BERNOULLI CONVOLUTIONS ASSOCIATED WITH CERTAIN NON-PISOT NUMBERS

DE-JUN FENG AND YANG WANG

1. Introduction

For any $0 < \lambda < 1$ let ν_λ denote the distribution of $\sum_{n=0}^{\infty} \varepsilon_n \lambda^n$ where the coefficients ε_n are either 0 or 1, chosen independently with probability $\frac{1}{2}$ for each.[1] It is the infinite convolution product of the distributions $\frac{1}{2}(\delta_0 + \delta_{\lambda^n})$, giving rise to the term "infinite Bernoulli convolution" or simply "Bernoulli convolution." The Bernoulli convolution can be expressed as a self-similar measure ν_λ satisfying the equation

$$\nu_\lambda = \frac{1}{2}\nu_\lambda \circ \phi_0^{-1} + \frac{1}{2}\nu_\lambda \circ \phi_1^{-1}, \tag{1.1}$$

where $\phi_0(x) = \lambda x$ and $\phi_1(x) = \lambda x + 1$. This measure has surprising connections with a number of areas in mathematics, such as harmonic analysis, fractal geometry, number theory, dynamical systems, and others, see [10].

One of the fundamental questions is for which values λ is the Bernoulli convolution ν_λ singular. The study of this questions goes back to the 1930's. It is not hard to show that for $0 < \lambda < \frac{1}{2}$ the measure is not absolutely continuous (Kershner and Wintner [7]), since the support of ν_λ is a Cantor set of Lebesgue measure 0. Jessen and Wintner [5] showed that ν_λ must be of pure type, i.e. it is either singular or absolutely continuous. The problem becomes much more interesting if $\frac{1}{2} \le \lambda < 1$. Erdös [2] proved that if $\lambda^{-1} < 2$ is a Pisot number, i.e. an algebraic integer whose algebraic conjugates are all inside the unit disk, then ν_λ is singular. This was done by showing that $\widehat{\nu_\lambda}(\xi) \not\to 0$ as $\xi \to \infty$. Such a

The first author is suported in part by the Special Funds for Major State Basic Research Projects in China.

The second author is supported in part by the National Science Foundation, grant DMS-0070586 and DMS-0139261.

[1] In most papers the coefficients ε_n are 1 or -1 instead of 0 and 1. But it is well known and easily shown that the two definitions are equivalent, with the resulting measures differ only by a suitable translation. We use 0 and 1 for simplicity.

technique is unable to produce other singular Bernoulli convolutions, as Salem [13] proved that $\widehat{\nu_\lambda}(\xi)\longrightarrow 0$ as $\xi\longrightarrow\infty$ for all other $\frac{1}{2}\le\lambda<1$.

In the opposite direction the first important result was due to Erdös [3], who proved that there exists an $a<1$ such that for almost all $\lambda\in(a,1)$ the Bernoulli convolution ν_λ is absolutely continuous. Later Kahane [6] (see also [10]) indicated that the argument of [3] actually implies that the Hausdorff dimension of the set of exceptional λ's in $(a,1)$ tends to 0 as $a\longrightarrow 1$. In 1995 Solomyak [14] proved the following fundamental theorem:

Theorem. (Solomyak 1995) *For almost all $\lambda\in(\frac{1}{2},1)$ the Bernoulli convolution ν_λ is absolutely continuous, with the density function $\frac{d\nu_\lambda}{dx}\in L^2(\mathbb{R})$.*

A simpler proof of this theorem was later given in Peres and Solomyak [11]. Peres and Schlag [9] showed as a corollary of a more general result that the Hausdorff dimension of the exceptional λ's in $[a,1)$ is strictly smaller than 1 for any $a>\frac{1}{2}$ (see also [10]). Another important work was due to Garsia [4], who proved that ν_λ is absolutely continuous with *bounded* density function if λ^{-1} is an algebraic integer whose algebraic conjugates are all outside the unit disk and whose minimal polynomial has constant term ± 2.

The reciprocal of Pisot numbers remain today the only known class of λ's in $(\frac{1}{2},1)$ for which ν_λ is singular. It raises the following fundamental question:

Open Question: *Is it true that if $\lambda\in(\frac{1}{2},1)$ and ν_λ is singular then λ^{-1} is a Pisot number?*

This question is far from being answered. Since ν_λ has its density function in $L^2(\mathbb{R})$ for almost all $\lambda\in(\frac{1}{2},1)$, one may ask a weaker question: *Are there any $\lambda\in(\frac{1}{2},1)$ such that ν_λ doesn't have an L^2 density and λ^{-1} is not Pisot?* Even this weaker question had not been answered. In fact, to our knowledge there had not been a published example of a non-Pisot type Bernoulli convolution ν_λ whose density is unbounded.

