ with $k = [\mathbb{K} : \mathbb{Q}] < \infty$. If $\alpha \neq 0$ is in $\mathbb{K}$, then $|\alpha| \geq H(\alpha)^{-k}$.*

Lemma 2.6 ([21, Corollary 5.8])**.** *Let $\alpha_1, \ldots, \alpha_k$ be algebraic numbers. Then there exists an algebraic integer θ such that $\mathbb{Q}(\alpha_1, \ldots, \alpha_k) = \mathbb{Q}(\theta)$.*

2.2. **The distribution of roots of polynomials with coefficients in $\mathcal{D}$.** Recall that $\mathcal{D}$ is a fixed finite set of algebraic numbers, and $\mathcal{P}^n$, $n \in \mathbb{N}$, are defined as in (2.1).

The following proposition is a slight extension of [24, Lemma 4.2] (see also, [6] and [8, Lemma 26]).

Proposition 2.7. *For $\varepsilon > 0$, there exists $k = k(\varepsilon, \mathcal{D}) > 0$ such that every polynomial P in $\bigcup_{n=1}^{\infty} \mathcal{P}^n$ has at most k nonzero roots of modulus at most $1 - \varepsilon$.*

Proof. The proposition was proved in [24, Lemma 4.2] in the case when $\mathcal{D}$ is a finite set of integers. Here we follow that proof with slight modifications. Let $\varepsilon > 0$. Set

$$\rho = \frac{\max\{|s| : \ 0 \neq s \in \mathcal{D}\}}{\min\{|s| : \ 0 \neq s \in \mathcal{D}\}}.$$

Define

$$a(j) = \frac{j}{j+1} \cdot \frac{1}{[\rho(j+1)]^{1/j}} \quad \text{for } j \in \mathbb{N}.$$

Then $a(j) \to 1$ as $j \to \infty$. So we may pick a large integer k such that $a(k) > 1 - \varepsilon/2$.

Let $P \in \mathcal{P}^n$ for some $n \in \mathbb{N}$, and let $z_1, \ldots, z_N$ be the nonzero roots, repeated according to multiplicity, of P in the open disc $|z| < 1 - \varepsilon$. Below we prove that $N \leq k$.

Set $u = \min\{|s| : \ 0 \neq s \in \mathcal{D}\}$ and

$$Q = \frac{P}{uX^m},$$

where m is the lowest degree among the monomials appearing in P. Then $|Q(0)| \geq 1$ and $\ell_\infty(Q) \leq \rho$. Write $r = k/(k+1)$. Then $1 - \varepsilon/2 < a(k) < r$. By Jensen's formula and since $|Q(0)| \geq 1$,

$$\sum_{j=1}^{N} \log \frac{r}{|z_j|} \leq \int_0^1 \log |Q(re^{2\pi i t})| dt. \tag{2.4}$$

Note that for each $t \in [0, 1]$,

$$|Q(re^{2\pi i t})| \leq \rho(1 + r + r^2 + \cdots) = \frac{\rho}{1-r} = \rho(k+1).$$

Since $|z_j| \leq a(k)$ for $1 \leq j \leq N$, by (2.4),

$$N \log \left((\rho(k+1))^{1/k}\right) \leq \log(\rho(k+1)).$$

This concludes $N \leq k$. □

As a consequence of the above proposition, we give the following corollary which is an analogue of [24, Lemma 4.3] in our setting.

Corollary 2.8. *For any $\varepsilon \in (0, 1/2)$, there exists $c = c(\varepsilon) \in (0, 1)$ such that the following holds. Let $n \in \mathbb{N}$, $0 \neq P \in \mathcal{P}^n$ and $0 < r < u\varepsilon^n 2^{-n}$, where*

$$u = \min\{|s| : \ 0 \neq s \in \mathcal{D}\}.$$

Suppose that $|P(\lambda)| \le r$ *for some* $\lambda \in \mathbb{C}$ *with* $\varepsilon \le |\lambda| \le 1-\varepsilon$. *Then there exists* $\eta \in \mathbb{C}$ *such that* $P(\eta) = 0$ *and*

$$|\lambda - \eta| \le (u 2^n \epsilon^{-n} r)^c.$$

Proof. The proof is nearly identical to that of [24, Lemma 4.3]. For the reader's convenience, we include the details.

Let $\varepsilon \in (0, 1/2)$, and let n, P, r be given as in the statement of the corollary. By Proposition 2.7, there exists an integer $k = k(\epsilon, \mathcal{D})$ (independent of P) such that P has at most k nonzero roots of modulus at most $1 - \varepsilon/2$ (repeated according to multiplicity). Denote these roots by $\eta_1, \dots, \eta_m$. Then $m \le k$ and

$$r \ge |P(\lambda)| \ge u(\varepsilon/2)^{n-m} \cdot \prod_{j=1}^{m} |\eta_j - \lambda|.$$

Hence there exists some $1 \le j \le m$ such that

$$|\eta_j - \lambda| \le (u^{-1}(\varepsilon/2)^{-n} r)^{1/m} \le (u^{-1} 2^n \epsilon^{-n} r)^{1/k}.$$

This completes the proof of the corollary by taking $c = 1/k$. □

2.3. **A separation property of roots of polynomials in $\mathcal{P}^n$.** The main result in the subsection is the following proposition, which is an analogue of [24, Lemma 4.1].

Proposition 2.9. *There exists* $M = M(\mathcal{D}) > 0$ *such that the following holds for all sufficiently large* n. *Let* $\lambda_1 \ne \lambda_2$ *be two algebraic numbers each of which is a root of a polynomial in* $\mathcal{P}^n$. *Then* $|\lambda_1 - \lambda_2| > 2n^{-Mn}$.

To prove the above proposition, we start with a result essentially due to Mahler [19]. For $n \in \mathbb{N}$ and $a > 0$, write

$$\mathcal{F}(n, a) := \{P \in \mathbb{Z}[X] : \ell_\infty(P) \le a \text{ and } \deg P \le n\}. \tag{2.5}$$

Lemma 2.10. *Let* $n \ge 4$, $a > 0$, *and let* λ_1, λ_2 *be two distinct algebraic numbers each of which is a root of a polynomial in* $\mathcal{F}(n, a)$. *Then*

$$|\lambda_1 - \lambda_2| > 2n^{-4n} a^{-4n+2}.$$

Proof. We follow the proof of [24, Lemma 4.1] with minor modifications. Let $n, a, \lambda_1, \lambda_2$ be given as in the statement of the lemma. Suppose $P_1(\lambda_1) = 0$ and $P_2(\lambda_2) = 0$ for some $P_1, P_2 \in \mathcal{F}(n, a)$. Let $P_1 = R_1^{k_1} \cdots R_s^{k_s}$ and $P_2 = Q_1^{h_1} \cdots Q_t^{h_t}$ be respectively the irreducible factorizations of P_1 and P_2 in $\mathbb{Z}[X]$.

Without loss of generality assume that $R_1(\lambda_1) = 0$ and $Q_1(\lambda_2) = 0$. If λ_1 and λ_2 are Galois conjugates, then take $P = R_1$ or Q_1, otherwise, take $P = R_1 Q_1$. In both cases, $\deg P \le 2n$, and $M(P) \le M(P_1)M(P_2) \le (n+1)a^2$ by Lemma 2.2. Hence applying [19, Theorem 2] to P gives

$$\begin{aligned} |\lambda_1 - \lambda_2| &> \sqrt{3}(2n)^{-(n+1)}((n+1)a^2)^{-2n+1} \\ &> \sqrt{3}\, 2^{-3n} n^{-3n} a^{-4n+2} \end{aligned}$$

$$> 2n^{-4n}a^{-4n+2},$$

where we use the assumption $n \geq 4$ in the last inequality. □

To apply Lemma 2.10, we need the following result.

Lemma 2.11. *Let $\lambda \in \mathbb{C}$. If λ is a root of a polynomial $f \in \mathcal{P}^n$ for some $n \geq 1$, then λ is a root of a polynomial F with integer coefficients bounded by $C(n+1)^d)$ and degree at most dn, where C is a constant depending on $\mathcal{D}$ and $d = [\mathbb{Q}(\mathcal{D}) : \mathbb{Q}]$.*

Proof. By Lemma 2.6, we can choose an algebraic integer θ such that $\mathbb{Q}(\theta) = \mathbb{Q}(\mathcal{D})$. Then we can take $L \in \mathbb{N}$ such that

$$Ls \in \mathbb{Z}[\theta] \quad \text{for all } s \in \mathcal{D}. \tag{2.6}$$

Notice that the degree of $\mathbb{Q}(\mathcal{D})$ over $\mathbb{Q}$ is d, so $\deg(\theta) = d$. Let $\theta_1, \dots, \theta_d$ be the algebraic conjugates of θ, with $\theta_1 = \theta$. For $j = 1, \dots, d$, let

$$\sigma_j : \mathbb{Q}(\theta) \to \mathbb{Q}(\theta_j)$$

be the field isomorphism such that $\sigma_j(a) = a$ for all $a \in \mathbb{Q}$ and $\sigma_j(\theta) = \theta_j$.

Let $\lambda \in \mathbb{C}$ and suppose that $P(\lambda) = 0$ for some $P = \sum_{i=0}^n a_i X^i \in \mathcal{P}^n$. Define $F \in \mathbb{C}[X]$ by

$$F = L^d \prod_{j=1}^d \left(\sum_{i=0}^n \sigma_j(a_i) X^i \right) = \prod_{j=1}^d \left(\sum_{i=0}^n \sigma_j(La_i) X^i \right).$$

Clearly $F(\lambda) = 0$, $\deg(F) \leq dn$ and

$$\ell_\infty(F) \leq \ell_1(F) \leq L^d \prod_{j=1}^d \left(\sum_{i=0}^n |\sigma_j(a_i)| \right) \leq C(n+1)^d,$$

where $C := L^d \max\{|\sigma_j(s)|^d : \ 1 \leq j \leq d, \ s \in \mathcal{D}\}$.

To conclude the lemma, it remains to show that $F \in \mathbb{Z}[X]$. By (2.6), $La_i \in \mathbb{Z}[\theta]$ for all $0 \leq i \leq n$. Hence there exist integer polynomials Q_i, $i = 0, \dots, n$, such that $La_i = Q_i(\theta)$. It follows that for $0 \leq i \leq n$ and $1 \leq j \leq d$,

$$\sigma_j(La_i) = \sigma_j(Q_i(\theta)) = Q_i(\theta_j).$$

Therefore

$$F = \prod_{j=1}^d \left(\sum_{i=0}^n Q_i(\theta_j) X^i \right) = \sum_{k=0}^{dn} \beta_k X^k.$$

Notice that each β_k is a symmetric polynomial in $\theta_1, \dots, \theta_d$ with integer coefficients. Since θ is an algebraic integer, it follows from the fundamental theorem of symmetric polynomials (see e.g. [1, Ch. 14, Theorem (3.4)]) and Vieta's formulas that $\beta_k \in \mathbb{Z}$, and so $F \in \mathbb{Z}[X]$. □

Now we are ready to prove Proposition 2.9.

Proof of Proposition 2.9. Let d, C be the constants as in Lemma 2.11. A combination of Lemmas 2.10 and 2.11 gives

$$|\lambda_1 - \lambda_2| > 2(dn)^{-4dn}[(C(n+1))^d]^{-4dn+2} > 2n^{-(4d^2+4d+1)n},$$

for n sufficiently large depending on $\mathcal{D}$. Taking $M = 4d^2+4d+1$ finishes the proof. □

2.4. **The order of zeros of polynomials under certain constrains.** The following proposition is an analogue of [24, Lemma 2.3] in our setting.

Proposition 2.12. *Let $\varepsilon > 0$ and $k \in \mathbb{N}$. There exists a positive integer $N = N(\epsilon, k, \mathcal{D})$ such that the following holds. Let $n \in \mathbb{N}$ with $n \geq N$, and $\lambda, \eta \in (-1, 1)$ such that η is algebraic with $\deg \eta \leq n$ and $0 < |\lambda - \eta| \leq 1$. Let $n' \in \mathbb{N}$ such that $n' \geq (k+2)n^{1+\varepsilon}$. Let $0 \neq P \in \mathcal{P}^{n'}$. Suppose that*

$$(2M(\eta))^{dn'/k}|P(\lambda)|^{1/k} \leq |\lambda - \eta| \leq (2M(\eta))^{-dn'}, \tag{2.7}$$

where $d := [\mathbb{Q}(\mathcal{D}) : \mathbb{Q}]$. Then η is a zero of P of order at least k.

The proof of Proposition 2.12 relies on the following result which is contained in the proof of [24, Lemma 4.6]. As noted by Rapaport and Varjú [24], this result was communicated privately by V. Dimitrov.

Lemma 2.13. *Let λ, η be distinct nonzero numbers in $(-1, 1)$. Let $0 \neq P \in \mathcal{P}^{n'}$ for some $n' \in \mathbb{N}$. Let m be the order of vanishing of P at η, where we allow $m = 0$. Then*

$$\left|\frac{P^{(m)}(\eta)}{m!}\right| \leq 2(n')^{m+2}D|\lambda - \eta| + \frac{2|P(\lambda)|}{|\lambda - \eta|^m}, \tag{2.8}$$

where $P^{(i)}$ denotes the i-th derivative of P for $i \in \mathbb{N}$, and $D := \max\{|s| : s \in \mathcal{D}\}$.

Proof. We follow the proof of [24, Lemma 4.6]. Write $P = \sum_{j=0}^{n'} a_j X^j$ with $a_j \in \mathcal{D}$. Set

$$P_1 = \sum_{j=0}^{n'} \mathrm{Re}(a_j)X^j, \qquad P_2 = \sum_{j=0}^{n'} \mathrm{Im}(a_j)X^j,$$

where $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ stand for the real and imaginary parts of a complex number z. Then $P = P_1 + iP_2$. Below we prove that

$$\left|\frac{P_k^{(m)}(\eta)}{m!}\right| \leq (n')^{m+2}D|\lambda - \eta| + \frac{|P_k(\lambda)|}{|\lambda - \eta|^m}, \quad k = 1, 2. \tag{2.9}$$

Clearly (2.8) follows from (2.9).

Without loss of generality, we only prove (2.9) for $k = 1$. To this end, denote

$$Q = \frac{P_1^{(m)}}{m!} = \sum_{j=m}^{n'} \binom{j}{m} \mathrm{Re}(a_j)X^{j-m}. \tag{2.10}$$

Since η is vanishing of order m, by Taylor's theorem with Lagrange remainder term, there exists some ξ between λ and η such that

$$P_1(\lambda) = Q(\xi)(\lambda - \eta)^m. \tag{2.11}$$

If $m = 0$, this holds with $\xi = \lambda$.

It follows from (2.10) that $\deg Q \le n'$ and $\ell_\infty(Q) \le (n')^m D$. Then

$$|Q^{(1)}(x)| \le (n')^2 \ell_\infty(Q) \le (n')^{m+2} D \quad \text{for } x \in (-1, 1).$$

Hence by the mean value theorem,

$$|Q(\eta)| \le (n')^{m+2} D|\xi - \eta| + |Q(\xi)|.$$

From this, $|\xi - \eta| \le |\lambda - \eta|$ and (2.11), we obtain that

$$|Q(\eta)| \le (n')^{m+2} D|\lambda - \eta| + \frac{|P_1(\lambda)|}{|\lambda - \eta|^m},$$

which proves (2.9) for $k = 1$. □

Now we are ready to prove Proposition 2.12.

Proof of Proposition 2.12. Our argument is adapted from the proof of [24, Lemma 4.6]. Let $\varepsilon, k, \lambda, \eta, n', P$ be given as in the statement of the proposition. Let m be the order of vanishing of P at η. We show below that $m \ge k$.

Write

$$Q = \frac{P^{(m)}}{m!} = \sum_{j=m}^{n'} \binom{j}{m} a_j X^{j-m}. \tag{2.12}$$

Then $Q(\eta) \ne 0$. From (2.12) we see that

$$Q(\eta) = \widetilde{Q}(s_1, \dots, s_r, \eta)$$

for an integer multivariate polynomial $\widetilde{Q}$ with $\ell_1(\widetilde{Q}) \le (n'+1)(n')^m$, where $s_1, \dots, s_r$ are the elements in $\mathcal{D}\backslash\{0\}$. By Lemma 2.4, the height of $\widetilde{Q}(\eta)$ (cf. (2.3)) satisfies

$$\begin{aligned} H(Q(\eta)) &= H\left(\widetilde{Q}(s_1, \dots, s_r, \eta)\right) \\ &\le \ell_1(\widetilde{Q}) \left(\prod_{s \in \mathcal{D}\backslash\{0\}} H(s)\right) \cdot H(\eta)^{n'} \\ &\le (n'+1)(n')^m C_2 M(\eta)^{n'/\deg \eta}, \end{aligned} \tag{2.13}$$

where $C_2 := \prod_{s \in \mathcal{D}\backslash\{0\}} H(s)$.