This paper addresses these questions. There appears to be a general belief that the best candidates for counter-examples — singular or non-L^2 density ν_λ — are the reciprocals of *Salem numbers.* A number ρ is a Salem number if it is an algebraic integer whose algebraic conjugates all have modulus no greater than 1, with at least one of which on the unit circle. In this paper, however, we construct Bernoulli convolutions ν_λ with density not in $L^2(\mathbb{R})$

for a class of λ's, whose reciprocals are neither Pisot nor Salem. We also study Bernoulli convolutions with unbounded density. The main theorem of ours is:

Theorem 1.1. *Let $\lambda_{n,k}$ denote the reciprocal of the largest real root of the polynomial $P_{n,k}(x) = x^n - x^{n-1} - \cdots - x^k - 1$. For any $k \geq 3$ there exists an $N(k) > 0$ such that for all $n \geq N(k)$ the density of the Bernoulli convolution $\nu_{\lambda_{n,k}}$, if it exists, is not in $L^2(\mathbb{R})$. Particularly if $k = 3$ then the density of $\nu_{\lambda_{n,3}}$, if it exists, is not in $L^2(\mathbb{R})$ for all $n \geq 17$.*

It is possible that $\nu_{\lambda_{n,k}}$ is singular, but our technique is unable to settle this question. Theorem 1.1 actually holds for all $k \geq 1$. However, for $k = 1$ or $k = 2$ the largest real roots of $P_{n,k}(x)$ are Pisot numbers. For $k \geq 3$ and large n the polynomial $P_{n,k}(x)$ has approximately $c_k n$ roots that are outside the unit disk, where c_k is the proportion of the unit circle $\{x \in \mathbb{C} : \ |x| = 1\}$ on which $|\frac{x^k - x + 1}{2 - x}| > 1$. This implies that for $k = 3$ approximately $\frac{n}{6}$ roots of $P_{n,k}(x)$ are outside the unit disk. We shall prove this fact in the appendix, as well as the following theorem:

Theorem 1.2. *For each $k \geq 3$ the largest real root $\lambda_{n,k}^{-1}$ of $P_{n,k}(x)$ is neither Pisot nor Salem for all sufficently large n. In particular, $\lambda_{n,3}^{-1}$ is neither Pisot nor Salem for $n \geq 14$.*

A well known class of Salem numbers are the largest real roots of the polynomials $Q_n(x) = x^n - x^{n-1} - \cdots - x + 1$, where $n \geq 4$ (see, e.g., [1, Theorem 5.3]). The reciprocals of these Salem numbers have been thought of as potential candidates for producing non-Pisot type singular Bernoulli convolutions. Our technique yields:

Theorem 1.3. *Let λ_n denote the reciprocal of the largest real root of the polynomial $Q_{n,k}(x) = x^n - x^{n-1} - \cdots - x + 1$. For any $\alpha > 3$ the density of the Bernoulli convolution ν_{λ_n}, if it exists, is not in $L^\alpha(\mathbb{R})$ for all sufficiently large n.*

Our next theorem concerns Bernoulli convolutions with unbounded densities. It is actually a corollary of a more general theorem in §3, and it presents a general construction for such Bernoulli convolutions.

Theorem 1.4. *Let $\frac{1}{2} < \lambda < 1$ be a real root of a $\{0, 1, -1\}$-polynomial of degree n (namely all its coefficients are 0, 1 or -1). Suppose that $\lambda < 2^{-\frac{n}{n+1}}$. Then the density of the Bernoulli convolution ν_λ, if it exists, is unbounded.*

We point out that there are many numbers satisfying the condition of the assumption of the above theorem. For example, the largest real root of polynomial $x^n - x^{n-1} - \cdots - x^p + P(x)$ satisfies the condition for every $\{0, 1, -1\}$-polynomial $P(x)$ with degree less than p and every sufficient large n.

While the fundamental question whether all singular Bernoulli convolutions ν_λ for $\lambda \in (\frac{1}{2}, 1)$ come from the reciprocals of Pisot numbers remain unresolved, our results may have hinted that perhaps there are other algebraic numbers among roots of $\{0, 1, -1\}$-polynomials that also give rise to singular Bernoulli convolutions. Another evidence comes from biased Bernoulli convolutions. A Bernoulli convolution is *biased* if the coefficients $\varepsilon_n \in \{0, 1\}$ in $\sum_{n=0}^\infty \varepsilon_n \lambda^n$ are not chosen with equal probabilty $\frac{1}{2}$. Let $\nu_{\lambda,p}$ denote the biased Bernoulli convolution that is the distribution of $\sum_{n=0}^\infty \varepsilon_n \lambda^n$, where $\varepsilon_n = 0$ and $\varepsilon_n = 1$ are chosen independently with probability p and $1 - p$, respectively. Peres and Solomyak [12] proved the following theorem:

Theorem (Peres and Solomyak 1998) The biased Bernoulli convolution $\nu_{\lambda,p}$ is singular if $\lambda < p^p(1-p)^{1-p}$. For any $\frac{1}{3} \leq p \leq \frac{2}{3}$, $\nu_{\lambda,p}$ is absolutely continuous for almost all $p^p(1-p)^{1-p} < \lambda < 1$.

The proof of the following theorem is essentially trivial. However, it provides a different angle for examining biased Bernoulli convolutions.

Theorem 1.5. *Let $\frac{1}{2} < \lambda < 1$ be any real root of a $\{0, 1, -1\}$-polynomial. Then there exists an interval $I \subset (0, 1)$ such that for any $p \in I$ we have $\lambda > p^p(1-p)^{1-p}$ and $\nu_{\lambda,p}$ is singular.*

We prove these theorems in §3. Along the way we establish other related results. In the appendix we study the roots of certain type of polynomials.