Note that by (2.10), $Q(\eta) \in \mathbb{Q}(\mathcal{D}, \eta)$. It follows from [21, Propositions 1.15 and 1.21] that

$$\deg Q(\eta) = [\mathbb{Q}(Q(\eta)) : \mathbb{Q}] \le [\mathbb{Q}(\mathcal{D}, \eta) : \mathbb{Q}] \le [\mathbb{Q}(\mathcal{D}) : \mathbb{Q}][\mathbb{Q}(\eta) : \mathbb{Q}] = d \deg(\eta), \tag{2.14}$$

where $d := [\mathbb{Q}(\mathcal{D}) : \mathbb{Q}]$. Since $Q(\eta) \ne 0$, $\deg(\eta) \le n$ and $H(Q(\eta)) \ge 1$, it follows from Lemma 2.5 , (2.14) and (2.13) that

$$|Q(\eta)| \ge H(Q(\eta))^{-\deg Q(\eta)} \ge H(Q(\eta))^{-d \deg \eta} \ge ((n'+1)(n')^m C_2)^{-dn} M(\eta)^{-dn'}. \tag{2.15}$$

Suppose on the contrary that $m < k$. Set $D = \max\{|s| : s \in \mathcal{D}\}$. Then by Lemma 2.13 and $|\lambda - \eta| \le 1$,

$$|Q(\eta)| = \frac{|P^{(m)}(\eta)|}{m!} \le 2(n')^{m+2} D|\lambda - \eta| + \frac{2|P(\lambda)|}{|\lambda - \eta|^m} \le 2(n')^{k+2} D|\lambda - \eta| + \frac{2|P(\lambda)|}{|\lambda - \eta|^k}.$$

Hence

$$\begin{aligned} &\left((n'+1)(n')^k C_2\right)^{-dn} M(\eta)^{-dn'} \\ &\quad \le |Q(\eta)| && \text{(by (2.15) and } m < k) \\ &\quad \le 2(n')^{k+2} D|\lambda - \eta| + \frac{2|P(\lambda)|}{|\lambda - \eta|^k} \\ &\quad \le 2\left((n')^{k+2} D + 1\right)(2M(\eta))^{-dn'} && \text{(by (2.7)),} \end{aligned}$$

which is simplified to

$$\left((n')^{k+2} D + 1\right)\left((n'+1)(n')^k C_2\right)^{dn} \ge 2^{dn'-1}.$$

This contradicts the assumption that $n' \ge (k+2)n^{1+\varepsilon}$, when n is large enough depending on $\mathcal{D}$, k and ε. □

3. Preliminaries on Shannon and differential entropies

In this section, we briefly review two notions of entropy, and present some of their properties which will be used in the later sections. For more background material on entropy, the reader is referred to [9, 11].

Let X be a random variable in $\mathbb{R}^d$. If X is a discrete random variable taking values in a countable set $\{x_j\}_j$, the *Shannon entropy* of X is defined as

$$H(X) := \sum_j -\mathbb{P}\{X = x_j\} \log \mathbb{P}\{X = x_j\}, \tag{3.1}$$

with convention $0 \log 0 = 0$. If X is an absolutely continuous random variable with density $f\colon \mathbb{R}^d \to [0, \infty)$, the *differential entropy* of X is defined as

$$H(X) := \int_{\mathbb{R}^d} -f(x) \log f(x)\, dx. \tag{3.2}$$

Recall the definition of $H(p)$ for a probability vector p (see (1.1)) and the height $H(\alpha)$ for an algebraic number α (see (2.3)). The multiple use of $H(\cdot)$ should cause no confusion, as it will always be clear from the context which type of the input is.

Let A be an invertible $d \times d$ real matrix and $b \in \mathbb{R}^d$. If X is a discrete random variable, it is easy to see that

$$H(AX + b) = H(X).$$

If X is an absolutely continuous random variable with finite differential entropy, then it follows from the change of variables formula that

$$H(AX+b) = H(X) + \log|\det A|\,. \tag{3.3}$$

4. Proof of Theorem 1.2

In this section, we prove Theorem 1.2 by assuming Theorems 1.3 and 1.4. The proofs of Theorems 1.3 and 1.4 will be given in the next two sections respectively.

Throughout this section, let $t_1,\dots,t_m$ be real algebraic numbers and $p=(p_1,\dots,p_m)$ a probability vector. Let ν denote the atomic probability measure $\sum_{j=1}^m p_j\delta_{t_j}$. For any $0\neq\eta\in(-1,1)$, let μ_η denote the self-similar measure associated with the IFS $\{\eta x+t_j\}_{j=1}^m$ and p, and let $h_{\eta,\nu}$ denote the Garsia entropy of η and ν (see (1.5)). Let $\mathcal{P}^n$, $n\in\mathbb{N}$, be defined as in (1.4).

The proof of Theorem 1.2 also relies on the following two results of Hochman.

Theorem 4.1 ([14]). *Let $0\neq\lambda\in(-1,1)$. Suppose that*

$$\dim\mu_\lambda < \min\left\{1,\ \frac{H(p)}{-\log|\lambda|}\right\}.$$

Then for every $\delta\in(0,1)$ and for all sufficiently large n (depending on λ and δ), there is a polynomial $0\neq P\in\mathcal{P}^n$ such that $|P(\lambda)|<\delta^n$.

Theorem 4.2 ([14]). *For each nonzero algebraic number $\eta\in(-1,1)$,*

$$\dim\mu_\eta = \min\left\{1,\ \frac{h_{\eta,\nu}}{-\log|\eta|}\right\}.$$

If in addition the IFS $\{\eta x+t_j\}_{j=1}^m$ has no exact overlaps, then

$$\dim\mu_\eta = \min\left\{1,\ \frac{H(p)}{-\log|\eta|}\right\}.$$

Now we are ready to prove Theorem 1.2.

Proof of Theorem 1.2. We follow the proof of [24, Theorem A.8] with slight modifications. For the reader's convenience, we provide the full details.

By Theorem 4.2, we may assume that λ is transcendental. Since all t_j are algebraic numbers, the IFS $\{\lambda x+t_j\}_{j=1}^m$ has no exact overlaps. Suppose on the contrary that

$$\dim\mu_\lambda < \min\left\{1,\ \frac{H(p)}{-\log|\lambda|}\right\}.$$

Let $\varepsilon>0$ such that

$$\dim\mu_\lambda + 2\varepsilon < \min\left\{1,\ \frac{H(p)}{-\log|\lambda|}\right\}. \tag{4.1}$$

Let $M > 0$ be large with respect to $t_1, \dots, t_m, \lambda, p$ and ε. Let $n_0 \in \mathbb{N}$ be large depending on M.

By Theorem 1.3, there exist an integer $n > n_0$ and a nonzero algebraic number η in $(-1, 1)$ such that η is a root of some polynomial in $\mathcal{P}^n$ satisfying

$$\dim \mu_\eta < \dim \mu_\lambda + \varepsilon \quad \text{and} \quad |\lambda - \eta| < \exp(-n^3). \tag{4.2}$$

We can assume n_0 is large enough so that $|\eta| < (1 + |\lambda|)/2$. By (4.1) and (4.2),

$$\dim \mu_\eta < \min\left\{1, \frac{H(p)}{-\log|\lambda|}\right\} - \epsilon. \tag{4.3}$$

Since η is algebraic, it follows from Theorem 4.2 that

$$\dim \mu_\eta = -\frac{h_{\eta,\nu}}{\log|\eta|}.$$

From this, (4.3) and Theorem 1.4, we may assume that $M(\eta) < M$.

Set $d = [\mathbb{Q}(\mathcal{D}) : \mathbb{Q}]$. Since $Q(\eta) = 0$ for some $Q \in \mathcal{P}^n$, it follows that

$$[\mathbb{Q}(\mathcal{D}, \eta) : \mathbb{Q}(\mathcal{D})] \le n.$$

Hence

$$\deg(\eta) = [\mathbb{Q}(\eta) : \mathbb{Q}] \le [\mathbb{Q}(\mathcal{D}, \eta) : \mathbb{Q}] = [\mathbb{Q}(\mathcal{D}, \eta) : \mathbb{Q}(\mathcal{D})][\mathbb{Q}(\mathcal{D}) : \mathbb{Q}] \le dn,$$

where the second equality is a basic property about the degree of field extension (see e.g. [21, Proposition 1.20]). Since λ is transcendental and η is algebraic, it follows that $|\lambda - \eta| > 0$. Let n' be the unique integer such that

$$(2M)^{-d(n'+1)} \le |\lambda - \eta| < (2M)^{-dn'}. \tag{4.4}$$

Then by this and the second inequality in (4.2),

$$d(n' + 1)\log(2M) \ge -\log|\lambda - \eta| > n^3.$$

Thus for n_0 large enough,

$$n' \ge (M + 2)(dn)^2. \tag{4.5}$$

Applying Theorem 4.1 with $\delta = (2M)^{-3Md}$, we see that for n_0 large enough there exists $0 \ne P \in \mathcal{P}^{n'}$ such that $|P(\lambda)| \le (2M)^{-3Mdn'}$. Then by (4.4),

$$|\lambda - \eta| \ge (2M)^{-d(n'+1)} \ge (2M)^{dn'}|P(\lambda)|^{1/M} \ge (2M)^{dn'/M}|P(\lambda)|^{1/M}.$$

Combining this with (4.4) yields

$$(2M)^{dn'/M}|P(\lambda)|^{1/M} \le |\lambda - \eta| \le (2M)^{-dn'}.$$

From this, (4.5) and Proposition 2.12 (in which we replace n by dn, and k by M), we conclude that η is a zero of P of order at least M.

On the other hand, since $|\eta| < (1 + |\lambda|)/2 < 1$, it follows from Proposition 2.7 that there exists $N > 0$ (only depending on $t_1, \dots, t_m$ and λ) such that η is a zero of P of order at most N. Recall that M is allowed to depend upon $t_1, \dots, t_m$ and λ. This leads to a contradiction by letting $M > N$. □

5. Proof of Theorem 1.3

Proof of Theorem 1.3. The proof is almost identical to that of [24, Theorem A.2], except some of the involved auxiliary results in Appendix A.1 of [24, Theorem A.2] should be slightly revised.

Indeed, Theorem 1.3 is proved in [24, Theorem A.2] in the setting of homogeneous IFSs on $\mathbb{R}$ with positive contraction ratio and rational translations. Actually, the proof of [24, Theorem A.2] is based on some auxiliary results (see Propositions A.3, A.5 and A.7, Lemmas A.4 and A.6 in [24]), which adapts and generalizes the main results of [8]. As was pointed out in [24, Appendix A.1], the statements of [24, Proposition A.3, Lemmas A.4 and A.6] hold in the general setting of homogeneous IFSs on $\mathbb{R}$. Notice that [24, Proposition A.5 and A.7] are built on a separation property (see [24, Lemma 4.1]) for roots of integer polynomials. By a similar separation property (see Proposition 2.9) for roots of polynomials with algebraic coefficients, we are able to extend [24, Proposition A.5 and A.7] to the general setting of Theorem 1.3; see Propositions 5.1 and 5.2. Using these (revised) auxiliary results, the exact arguments of [24, Theorem A.2] give the proof Theorem 1.3. Since the proof of [24, Theorem A.2] is rather long, we do not repeat it here and we refer to [24] for the full details. □

Below we present two propositions, which are the extensions of [24, Propositions A.5 and A.7] to our setting respectively.

Following [8, Section 2] and [27, Section 2], we introduce the notion of entropy at a given scale. Let X be a bounded random variable in $\mathbb{R}$ with law ν. For $r > 0$, the *entropy of X at scale r* is defined by

$$H(X; r) := \int_0^1 H(\lfloor X/r + t \rfloor)\, dt,$$

where $\lfloor x \rfloor$ stands for the largest integer not exceeding x. We also write $H(\nu; r)$ in place of $H(X; r)$. As was noted in [8, 27], the idea of the above averaging procedure originates in Wang's paper [29].

In the remaining part of this section, let $t_1, \ldots, t_m$ be real algebraic numbers and $p = (p_1, \ldots, p_m)$ a probability vector. Let ν denote the atomic probability measure $\sum_{j=1}^m p_j \delta_{t_j}$. For any $0 \neq \eta \in (-1, 1)$, let μ_η denote the self-similar measure associated with the IFS $\{\eta x + t_j\}_{j=1}^m$ and p, and let $\mu_\eta^{(n)}$ $(n \in \mathbb{N})$ denote the probability law of $\sum_{k=0}^{n-1} \xi_k \eta^k$, where $(\xi_k)_{k=0}^{n-1}$ is a finite sequence of independent random variables with common law $\nu = \sum_{j=1}^m p_j \delta_{t_j}$. Let $\mathcal{D}$ and $\mathcal{P}^n$, $n \in \mathbb{N}$, be defined as in (1.3) and (1.4), respectively.

Proposition 5.1. *For every $\varepsilon \in (0, 1/2)$ there exists $C > 0$ (depending on ε, $t_1, \ldots, t_m$ and $p_1, \ldots, p_m$) such that the following holds for all $n \geq N(\varepsilon, C)$. Let $0 < r < n^{-Cn}$ and $\lambda \in (\varepsilon - 1, -\varepsilon) \cup (\varepsilon, 1 - \varepsilon)$ be given, and suppose that $\frac{1}{n} H(\mu_\lambda^{(n)}; r) < H(p)$. Then there exists $0 \neq \eta \in (-1, 1)$, which is a root of a nonzero polynomial in*

$\mathcal{P}^n$ *such that*

$$|\lambda - \eta| < r^{1/C} \quad \text{and} \quad H(\mu_\eta^{(n)}) \leq H(\mu_\lambda^{(n)}; r).$$

Proof. Here we follow the arguments in the proof of [24, Proposition A.5] with slight modifications.

Let $\varepsilon \in (0,1)$ and let $C > 1$ be large with respect to ε. Let n be large with C and let r and λ be given as in the statement of the proposition.

Let $0 \leq s \leq 1$ be with

$$H\left(\left\lfloor r^{-1} \sum_{k=0}^{n-1} \xi_k \lambda^k + s \right\rfloor\right) \leq H(\mu_\lambda^{(n)}; r) < nH(p), \tag{5.1}$$

where $(\xi_k)_{k=0}^{n-1}$ is a finite sequence of independent random variables with common law $\nu = \sum_{j=1}^m p_j \delta_{t_j}$. Let $\mathcal{A}$ be the set of all nonzero $P \in \mathcal{P}^n$ with $|P(\lambda)| \leq r$. By (5.1), $\mathcal{A}$ is nonempty.

Given $P \in \mathcal{A}$ it follows that Corollary 2.8 that there exists $\eta_P \in \mathbb{C}$ with $P(\eta_P) = 0$ and

$$|\eta_P - \lambda| \leq (u 2^n \varepsilon^{-n} r)^{C^{-1/4}},$$

$u = \min\{|s| : 0 \neq s \in \mathcal{D}\}$.

From $r < n^{-Cn}$, since C is large with respect to ε, and since n is large with respect to C, it follows that we may assume $|\lambda - \eta_P| < r^{C^{-1/2}}$.

For $Q, P \in \mathcal{A}$,

$$|\eta_P - \eta_Q| \leq |\eta_P - \lambda| + |\lambda - \eta_Q| \leq 2r^{C^{-1/2}} < 2n^{-C^{1/2}n}.$$

Thus, by Proposition 2.9 and by assuming that C is large enough, it follows that $\eta_P = \eta_Q$. Write η for this common value, then $P(\eta) = 0$ for all $P \in \mathcal{A}$. From this, (5.1) and by the definition of $\mathcal{A}$,

$$H(\mu_\eta^{(n)}) \leq H(\mu_\lambda^{(n)}; r).$$

Since $\lambda \in \mathbb{R}$ we have $|\overline{\eta} - \lambda| = |\eta - \lambda|$, and so $|\eta - \overline{\eta}| \leq 2n^{-C^{1/2}n}$. For $P \in \mathcal{A}$ we clearly have $P(\overline{\eta}) = 0$. Thus, another application of Proposition 2.9 gives $\eta = \overline{\eta}$. Since $\lambda \in (\varepsilon - 1, -\varepsilon) \cup (\varepsilon, 1 - \varepsilon)$ we may assume $0 \neq \eta \in (-1, 1)$, which completes the proof of the proposition. □

Proposition 5.2. *For every $\varepsilon \in (0, 1/2)$ there exists $C > 0$ (depending on ε, $t_1, \ldots, t_m$ and $p_1, \ldots, p_m$) such that the following holds for all $n \geq N(\varepsilon, C)$. Let $\lambda \in (\varepsilon - 1, -\varepsilon) \cup (\varepsilon, 1 - \varepsilon)$ and suppose that there exists $0 \neq \eta \in \mathbb{C}$, which is a root of a nonzero polynomial in $\mathcal{P}^n$, such that $|\lambda - \eta| < n^{-Cn}$. Then $\frac{1}{n} H(\mu_\lambda^{(n)}; r) = H(p)$ for all $r \leq |\lambda - \eta|^C$.*

Proof. Here we repeat the proof of [24, Proposition A.7] with trivial modifications. Let $\varepsilon \in (0, 1/2)$, and let $C > 1$ be large with respect to ε, let $n \geq 1$ be large with respect to C, and let λ and η be as in the statement of the proposition.