The first author is indebted to the generous support by the School of Mathematics, Georgia Institute of Technology, where he is currently a visitor. We also wish to thank Ka-Sing Lau for helpful discussions.

2. Some General Results

In this section we introduce some general results on self-similar measures. These results are used later on to prove the main theorems in the paper. We first consider general self-similar measures in $\mathbb{R}$, for which Bernoulli convolutions are an example. Let $\{\phi_j(x) := \lambda x +$

$b_j\}_{j=1}^m$ be an iterated functions system (IFS), where $|\lambda| < 1$ and $m > 1$. Let $p_1, \dots, p_m > 0$ with $\sum_{j=1}^m p_j = 1$. Then there is a unique compactly supported measure μ satisfying the equation

$$\mu = \sum_{j=1}^{m} p_j \mu \circ \phi_j^{-1}. \tag{2.1}$$

The measure μ is called the *self-similar measure assciated with the IFS* $\{\phi_j(x)\}$ *and probability weights* $\{p_j\}$.

Lemma 2.1. *Let μ be the self-similar measure given by (2.1).*

(a) *Suppose that $|\lambda| < \prod_{j=1}^m p_j^{p_j}$. Then μ is singular.*

(b) *For any $\alpha > 1$ suppose that $|\lambda| < \left(\sum_{j=1}^m p_j^\alpha\right)^{\frac{1}{\alpha-1}}$. Then the density of μ, if it exists, is not in $L^\alpha(\mathbb{R})$.*

(c) *Suppose that $p_{j_0} > |\lambda|$ for some $1 \le j_0 \le m$. Then the density of μ, if it exists, is not in $L^\alpha(\mathbb{R})$ for sufficiently large α. In particular, the density of μ, if it exists, is unbounded.*

Proof. (a) and (b) are given in Peres and Solomyak [12], Theorem 1.3. To prove (c), simply observe that $(\sum_{j=1}^m p_j^\alpha)^{\frac{1}{\alpha-1}}$ tends to $\max_j p_j$ as $\alpha \to \infty$. ∎

Going back to the Bernoulli convolutions, let $\phi_0(x) = \lambda x$ and $\phi_1(x) = \lambda x + 1$, where $0 < \lambda < 1$. Then the Bernoulli convolution ν_λ is the self-similar measure satisfying

$$\nu_\lambda = \frac{1}{2}\nu_\lambda \circ \phi_0^{-1} + \frac{1}{2}\nu_\lambda \circ \phi_1^{-1}. \tag{2.2}$$

We introduce some notations in symbolic space. Let $\mathcal{A} = \{0, 1\}$ be the alphabet. We use $\mathcal{A}^n$ to denote the set of words in $\mathcal{A}$ of length n, $n \ge 0$, and $\mathcal{A}^* := \bigcup_{n\ge 0} \mathcal{A}^n$. For $\mathbf{j} = j_0 j_1 \cdots j_{n-1} \in \mathcal{A}^n$ we denote $|\mathbf{j}| = n$ and $\mathbf{j}(k) = j_k$. Using these notations we iterate (2.2) n times to yield

$$\nu_\lambda = \sum_{\mathbf{j}\in\mathcal{A}^n} \frac{1}{2^n}\nu_\lambda \circ \phi_{\mathbf{j}}^{-1}, \tag{2.3}$$

where $\phi_{\mathbf{j}} := \phi_{j_0} \circ \phi_{j_1} \circ \cdots \circ \phi_{j_{n-1}}$ for any $\mathbf{j} = j_0 j_1 \cdots j_{n-1} \in \mathcal{A}^n$. We define an equivalence relation $\sim_\lambda$ on $\mathcal{A}^*$: For any $\mathbf{i}, \mathbf{j} \in \mathcal{A}^*$ we denote $\mathbf{i} \sim_\lambda \mathbf{j}$ if and only if $\phi_{\mathbf{i}} = \phi_{\mathbf{j}}$. Let $\Pi_\lambda(\mathbf{j})$ be the projection from $\mathcal{A}^*$ to $\mathbb{R}$ given by

$$\Pi_\lambda(\mathbf{j}) = j_0 + j_1\lambda + \cdots + j_{n-1}\lambda^{n-1}$$

for $\mathbf{j} = j_0 j_1 \cdots j_{n-1} \in \mathcal{A}^n$.

Lemma 2.2. *Let* $\mathbf{i}, \mathbf{j} \in \mathcal{A}^*$. *Then* $\mathbf{i} \sim_\lambda \mathbf{j}$ *if and only if* $|\mathbf{i}| = |\mathbf{j}|$ *and* $\Pi_\lambda(\mathbf{i}) = \Pi_\lambda(\mathbf{j})$.

Proof. The lemma follows immediately from the fact

$$\phi_{\mathbf{i}}(x) = \lambda^{|\mathbf{i}|}x + \Pi_\lambda(\mathbf{i}), \quad \phi_{\mathbf{j}}(x) = \lambda^{|\mathbf{j}|}x + \Pi_\lambda(\mathbf{j}).$$

∎

Suppose now the $\sim_\lambda$ equivalent classes in $\mathcal{A}^n$ are $\{\mathcal{A}_{n,k} :\ 1 \le k \le L\}$. Then (2.3) can be re-written as

$$\nu_\lambda = \sum_{k=1}^{L} \frac{|\mathcal{A}_{n,k}|}{2^n} \nu_\lambda \circ \phi_{\mathbf{j}_k}^{-1} \tag{2.4}$$

where $\mathbf{j}_k$ is any element in $\mathcal{A}_{n,k}$. This leads to the following corollary of Lemma 2.1.