Suppose to the contrary that there exists $0 < r \le |\lambda-\eta|^C$ with $\frac{1}{n}H(\mu_\lambda^{(n)}; r) < H(p)$. By Corollary 2.8, there exists $\eta' \in (0,1)$, which is a root of a nonzero polynomial in $\mathcal{P}^n$, such that $|\lambda - \eta'| < r^{1/C} \le |\lambda - \eta|$. In particular $\eta \neq \eta'$ and

$$|\eta - \eta'| \le |\eta - \lambda| + |\lambda - \eta'| \le 2n^{-Cn}.$$

However, as C is assumed to be large enough, this contradicts Proposition 2.9, which completes the proof of the proposition. □

6. Proof of Theorem 1.4

In this section, we prove the following slight extension of Theorem 1.4, in which we allow $t_1, \dots, t_m$ and η to be complex algebraic numbers.

Theorem 6.1. *Let $t_1, \dots, t_m$ be algebraic numbers in $\mathbb{C}$, and let $p = (p_j)_{j=1}^m$ be a probability vector. Set $\nu = \sum_{j=1}^m p_j \delta_{t_j}$. Then for any $h \in (0, H(p))$, there is a positive number M depending on h, $t_1, \dots, t_m$ and $p_1, \dots, p_m$, such that*

$$h_{\eta,\nu} \ge h$$

for every algebraic number $\eta \in \mathbb{C}$ with $0 < |\eta| < 1$ and $M(\eta) \ge M$, where $h_{\eta,\nu}$ and $M(\cdot)$ are defined as in (1.5) and (1.6) respectively.

Proof of Theorem 6.1 assuming Proposition 1.5. Let $t_1, \dots, t_m, p_1, \dots, p_m, \nu$ and h be given as in the statement of the theorem. Set

$$\mathcal{D} = \{t_i - t_j : \ 1 \le i, j \le m\}.$$

Recall that for $n \in \mathbb{N}$, $\mathcal{P}^n$ denotes the collection of polynomials in variable X of degree not exceeding n and with coefficients in $\mathcal{D}$.

By Lemma 2.6, we can choose an algebraic integer θ such that $\mathbb{Q}(\theta) = \mathbb{Q}(\mathcal{D})$. Then we can take a suitable $L \in \mathbb{N}$ such that

$$Ls \in \mathbb{Z}[\theta] \quad \text{ for all } s \in \mathcal{D}.$$

Let $d = \deg(\theta)$, and let $\theta_1, \dots, \theta_d$ be the algebraic conjugates of θ, with $\theta_1 = \theta$. For $k = 1, \dots, d$, let

$$\sigma_k : \mathbb{Q}(\theta) \to \mathbb{Q}(\theta_k)$$

be the field isomorphism such that $\sigma_k(a) = a$ for all $a \in \mathbb{Q}$ and $\sigma_k(\theta) = \theta_k$.

Let $\eta \in \mathbb{C}$ be an algebraic number with $0 < |\eta| < 1$. We aim to show that $h_{\eta,\nu} \ge h$ if the Mahler measure of η is large enough. For this purpose, we may assume that that $P(\eta) = 0$ for some $P = \sum_{i=0}^n a_i X^i \in \mathcal{P}^n$ with $n \in \mathbb{N}$; otherwise $h_{\eta,\nu} = H(p) > h$ and there is nothing left to prove.

Fix such a polynomial $P = \sum_{i=0}^n a_i X^i$ and define $F \in \mathbb{C}[X]$ by

$$F = L^d \prod_{k=1}^d \left(\sum_{i=0}^n \sigma_k(a_i) X^i \right) = \prod_{k=1}^d \left(\sum_{i=0}^n \sigma_k(La_i) X^i \right). \tag{6.1}$$

Then $\deg(F) \le dn$, and $F \in \mathbb{Z}[X]$ which follows from the same argument as in the last paragraph of the proof of Lemma 2.11. Clearly $F(\eta) = 0$. Writing $F = \sum_{i=0}^{dn} \beta_i X^i$ and letting τ be the smallest integer so that $\beta_\tau \neq 0$, by (6.1) we see that $\beta_\tau = L^d \sigma_1(a_j) \cdots \sigma_d(a_j)$, where j is the smallest integer so that $a_j \neq 0$. Hence $|\beta_\tau| \le C$, where

$$C := L^d \max\{|\sigma_k(s)|^d : \ 1 \le k \le d, \ s \in \mathcal{D}\}.$$

We claim that

$$\widetilde{M}_{\mathbb{Q}}(\eta) \ge C^{-1} M(\eta), \tag{6.2}$$

where $\widetilde{M}_{\mathbb{Q}}(\eta)$ is defined as in (1.8). To see this, let $f = \sum_{i=0}^{q} b_i X^i$ be the minimal polynomial of η in $\mathbb{Z}[X]$. Then f divides F in $\mathbb{Z}[X]$ since $F(\eta) = 0$. It follows that b_0 divides β_τ. Consequently, $|b_0| \le |\beta_\tau| \le C$. By Vieta's formulas,

$$\frac{M(\eta)}{\widetilde{M}_{\mathbb{Q}}(\eta)} = |b_q| \prod_{i=1}^{q} |\alpha_i| = |b_0| \le C, \tag{6.3}$$

where $\alpha_1, \ldots, \alpha_q$ are the roots of f. This yields (6.2).

Now let $g = \sum_{i=0}^{r} c_i X^i$ be a minimal polynomial of η in $\mathbb{Q}(\theta)[X]$ with coefficients in $\mathbb{Z}[\theta]$. For $1 \le k \le d$, define

$$g_k = \sigma_k(g) := \sum_{i=0}^{r} \sigma_k(c_i) X^i. \tag{6.4}$$

Since g is irreducible in $\mathbb{Q}(\theta)[X]$, g_k is irreducible in $\mathbb{Q}(\theta_k)[X]$ for each $1 \le k \le d$. Define

$$G = \prod_{k=1}^{d} g_k.$$

Then $G \in \mathbb{Z}[X]$, which follows from a similar argument as in the last paragraph of the proof of Lemma 2.11. Since $G(\eta) = 0$ and f is the minimal polynomial of η in $\mathbb{Z}[X]$, it follows that f divides G.

For a nonzero polynomial $Q = \sum_{i=0}^{l} e_i X^i = e_l \prod_{i=1}^{l} (X - \gamma_l)$ in $\mathbb{C}[X]$, define

$$\widetilde{M}(Q) = \prod_{1 \le i \le l: \ \gamma_i \neq 0} \frac{1}{\min\{1, |\gamma_i|\}},$$

with convention $\widetilde{M}(Q) = 1$ if Q does not have a nonzero root. It is easily checked that $\widetilde{M}(Q_1 Q_2) = \widetilde{M}(Q_1)\widetilde{M}(Q_2)$ for any $Q_1, Q_2 \in \mathbb{C}[X]$, and $\widetilde{M}(Q_2) \ge \widetilde{M}(Q_1)$ if Q_1 divides Q_2.

For $1 \le k \le d$, let η_k be a root of g_k. Using the above notation, we have

$$\prod_{k=1}^{d} \widetilde{M}_{\mathbb{Q}(\theta_k)}(\eta_k) = \prod_{k=1}^{d} \widetilde{M}(g_k) = \widetilde{M}(G) \ge \widetilde{M}(f) = \widetilde{M}_{\mathbb{Q}}(\eta),$$

where the first equality holds since g_k is irreducible in $\mathbb{Q}(\theta_k)[X]$, the second equality holds by $G = \prod_{k=1}^{d} g_k$, the inequality is by f dividing G, and the last equality holds

since $f \in \mathbb{Z}[X]$ is the minimal polynomial of η over $\mathbb{Q}$. Thus there exists some $1 \le k_0 \le d$ such that

$$\widetilde{M}_{\mathbb{Q}(\theta_{k_0})}(\eta_{k_0}) \ge (\widetilde{M}_{\mathbb{Q}}(\eta))^{1/d}.$$

From this and (6.2), we conclude that

$$\widetilde{M}_{\mathbb{Q}(\theta_{k_0})}(\eta_{k_0}) \ge \frac{M(\eta)^{1/d}}{C} > 1 \tag{6.5}$$

when $M(\eta)$ is sufficiently large depending on $\mathcal{D}$. Due to this and (1.8), replacing η_{k_0} by another root of g_{k_0} if necessary, we may assume that $|\eta_{k_0}| < 1$.

Define $\widetilde{t}_j = \sigma_{k_0}(t_j)$ for $1 \le j \le m$. Let $\widetilde{\nu} = \sum_{j=1}^m p_j \delta_{\widetilde{t}_j}$ and let $(\widetilde{\xi}_j)_{j=0}^\infty$ be a sequence of i.i.d. random variables with common law $\widetilde{\nu}$. Next we show that

$$h_{\eta,\nu} = h_{\eta_{k_0},\widetilde{\nu}}. \tag{6.6}$$

To see this, it is sufficient to prove that for each polynomial Q with coefficients in $\mathcal{D}$,

$$Q(\eta) = 0 \Longleftrightarrow \sigma_{k_0}(Q)(\eta_{k_0}) = 0, \tag{6.7}$$

where $\sigma_{k_0}(Q)$ is defined in a way similar to the definition of $\sigma_k(g)$ (see (6.4)). Below we only prove the direction '$\Longrightarrow$' in (6.7), the other direction follows from a similar argument.

Suppose that $Q(\eta) = 0$ for some polynomial Q with coefficients in $\mathcal{D}$. Since $\mathcal{D} \subset \mathbb{Q}(\theta)$ and g is a minimal polynomial of η in $\mathbb{Q}(\theta)$, g divides Q in $\mathbb{Q}(\theta)[X]$. It follows that g_{k_0} divides $\sigma_{k_0}(Q)$ in $\mathbb{Q}(\theta_{k_0})[X]$. Hence $\sigma_{k_0}(Q)(\eta_{k_0}) = 0$ by $g_{k_0}(\eta_{k_0}) = 0$. This proves the direction '$\Longrightarrow$' in (6.7).

By (6.6) and Proposition 1.5 applied to $h_{\eta_{k_0},\widetilde{\nu}}$, we obtain that

$$h_{\eta,\nu} = h_{\eta_{k_0},\widetilde{\nu}} \ge \Phi_{\widetilde{\nu}}\left(\widetilde{M}_{\mathbb{Q}(\theta_{k_0})}(\eta_{k_0})\right), \tag{6.8}$$

where $\Phi_{\widetilde{\nu}}(\cdot)$ is as defined in (1.7) (in which we replace ν by $\widetilde{\nu}$).

Let G be a complex Gaussian random variable with density $\exp(-|z|^2/2)/(2\pi)$ for $z \in \mathbb{C}$. It is easily checked that

$$\lim_{a\to\infty} H(a^{-1/2}\zeta + G) = H(G) \quad \text{and} \quad \lim_{a\to\infty} H(a^{1/2}\zeta + G) = H(G) + H(p),$$

where ζ is a random variable with law $\widetilde{\nu}$ and independent of G. Taking $t = a^{-1/2}$ in (1.7) (in which we replace ν by $\widetilde{\nu}$) yields that

$$\lim_{a\to\infty} \Phi_{\widetilde{\nu}}(a) \ge \lim_{a\to\infty} H(a^{1/2}\zeta + G) - H(a^{-1/2}\zeta + G) = H(p).$$

Since $h < H(p)$, there exists $A > 0$ such that $\Phi_{\widetilde{\nu}}(a) \ge h$ for all $a \ge A$. By (6.5), $\widetilde{M}_{\mathbb{Q}(\theta_{k_0})}(\eta_{k_0}) \ge A$ if $M(\eta) > (AC)^d$. Then (6.8) implies that

$$h_{\eta,\nu} \ge h$$

if $M(\eta) > (AC)^d$, which finishes the proof of the theorem. □

In the remaining part of this section, we prove Proposition 1.5. This is done by adapting the ideas of [9, Proposition 13].

Let $\mathrm{M}_d(\mathbb{C})$, $\mathrm{M}_d(\mathbb{R})$, $\mathrm{GL}_d(\mathbb{R})$, $\mathrm{U}_d(\mathbb{C})$ and $\mathrm{O}_d(\mathbb{R})$ denote the sets of $d\times d$ complex, real, real invertible, unitary and orthogonal matrices, respectively. Since $\mathbb{K}$ in Proposition 1.5 is assumed to be a subfield of $\mathbb{C}$, to apply the ideas of [9, Proposition 13], we need to transfer certain series of complex matrices into the series of real matrices. For this purpose, let us introduce a natural operator transferring complex matrices into real matrices.

Define $\widehat{\cdot}\,: \mathrm{M}_d(\mathbb{C}) \to \mathrm{M}_{2d}(\mathbb{R})$ by

$$\widehat{A} = \left(\begin{bmatrix} a_{jk} & -b_{jk} \\ b_{jk} & a_{jk} \end{bmatrix} \right)_{1\le j,k\le d} \tag{6.9}$$

if $A = (a_{jk} + ib_{jk})_{1\le j,k\le d}$, where $a_{jk}, b_{jk} \in \mathbb{R}$. For $z \in \mathbb{C}$, write $\widehat{z} := \widehat{zI_d}$, where I_d stands for the $d \times d$ identity matrix.

Lemma 6.2. *The operator $\widehat{\cdot}$ has the following properties.*

(i) $\widehat{\cdot}$ is an injective $\mathbb{R}$-algebra homomorphism, that is, for $\alpha \in \mathbb{R}$, $A, B \in \mathrm{M}_d(\mathbb{C})$,

$$\widehat{\alpha A + B} = \alpha\widehat{A} + \widehat{B} \quad \text{and} \quad \widehat{AB} = \widehat{A}\widehat{B}.$$

(ii) $\widehat{\cdot}$ is an isometry with respect to the norms induced by the standard inner products respectively, that is, $\|\widehat{A}\| = \|A\|$ for $A \in \mathrm{M}_d(\mathbb{C})$. If $U \in \mathrm{U}_d(\mathbb{C})$, then $\widehat{U} \in \mathrm{O}_{2d}(\mathbb{R})$.

(iii) For any $z \in \mathbb{C}$ and $B \in \mathrm{M}_d(\mathbb{C})$,

$$\widehat{z}\widehat{B} = \widehat{zB} = \widehat{B}\widehat{z}. \tag{6.10}$$

Proof. Consider the natural map $\tau\colon \mathbb{C}^d \to \mathbb{R}^{2d}$ defined by

$$\tau\left(x_1 + iy_1, \dots, x_d + iy_d\right) = (x_1, y_1, \dots, x_d, y_d),$$

where $x_j, y_j \in \mathbb{R}$ for $1 \le j \le d$. Then the statements of the lemma are readily justified from the fact that $\widehat{A}x = \tau A \tau^{-1} x$ for $A \in \mathrm{M}_d(\mathbb{C})$ and $x \in \mathbb{R}^{2d}$. We leave the details to the reader. □

To prove Proposition 1.5, we need to use some entropy quantities introduced in [9]. Let X, Y be two independent bounded random variables in $\mathbb{R}^d$. Following Breuillard and Varjú [9], we define

$$H(X; B) = H(X + G_B) - H(G_B) \quad \text{for } B \in \mathrm{GL}_d(\mathbb{R}), \tag{6.11}$$

where G_B is a Gaussian random variable in $\mathbb{R}^d$, independent of X, with mean 0 and covariance matrix BB^t. Here B^t stands for the transpose of B. For $B_1, B_2 \in \mathrm{GL}_d(\mathbb{R})$, write

$$H(X; B_1|B_2) = H(X; B_1) - H(X; B_2). \tag{6.12}$$

It follows from (3.3) that for each $B \in \mathrm{GL}_d(\mathbb{R})$,

$$H(X; B_1) = H(BX; BB_1) \quad \text{and} \quad H(X; B_1|B_2) = H(BX; BB_1|BB_2). \tag{6.13}$$

The quantities defined above have the following properties, which are crucial for our proof.