Lemma 2.3. *Let* $0 < \lambda < 1$ *and* ν_λ *be the Bernoulli convolution.*

(a) *Suppose that*

$$n(\log_2 \lambda + 1) < \frac{1}{2^n} \sum_{k=1}^{L} |\mathcal{A}_{n,k}| \log_2(|\mathcal{A}_{n,k}|)$$

for some $n \ge 1$. *Then* ν_λ *is singular.*

(b) *For any* $\alpha > 1$ *suppose that*

$$(2\lambda)^{(\alpha-1)n} < \frac{1}{2^n} \sum_{k=1}^{L} |\mathcal{A}_{n,k}|^\alpha$$

for some $n \ge 1$. *Then the density of* ν_λ, *if it exists, is not in* $L^\alpha(\mathbb{R})$.

(c) *Suppose that* $\max_{1\le k\le L} |\mathcal{A}_{n,k}| > (2\lambda)^n$ *for some* $n \ge 1$. *Then the density of* ν_λ, *if it exists, is not in* $L^\alpha(\mathbb{R})$ *for sufficiently large* α. *In particular, the density of* ν_λ, *if it exists, is unbounded.*

Proof. For (a), apply Lemma 2.1 to (2.4) we see that ν_λ is singular if

$$\lambda^n < \prod_{k=1}^{L} \Big(\frac{|\mathcal{A}_{n,k}|}{2^n}\Big)^{\frac{|\mathcal{A}_{n,k}|}{2^n}}.$$

Taking base 2 logarithm and (a) follows immediately by observing that $\sum_{k=1}^{l} |\mathcal{A}_{n,k}| = 2^n$. Part (b) of this lemma comes directly from part (b) of Lemma 2.1, as does part (c). ∎

Lemma 2.4. *Let* $\{\mathcal{B}_{n,k}:\ 1\le k\le M\}$ *be a partition of* $\mathcal{A}^n$ *such that for any* k *the elements in* $\mathcal{B}_{n,k}$ *are* $\sim_\lambda$ *equivalent.*

(a) *Suppose that*

$$n(\log_2\lambda+1)<\frac{1}{2^n}\sum_{k=1}^{L}|\mathcal{B}_{n,k}|\log_2(|\mathcal{B}_{n,k}|)$$

for some $n\ge 1$*. Then* ν_λ *is singular.*

(b) *For any* $\alpha>1$ *suppose that*

$$(2\lambda)^{(\alpha-1)n}<\frac{1}{2^n}\sum_{k=1}^{L}|\mathcal{B}_{n,k}|^\alpha$$

for some $n\ge 1$*. Then the density of* ν_λ*, if it exists, is not in* $L^\alpha(\mathbb{R})$*.*

(c) *Suppose that* $\max_{1\le k\le L}|\mathcal{B}_{n,k}|>(2\lambda)^n$ *for some* $n\ge 1$*. Then the density of* ν_λ*, if it exists, is not in* $L^\alpha(\mathbb{R})$ *for sufficiently large* α*. In particular, the density of* ν_λ*, if it exists, is unbounded.*

Proof. We only need to show that

$$\sum_{k=1}^{M}|\mathcal{B}_{n,k}|\log_2(|\mathcal{B}_{n,k}|)\le\sum_{k=1}^{L}|\mathcal{A}_{n,k}|\log_2(|\mathcal{A}_{n,k}|)$$

and for any $\alpha>1$,

$$\sum_{k=1}^{M}|\mathcal{B}_{n,k}|^\alpha\le\sum_{k=1}^{L}|\mathcal{A}_{n,k}|^\alpha.$$

First we note that each $\mathcal{B}_{n,k}$ is contained in some $\mathcal{A}_{n,l}$, so each $\mathcal{A}_{n,l}$ is the disjoint union of some $\mathcal{B}_{n,k}$'s. Suppose that $\mathcal{A}_{n,l}$ is the union of $\mathcal{B}_{n,k_1},\dots,\mathcal{B}_{n,k_m}$. We have

$$\sum_{j=1}^{m}|\mathcal{B}_{n,k_j}|\log_2(|\mathcal{B}_{n,k_j}|)\le\sum_{j=1}^{m}|\mathcal{B}_{n,k_j}|\log_2(|\mathcal{A}_{n,k_j}|)=|\mathcal{A}_{n,l}|\log_2(|\mathcal{A}_{n,l}|).$$

This proves the first inequality. The second inequality follows from

$$\sum_{j=1}^{m}|\mathcal{B}_{n,k_j}|^\alpha\le|\mathcal{A}_{n,l}|^\alpha.$$

∎

3. Proof of Theorems

We first establish Theorem 1.3. Let ρ_n be the largest real root of the polynomial $Q_n(x) = x^n - x^{n-1} - \cdots - x + 1$, which is known to be a Salem number for $n \geq 4$. Let $\lambda_n = \rho_n^{-1}$. We show that for any $\alpha > 3$ the density of ν_{λ_n}, if it exists, is not in $L^\alpha(\mathbb{R})$ for sufficiently large n. Note that λ_n is also a root of $Q_n(x)$. Consider the two words $\mathbf{u}^n$ and $\mathbf{v}^n$ in $\mathcal{A}^{n+1}$ given by

$$\mathbf{u}^n = 100\cdots 01, \quad \mathbf{v}^n = 011\cdots 10. \tag{3.1}$$

Observe that $\mathbf{u}^n \sim_{\lambda_n} \mathbf{v}^n$ since $\Pi_{\lambda_n}(\mathbf{u}^n) = 1 + \lambda_n^n = \lambda_n + \cdots + \lambda_n^{n-1} = \Pi_{\lambda_n}(\mathbf{v}^n)$.