Lemma 6.3 ([9]). *Let $B_1, B_2 \in \mathrm{GL}_d(\mathbb{R})$ be such that $\|B_1 x\| \le \|B_2 x\|$ for all $x \in \mathbb{R}^d$. Assume that X, Y are two bounded independent random variables taking values in $\mathbb{R}^d$. Then*

$$H(X+Y; B_1) \le H(X; B_1) + H(Y; B_1), \tag{6.14}$$

$$H(X; B_1 O) = H(X; B_1) \quad \text{for } O \in \mathrm{O}_d(\mathbb{R}), \tag{6.15}$$

and

$$H(X+Y; B_1|B_2) \ge H(X; B_1|B_2).$$

As a consequence, for any $B \in \mathrm{GL}_d(\mathbb{R})$ with $\|B\| \le 1$ and any nonnegative integer k,

$$H\left(X+Y; B^{k+1}|B^k\right) \ge H\left(X; B^{k+1}|B^k\right). \tag{6.16}$$

If in addition X is discrete, then

$$H(X) \ge H(X; B_1). \tag{6.17}$$

Proof. (6.15) follows from that $G_{B_1 O}$ and G_{B_1} have the same distribution for each $O \in \mathrm{O}_d(\mathbb{R})$. (6.17) follows from the definition of $H(X; B_1)$ and [9, Equation (2.10)]. The other statements of the lemma come from [9, Lemma 9]. □

Now we are ready to prove Proposition 1.5.

Proof of Proposition 1.5. We adapt the proof of [9, Proposition 13].

First we make some preparations. Let $\eta_1, \dots, \eta_d$ (including η itself) be all the conjugates of η over $\mathbb{K}$ with modulus strictly less than 1. By (1.8), $\widetilde{M}_{\mathbb{K}}(\eta) = 1/|\eta_1 \cdots \eta_d|$. Set

$$\alpha_1 = \cdots = \alpha_{d-1} = 1, \quad \alpha_d = |\eta_1 \cdots \eta_d| = \frac{1}{\widetilde{M}_{\mathbb{K}}(\eta)}.$$

By [15, Theorem 3] (see also [16, Theorem 3.6.6]), there exist $A \in \mathrm{M}_d(\mathbb{C})$ with eigenvalues $\eta_1, \dots, \eta_d$ and $U, V \in \mathrm{U}_d(\mathbb{C})$ such that $A = UDV$, where $D = \mathrm{diag}(\alpha_1, \dots, \alpha_d)$. Then the spectral radius $\rho(A)$ of A is strictly less than 1, and the operator norm of A (with respect to the standard inner product on $\mathbb{C}^d$) satisfies $\|A\| = 1$. Recall the operator $\widehat{\cdot}$ defined in (6.9). Applying Lemma 6.2 gives $\widehat{A} = \widehat{U}\widehat{D}\widehat{V}$, where $\widehat{U}, \widehat{V} \in \mathrm{O}_{2d}(\mathbb{R})$, and $\left\|\widehat{A}\right\| = 1$. Moreover by (6.9),

$$\widehat{D} = \mathrm{diag}\left(\underbrace{1, \dots, 1}_{2d-2}, 1/\widetilde{M}_{\mathbb{K}}(\eta), 1/\widetilde{M}_{\mathbb{K}}(\eta)\right).$$

Let $(\xi_k)_{k=0}^{\infty}$ be a sequence of i.i.d. complex random variables with common law ν. Since $|\eta| < 1$ and $\rho(A) < 1$, the random variables

$$X_\eta := \sum_{k=0}^{\infty} \xi_k \eta^k \quad \text{and} \quad X_A := \sum_{k=0}^{\infty} \xi_k A^k$$

are well-defined and bounded. For $n \in \mathbb{N}$, write

$$X_\eta^{(n)} := \sum_{k=0}^{n-1} \xi_k \eta^k \quad \text{and} \quad X_A^{(n)} := \sum_{k=0}^{n-1} \xi_k A^k.$$

We claim that

$$H(X_\eta^{(n)}) \geq H(X_A^{(n)}) \quad \text{for all } n \in \mathbb{N}, \tag{6.18}$$

where $H(\cdot)$ stands for Shannon entropy (see (3.1)). To see this, let $\{v_1, \ldots, v_d\}$ be a basis of $\mathbb{C}^d$, with v_j being an eigenvector of A corresponding to the eigenvalue η_j for each $1 \leq j \leq d$. Suppose that $P(\eta) = 0$ for some polynomial $P = \sum_{i=0}^n a_i X^i \in \mathcal{P}^n$, with coefficients in

$$\mathcal{D} = \{u_i - u_j : \ 1 \leq i, j \leq m\}.$$

Since $\eta_1, \ldots, \eta_d$ are conjugates of η over $\mathbb{K}$ and $\{u_i\}_{i=1}^m \subset \mathbb{K}$, it follows that $P(\eta_j) = 0$ for $1 \leq j \leq d$. Moreover for each $x \in \mathbb{C}^d$, writing $x = \sum_{j=1}^d x_j v_j$ with $x_j \in \mathbb{C}$, we see that

$$\sum_{i=0}^n a_i A^i x = \sum_{j=1}^d \sum_{i=0}^n a_i x_j A^i v_j = \sum_{j=1}^d \sum_{i=0}^n a_i x_j \eta_j^i v_j = \sum_{j=1}^d x_j P(\eta_j) v_j = 0.$$

Hence $\sum_{i=0}^n a_i A^i = 0$. This implies (6.18).

Let $\widehat{\cdot} : \mathrm{M}_d(\mathbb{C}) \to \mathrm{M}_{2d}(\mathbb{R})$ be as defined in (6.9). By (i) and (ii) of Lemma 6.2,

$$\widehat{X_A^{(n)}} = \sum_{k=0}^{n-1} \widehat{\xi_k} \widehat{A}^k. \tag{6.19}$$

Notice that for each $n \in \mathbb{N}$ and $x \in \mathbb{R}^{2d}$, $\widehat{X_A^{(n)}} x$ is a discrete random variable in $\mathbb{R}^{2d}$. Below we show that

$$H\left(\widehat{X_A^{(n)}} x\right) \geq n H\left(\widehat{\xi}_0 x; \widehat{A} | I_{2d}\right) \quad \text{for all } n \in \mathbb{N},\ x \in \mathbb{R}^{2d}, \tag{6.20}$$

where $H(\cdot, \cdot|\cdot)$ is the quantity defined as in (6.12), and I_{2d} stands for the identity matrix in $\mathrm{M}_{2d}(\mathbb{R})$.

To see (6.20), write

$$\widehat{X_{A,k}} = \sum_{j=0}^{\infty} \widehat{\xi_{k+j}} \widehat{A}^j.$$

Since ξ_j $(j \in \mathbb{N})$ are i.i.d., the random variable $\widehat{X_{A,k}}$ is independent of $\widehat{X_A^{(k)}}$ and has the same law as $\widehat{X_A}$. By (6.10) and (6.19), for each $k \in \mathbb{N}$,

$$\widehat{X_A} = \sum_{i=0}^{k-1} \widehat{\xi_i} \widehat{A}^i + \sum_{j=k}^{\infty} \widehat{\xi_j} \widehat{A}^j = \widehat{X_A^{(k)}} + \widehat{A}^k \sum_{j=0}^{\infty} \widehat{\xi_{k+j}} \widehat{A}^j = \widehat{X_A^{(k)}} + \widehat{A}^k \widehat{X_{A,k}}. \tag{6.21}$$

Hence for each $x \in \mathbb{R}^{2d}$,

$$
\begin{aligned}
& nH\left(\widehat{\xi}_0 x; \widehat{A} | I_{2d}\right) \\
& \leq \sum_{k=0}^{n-1} H\left(\widehat{X_A} x; \widehat{A} | I_{2d}\right) && \text{(by (6.16) and (6.21))} \\
& = \sum_{k=0}^{n-1} H\left(\widehat{A}^k \widehat{X_A} x; \widehat{A}^{k+1} | \widehat{A}^k\right) && \text{(by (6.13))} \\
& \leq \sum_{k=0}^{n-1} H\left(\widehat{X_A} x; \widehat{A}^{k+1} | \widehat{A}^k\right) && \text{(by (6.16) and (6.21))} \\
& = H\left(\widehat{X_A} x; \widehat{A}^n\right) - H\left(\widehat{X_A} x; I_{2d}\right) && \text{(by (6.12))} \\
& \leq H\left(\widehat{X_A^{(n)}} x; \widehat{A}^n\right) + H\left(\widehat{A}^n \widehat{X_A} x; \widehat{A}^n\right) - H\left(\widehat{X_A} x; I_{2d}\right) && \text{(by (6.14) and (6.21))} \\
& = H\left(\widehat{X_A^{(n)}} x; \widehat{A}^n\right) && \text{(by (6.13))} \\
& \leq H\left(\widehat{X_A^{(n)}} x\right) && \text{(by (6.17))}.
\end{aligned}
$$

This proves (6.20).