Lemma 3.1. *We have $\rho_n = 2 - \frac{3}{2^n} + O(\frac{n}{2^{2n}})$ and $\lambda_n = \frac{1}{2} + \frac{3}{2^{n+2}} + O(\frac{n}{2^{2n}})$.*

Proof. Multiplying $x-1$ by $Q_n(x)$ yields

$$(x-1)Q_n(x) = x^{n+1} - 2x^n + 2x - 1.$$

Hence $(2-\rho_n)\rho_n^n = 2\rho_n - 1$. It is easy to check that $\rho_n > 1.5$ for $n \geq 4$. This shows immediately that $\rho_n \longrightarrow 2$ as $n \longrightarrow \infty$, and $2 - \rho_n = \frac{2\rho_n - 1}{\rho_n^n} < 3 \times 1.5^{-n}$. Let $2 - \rho_n = \varepsilon_n$. Then $\varepsilon_n(2-\varepsilon_n)^n = 2(2-\varepsilon_n) - 1 = 3 - 2\varepsilon_n$. Now by the Mean Value Theorem, $(2-\varepsilon_n)^n = 2^n - n\varepsilon_n\, \xi_n^{n-1} \geq 2^n - n\, 2^{n-1}\varepsilon_n$. Hence

$$\begin{aligned} \varepsilon_n &\leq \frac{3 - 2\varepsilon_n}{2^n - n\, 2^{n-1}\varepsilon_n} \\ &= \frac{3-2\varepsilon_n}{2^n}\Big(1 + \frac{n\,\varepsilon_n}{2} + O(\frac{(n\,\varepsilon_n)^2}{2^2})\Big) \\ &= \frac{3}{2^n} + \epsilon_n \cdot \frac{3n-4}{2^{n+1}} + \frac{1}{2^n}\cdot O(n\,\varepsilon_n)^2 \\ &= \frac{3}{2^n} + o(\frac{1}{2^n}). \end{aligned} \tag{3.2}$$

Substituting ε_n on the right side of (3.2) with $\frac{3}{2^n} + o(\frac{1}{2^n})$ yields the estimate $\varepsilon_n \leq \frac{3}{2^n} + O(\frac{n}{2^{2n}})$. Using the fact $\varepsilon_n 2^n > \varepsilon_n(2-\varepsilon_n)^n = 3 - 2\varepsilon_n$ we obtain $\varepsilon_n \geq \frac{3}{2^n+2} = \frac{3}{2^n} - O(\frac{1}{2^{2n}})$. These two inequalities give our estimate for ρ_n. The estimate for λ_n comes directly by taking the reciprocal of ρ_n. ∎

We say that $\mathbf{w} \in \mathcal{A}^m$ *contains p times the subword $\mathbf{u}^n$ or $\mathbf{v}^n$* if the two words combine to appear in $\mathbf{w}$ exactly p times *nonoverlappingly.* This means, for example, the sequence $100\cdots 0100\cdots 01$ where in both places there are $(n-1)$ 0's counts only one appearance since the two $\mathbf{u}^n$'s overlap. The same goes for $10\cdots 011\cdots 10$.

Lemma 3.2. *Let $H_{m,p}$ be the number of elements in $\mathcal{A}^m$ that contain exact p times the subword $\mathbf{u}^n$ or $\mathbf{v}^n$. Then $H_{m,0} \geq 2^m \left(1 - \frac{m-n}{2^n}\right)$ and*

$$2^{m-np}\binom{m-pn}{p}\left(1-\frac{m-(n+1)p}{2^{n-2}}\right) \leq H_{m,p} \leq 2^{m-np}\binom{m-pn}{p} \tag{3.3}$$

for $p \geq 1$.

Proof. Denote by G_m be the number of elements in $\mathcal{A}^m$ that contain at least one of the subwords $\mathbf{u}^n$ and $\mathbf{v}^n$. If an element in $\mathcal{A}^m$ contains $\mathbf{u}^n$ or $\mathbf{v}^n$ at position j, then $0 \leq j \leq m-n-1$. For each such j, the numbers of elements in $\mathcal{A}^m$ containing $\mathbf{u}^n$ or $\mathbf{v}^n$ at position j is no more than $2 \cdot 2^{m-n-1} = 2^{m-n}$. Since there are only $m-n$ choices for j we conclude that $G_m \leq 2^{m-n}(m-n)$. Thus

$$H_{m,0} = 2^m - G_m \geq 2^m\left(1-\frac{m-n}{2^n}\right).$$

Now let $p \geq 1$. For any $\mathbf{w} \in \mathcal{A}^m$ that contains p times the subword $\mathbf{u}^n$ or $\mathbf{v}^n$, we may mark the positions where these subwords appear by integers $0 \leq k_1 < k_2 < \cdots < k_p < m$. To guarantee that the above marking is unique, we ask k_1 to be the first position in $\mathbf{w}$ for which $\mathbf{u}^n$ or $\mathbf{v}^n$ appears. And suppose the positions $k_1, \cdots, k_{i-1}$ have been marked, then k_i is the smallest position for which $\mathbf{u}^n$ or $\mathbf{v}^n$ appears and $k_i - k_{i-1} \geq n+1$. Clearly $m - k_p \geq n+1$ (recall that the indices begin with 0). Set $x_0 = k_1$, $x_p = m - k_p$ and $x_i = k_{i+1} - k_i$ for $1 \leq i < p$. Then $\sum_{i=0}^p x_i = m$ with $x_0 \geq 0$ and $x_i \geq n+1$ for $i \geq 1$. Hence the number of ways to choose the positions $k_1 < k_2 < \cdots < k_p$ equals the number of solutions to the above equation. This number is identical to the number of solutions for

$$y_0 + y_1 + \cdots y_p = m-(n+1)p, \quad y_i \geq 0, \tag{3.4}$$

after making a substitution $y_0 = x_0$ and $y_i = x_i - (n+1)$ for $i \geq 1$. For (3.4) it is well-known that the number of solutions is $\binom{m-np}{p}$.

For each chosen positions $k_1 < k_2 < \cdots < k_p$ for the subwords $\mathbf{u}^n$ and $\mathbf{v}^n$ there are at most $2^{m-(n+1)p}$ ways to choose the positions not occupied by these subwords. Since at each position k_i we may have either a $\mathbf{u}^n$ or a $\mathbf{v}^n$, we obtain

$$H_{m,p} \leq 2^{m-(n+1)p}\, 2^p \binom{m-pn}{p} = 2^{m-np}\binom{m-pn}{p}.$$

On the other hand, for each chosen positions $k_1 < k_2 < \cdots < k_p$ for the subwords $\mathbf{u}^n$ and $\mathbf{v}^n$ we may estimate the number of words in $\mathcal{A}^m$ such that there is another subword $\mathbf{u}^n$ or

$\mathbf{v}^n$ at some position j such that $j < k_1$ or $j > k_p+n$ or $k_{i-1} < j < k_i - n$ for some i between 2 and p. In such situation, if this subword has some overlaps with other subwords at k_i's, then the number of ovelapped letters does not exceeding 2. Therefore the freedom to choose the rest of the positions not occupied by these subwords is no more than $2^{m-(n+1)(p+1)+2}$. Since there are at most $m-(n+1)p$ ways to choose the index j, we have

$$\begin{aligned} H_{m,p} &\geq 2^{m-np}\binom{m-pn}{p} - (m-(n+1)p)\,2^{m-(n+1)(p+1)+2}\,2^{p+1}\binom{m-pn}{p} \\ &= 2^{m-np}\binom{m-pn}{p}\Big(1-\frac{m-(n+1)p+2}{2^{n-2}}\Big). \end{aligned}$$

∎

Proof of Theorem 1.3. We partition $\mathcal{A}^m$ into subsets $\{\mathcal{B}_{p,j}\}$, where each $\mathcal{B}_{p,j}$ is obtained as follows: There exist positions $k_1 < k_2 < \cdots < k_p < m$ such that all elements in $\mathcal{B}_{p,j}$ contain exactly p times the subword $\mathbf{u}^n$ or $\mathbf{v}^n$, and they are at the positions k_i's; furthermore they all have identical letters in the remaining positions not occupied by the p subwords of $\mathbf{u}^n$ or $\mathbf{v}^n$. Clearly $|\mathcal{B}_{p,j}| = 2^p$ and since $\Pi_{\lambda_n}(\mathbf{u}^n) = \Pi_{\lambda_n}(\mathbf{v}^n)$, all elements in $\mathcal{B}_{p,j}$ are $\sim_{\lambda_n}$ equivalent. For each fixed $p \geq 0$ there are exactly $2^{-p}H_{m,p}$ subsets $\mathcal{B}_{p,j}$'s, where $H_{m,p}$ is defined in Lemma 3.2.

We apply Lemma 2.4 to prove the theorem. Set $m = n^2$. We show that for any $\alpha > 3$ we have $(2\lambda_n)^{(\alpha-1)m} < \frac{1}{2^m}\sum_{p,j}|\mathcal{B}_{p,j}|^\alpha$ for sufficiently large n. It follows from Lemma 3.2 that

$$\begin{aligned} \frac{1}{2^m}\sum_{p,j}|\mathcal{B}_{p,j}|^\alpha &= \frac{1}{2^m}\sum_{p\geq 0} 2^{-p}\,H_{m,p}\,|\mathcal{B}_{p,j}|^\alpha \\ &\geq \frac{1}{2^m}\sum_{p=0}^{1} 2^{-p}\,H_{m,p}\,|\mathcal{B}_{p,j}|^\alpha \\ &= \frac{1}{2^m}H_{m,0} + \frac{1}{2^m}2^{\alpha-1}H_{m,1} \\ &\geq \Big(1-\frac{m-n}{2^n}\Big) + \frac{(m-n)\,2^{\alpha-1}}{2^n}\Big(1-\frac{m-(n+1)}{2^{n-2}}\Big) \\ &= 1 + \frac{(m-n)\,2^\alpha - 2m + 2n}{2^{n+1}} + O\Big(\frac{n^2}{2^{2n}}\Big). \end{aligned}$$