Now we are ready to estimate $h_{\eta,\nu}$. By (1.5), (6.18) and the injectivity of the operator $\widehat{\cdot}$,

$$
h_{\eta,\nu} = \lim_{n\to\infty} \frac{H\left(X_\eta^{(n)}\right)}{n} \geq \lim_{n\to\infty} \frac{H\left(X_A^{(n)}\right)}{n} = \lim_{n\to\infty} \frac{H\left(\widehat{X_A^{(n)}}\right)}{n}. \tag{6.22}
$$

For $n \in \mathbb{N}$ and $x \in \mathbb{R}^{2d}$,

$$
\begin{aligned}
\frac{H\left(\widehat{X_A^{(n)}}\right)}{n} &\geq \frac{H\left(\widehat{X_A^{(n)}} x\right)}{n} \\
&\geq H\left(\widehat{\xi}_0 x; \widehat{A} | I_{2d}\right) && \text{(by (6.20))} \\
&= H\left(\widehat{\xi}_0 x; \widehat{U}\widehat{D}\widehat{V} | I_{2d}\right) && \text{(by } \widehat{A} = \widehat{U}\widehat{D}\widehat{V}\text{)} \\
&= H\left(\widehat{U}^{-1}\widehat{\xi}_0 x; \widehat{D}\widehat{V} | \widehat{U}^{-1}\right) && \text{(by (6.13))} \\
&= H\left(\widehat{U}^{-1}\widehat{\xi}_0 x; \widehat{D}\widehat{V}\right) - H\left(\widehat{U}^{-1}\widehat{\xi}_0 x; \widehat{U}^{-1}\right) && \text{(by (6.12))} \\
&= H\left(\widehat{U}^{-1}\widehat{\xi}_0 x; \widehat{D}\right) - H\left(\widehat{U}^{-1}\widehat{\xi}_0 x; I_{2d}\right) && \text{(by (6.15))} \\
&= H\left(\widehat{\xi}_0 \widehat{U}^{-1} x; \widehat{D} | I_{2d}\right) && \text{(by (6.12) and (6.10))}
\end{aligned}
$$

$$= H\left(\widehat{\xi}_0 y; \widehat{D}|I_{2d}\right) \qquad \text{(by taking } y = \widehat{U}^{-1}x\text{)}.$$

From this and (6.22), we obtain that

$$h_{\eta,\nu} \geq H\left(\widehat{\xi}_0 x; \widehat{D}|I_{2d}\right) \quad \text{for all } x \in \mathbb{R}^{2d}.$$

For each $t > 0$, taking $x = te_{2d-1} = (0, \ldots, 0, t, 0)$ in the above inequality gives

$$\begin{aligned}
h_{\eta,\nu} &\geq H\left(t\widehat{\xi}_0 e_{2d-1}; \widehat{D}\right) - H\left(t\widehat{\xi}_0 e_{2d-1}; I_{2d}\right) && \text{(by (6.12))} \\
&= H\left(t\widehat{D}^{-1}\widehat{\xi}_0 e_{2d-1}; I_{2d}\right) - H\left(t\widehat{\xi}_0 e_{2d-1}; I_{2d}\right) && \text{(by (6.13))} \\
&= H\left(t\widehat{D}^{-1}\widehat{\xi}_0 e_{2d-1} + G_{I_{2d}}\right) - H\left(t\widehat{\xi}_0 e_{2d-1} + G_{I_{2d}}\right) && \text{(by (6.11))} \\
&= H\left(t\widetilde{M}_{\mathbb{K}}(\eta)\zeta + G_{I_2}\right) - H\left(t\zeta + G_{I_2}\right),
\end{aligned}$$

where in the last equality ζ is a random variable in $\mathbb{R}^2$ with law $\sum_{j\in\Lambda} p_j \delta_{(a_j,b_j)}$ in which a_j, b_j are the real and imaginary parts of u_j (i.e. $u_j = a_j + ib_j$), and ζ is independent of G_{I_2}. The last equality follows by integrating out the first $2d-2$ variables using $F(xy) = xF(y) + yF(x)$ where $F(x) = -x\log x$. This finishes the proof by taking supremum over $t > 0$. □

Acknowledgements. This work was partially supported by the General Research Funds (CUHK14305722, CUHK14308423) from the Hong Kong Research Grant Council, and by a direct grant for research from the Chinese University of Hong Kong.

References

[1] Michael Artin. *Algebra*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1991.

[2] Simon Baker. Iterated function systems with super-exponentially close cylinders. *Adv. Math.*, 379:Paper No. 107548, 13 pp., 2021.

[3] Simon Baker. Iterated function systems with super-exponentially close cylinders II. *Proc. Amer. Math. Soc.*, 150(1):245–256, 2022.

[4] Balázs Bárány and Antti Käenmäki. Super-exponential condensation without exact overlaps. *Adv. Math.*, 379:Paper No. 107549, 22 pp., 2021.

[5] Julien Barral and De-Jun Feng. On multifractal formalism for self-similar measures with overlaps. *Math. Z.*, 298(1-2):359–383, 2021.

[6] Frank Beaucoup, Peter Borwein, David W. Boyd, and Christopher Pinner. Multiple roots of $[-1, 1]$ power series. *J. Lond. Math. Soc. (2)*, 57(1):135–147, 1998.

[7] Enrico Bombieri and Walter Gubler. *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006.

[8] Emmanuel Breuillard and Péter P. Varjú. On the dimension of Bernoulli convolutions. *Ann. Probab.*, 47(4):2582–2617, 2019.

[9] Emmanuel Breuillard and Péter P. Varjú. Entropy of Bernoulli convolutions and uniform exponential growth for linear groups. *J. Anal. Math.*, 140(2):443–481, 2020.

[10] Changhao Chen. Self-similar sets with super-exponential close cylinders. *Ann. Fenn. Math.*, 46(2):727–738, 2021.

[11] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006.

[12] Gerald A. Edgar. *Integral, probability, and fractal measures*. Springer-Verlag, New York, 1998.

[13] De-Jun Feng and Huyi Hu. Dimension theory of iterated function systems. *Comm. Pure Appl. Math.*, 62(11):1435–1500, 2009.

[14] Michael Hochman. On self-similar sets with overlaps and inverse theorems for entropy. *Ann. of Math. (2)*, 180(2):773–822, 2014.

[15] Roger A. Horn. On the eigenvalues of a matrix with prescribed singular values. *Proc. Amer. Math. Soc.*, 5:4–7, 1954.

[16] Roger A. Horn and Charles R. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994.

[17] John E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981.

[18] Kurt Mahler. An application of Jensen's formula to polynomials. *Mathematika*, 7:98–100, 1960.

[19] Kurt Mahler. An inequality for the discriminant of a polynomial. *Michigan Math. J.*, 11:257–262, 1964.

[20] David Masser. *Auxiliary polynomials in number theory*, volume 207 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2016.

[21] Patrick Morandi. *Field and Galois Theory*. Springer New York, 1996.

[22] Ariel Rapaport. Proof of the exact overlaps conjecture for systems with algebraic contractions. *Ann. Sci. Éc. Norm. Supér. (4)*, 55(5):1357–1377, 2022.

[23] Ariel Rapaport. Dimension of self-conformal measures associated to an exponentially separated analytic IFS on $\mathbb{R}$. *Preprint, arXiv:2412.16753*, 2024.

[24] Ariel Rapaport and Péter P. Varjú. Self-similar measures associated to a homogeneous system of three maps. *Duke Math. J.*, 173(3):513–602, 2024.

[25] Pablo Shmerkin. On Furstenberg's intersection conjecture, self-similar measures, and the L^q norms of convolutions. *Ann. of Math. (2)*, 189(2):319–391, 2019.

[26] Károly Simon. Overlapping cylinders: the size of a dynamically defined Cantor-set. In *Ergodic theory of $\mathbf{Z}^d$ actions (Warwick, 1993–1994)*, volume 228 of *London Math. Soc. Lecture Note Ser.*, pages 259–272. Cambridge Univ. Press, Cambridge, 1996.

[27] Péter P. Varjú. Absolute continuity of Bernoulli convolutions for algebraic parameters. *J. Amer. Math. Soc.*, 32(2):351–397, 2019.

[28] Péter P. Varjú. On the dimension of Bernoulli convolutions for all transcendental parameters. *Ann. of Math. (2)*, 189(3):1001–1011, 2019.

[29] Zhiren Wang. Quantitative density under higher rank abelian algebraic toral actions. *Int. Math. Res. Not. IMRN*, (16):3744–3821, 2011.

Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong

Email address: djfeng@math.cuhk.edu.hk

Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong

Email address: zfeng@math.cuhk.edu.hk