On the other hand, $2\lambda_n = 1 + \frac{3}{2^{n+1}} + O(\frac{n}{2^{2n}})$ by Lemma 3.1. This yields

$$(2\lambda_n)^{(\alpha-1)m} = \Big(1+\frac{3}{2^{n+1}} + O(\frac{n}{2^{2n}})\Big)^{(\alpha-1)m} = 1 + \frac{3(\alpha-1)m}{2^{n+1}} + O(\frac{n^3}{2^{2n}}).$$

Since $\alpha > 3$, we have $2^\alpha - 2 > 3(\alpha-1)$ and as $n\longrightarrow\infty$,

$$\frac{(m-n)\,2^\alpha - 2m + 2n}{2^{n+1}} + O\Big(\frac{n^2}{2^{2n}}\Big) > \frac{3(\alpha-1)m}{2^{n+1}} + O(\frac{n^3}{2^{2n}}).$$

The theorem follows from Lemma 2.4. ∎

The proof of Theorem 1.1 is along the same lines as the proof of Theorem 1.3. Let $\rho_{n,k}$ be the largest real root of the polynomial $P_{n,k}(x) = x^n - x^{n-1} - \cdots - x^k - 1$, and $\lambda_{n,k} = \rho_{n,k}^{-1}$. Clearly, $\lambda_{n,k}$ is a real root of $1 - x - \cdots - x^{n-k} - x^n$.

Lemma 3.3. *For any fixed $k \geq 1$ we have $\rho_{n,k} = 2 - \frac{2^k-1}{2^n} + O(\frac{n}{2^{2n}})$. For $k = 3$ we have $0 < 2 - \rho_{n,3} \leq \frac{7}{2^n} + \frac{60n}{2^{2n}}$ for $n \geq 12$.*

Proof. Multiplying $x-1$ by $P_{n,k}(x)$ yields

$$(x-1)P_{n,k}(x) = x^{n+1} - 2x^n + x^k - x + 1.$$

Hence

$$(2-\rho_{n,k}) = \frac{\rho_{n,k}^k - \rho_{n,k} + 1}{\rho_{n,k}^n}.$$

It is easy to check that $2 > \rho_{n,k} > 1.5$ for $n \geq k+2$. This shows immediately that $\rho_{n,k}\longrightarrow 2$ as $n\longrightarrow\infty$, and $2-\rho_{n,k} = \frac{\rho_{n,k}^k - \rho_{n,k}+1}{\rho_{n,k}^n}$ tends to 0 exponentially. Set $2-\rho_{n,k} = \varepsilon_{n,k}$. Then $\varepsilon_{n,k} = f(2-\varepsilon_{n,k})$ where $f(x) := \frac{x^k - x + 1}{x^n}$. It follows from the Mean Value Theorem that $f(2-\varepsilon_{n,k}) = f(2) - f'(\xi_{n,k})\varepsilon_{n,k}$ for some $2-\varepsilon_{n,k} < \xi_{n,k} < 2$. So we have

$$\begin{aligned}\varepsilon_{n,k} &= f(2) - f'(\xi_{n,k})\varepsilon_{n,k}\\ &= \frac{2^k-1}{2^n} - \frac{(n-k)\,\xi_{n,k}^{k-1} - n + (n+1)\,\xi_{n,k}^{-1}}{\xi_{n,k}^n}\,\varepsilon_{n,k}\\ &\leq \frac{2^k-1}{2^n} - \frac{C\,n\,\varepsilon_{n,k}}{\xi_{n,k}^n} \qquad (3.5)\end{aligned}$$

where $C \leq \xi_{n,k}^{k-1} < 2^{k-1}$. Iterating (3.5) by substituting $\varepsilon_{n,k}$ on the right side with $\frac{2^k-1}{2^n} + o(\frac{1}{2^n})$ yields immediately $\rho_{n,k} = 2 - \frac{2^k-1}{2^n} + O(\frac{n}{2^{2n}})$.

For $k=3$ we make a more delicate estimate. Clearly $P_{n,3}(2) > 0$. Let $x_n = 2 - \frac{7}{2^n} - \frac{60\,n}{2^{2n}}$. We show that $P_{n,3}(x_n) < 0$ for $n \geq 12$. Set $\varepsilon_n = 2 - x_n$. By Taylor expansions,

$$x_n^n \geq 2^n - n\,2^{n-1}\,\varepsilon_n, \quad x_n^3 - x_n + 1 \leq 7 - 11\varepsilon_n + 6\varepsilon_n^2.$$

Recall that $(x-1)P_{n,3}(x) = (x-2)x^n + x^3 - x + 1$. Hence

$$\begin{aligned}
(x_n - 1)P_{n,3}(x_n) &\le -\varepsilon_n\,(2^n - n\,2^{n-1}\,\varepsilon_n) + 7 - 11\varepsilon_n + 6\varepsilon_n^2 \\
&= 7 - \varepsilon_n\,(2^n - 11) + (n\,2^{n-1} + 6)\,\varepsilon_n^2 \\
&\le 7 - \varepsilon_n\,(2^n - 11) + (n\,2^n)\,\varepsilon_n^2 \\
&= 7 - \left(7 + \frac{60\,n}{2^n} - \frac{77}{2^n} - \frac{660}{2^{2n}}\right) + \left(\frac{49\,n}{2^n} + \frac{840\,n^2}{2^{2\,n}} + \frac{3600\,n^3}{2^{3\,n}}\right) \\
&= -\frac{11\,n}{2^n} + \frac{77}{2^n} + \frac{840\,n^2 + 660\,n}{2^{2\,n}} + \frac{3600\,n^3}{2^{3\,n}} \\
&< 0
\end{aligned}$$

for $n \ge 12$. Therefore $x_n < \rho_{n,3} < 2$ for $n \ge 12$, proving the lemma. ∎

We shall introduce the term *pattern* as a generalization of word. Basically a pattern is a word with possibly nonconsecutive indices. To rigorously define it, a pattern is a map $\tau : \mathcal{S} \longrightarrow \{0,1\}$ where $\mathcal{S}$ is a finite set of nonnegative integers. For each pattern $\tau : \mathcal{S} \longrightarrow \{0,1\}$ we may define

$$\Pi_\lambda(\tau) = \sum_{j\in\mathcal{S}} \tau(j)\lambda^j.$$

Associated with $\lambda_{n,k}$ are two patterns $\tau_{n,k}$ and $\bar\tau_{n,k}$ defined on $\mathcal{S}_{n,k} = \{0,1,\dots,n-k,n\}$ by $\tau_{n,k}(0) = 1$ and $\tau_{n,k}(j) = 0$ for all other $j \in \mathcal{S}_{n,k}$, and $\bar\tau_{n,k}(0) = 0$ and $\bar\tau_{n,k}(j) = 1$ for all other $j \in \mathcal{S}_{n,k}$. It is easy to check that $\Pi_{\lambda_{n,k}}(\tau_{n,k}) = \Pi_{\lambda_{n,k}}(\bar\tau_{n,k})$.

Let $\tau : \mathcal{S} \longrightarrow \{0,1\}$ be a pattern. We say $\mathbf{w} \in \mathcal{A}^m$ *contains the pattern* τ *at position* k if $\mathbf{w}(k+j) = \tau(j)$ for all $j \in \mathcal{S}$.

Lemma 3.4. *Let F_m be the number of elements in $\mathcal{A}^m$ that contains no pattern $\tau_{n,k}$ or $\bar\tau_{n,k}$, and G_m be the number of elements in $\mathcal{A}^m$ that contains the pattern $\tau_{n,k}$ or $\bar\tau_{n,k}$. Then*

$$\begin{aligned}
F_m &\ge 2^m\left(1 - \frac{m-n}{2^{n-k+1}}\right), \\
G_m &\ge 2^{m-n+k-1}\,(m-n)\left(1 - \frac{m-n+k}{2^{n-k}}\right).
\end{aligned}$$

Proof. If an element in $\mathcal{A}^m$ contains the pattern $\tau_{n,k}$ or $\bar\tau_{n,k}$ at position j, then $0 \le j \le m-n-1$. For each such j, the number of elements in $\mathcal{A}^m$ containing the pattern $\tau_{n,k}$ or $\bar\tau_{n,k}$ at position j is no more than $2 \cdot 2^{m-(n-k+2)} = 2^{m-n+k-1}$. Since there are only $m-n$ choices for j, we conclude that $G_m \le 2^{m-n+k-1}(m-n)$. Hence

$$F_m = 2^m - G_m \ge 2^m\left(1 - \frac{m-n}{2^{n-k+1}}\right). \tag{3.6}$$

To find a lower bound for G_m we count the number of elements in $\mathcal{A}^m$ that contain a pattern $\tau_{n,k}$ or $\bar{\tau}_{n,k}$ at position j but nowhere else. If an element $\mathbf{w}$ in $\mathcal{A}^m$ contains $\tau_{n,k}$ or $\bar{\tau}_{n,k}$ at position j and at another position i, then the set $(j+\mathcal{S}_{n,k})\cap(i+\mathcal{S}_{n,k})$ contains at most one element. This means that for any j, there are no more than $m-(n-k)$ choices for i; and the freedom to choose the rest of the positions other than $(j+\mathcal{S}_{n,k})\cup(i+\mathcal{S}_{n,k})$ are at most $m-2n+2k-3$. The argument for (3.6) now applies to show that the number of elements in $\mathcal{A}^m$ containing the pattern $\tau_{n,k}$ or $\bar{\tau}_{n,k}$ at position j is at least $2\big(2^{m-n+k-2}-(m-n+k)2\cdot 2^{m-2n+2k-3}\big)$. Since there are only $n-m$ choices for j, we obtain

$$G_m \geq 2^{m-n+k-1}\,(m-n)\Big(1-\frac{m-n+k}{2^{n-k}}\Big).$$

∎

Lemma 3.5. *Let a,b be two real numbers such that $0<a<1$ and $0<ab<1$. Then*

$$(1+a)^b < 1+ab+(e-2)a^2b^2.$$

Proof. Since $0<a<1$, we have $\log(1+a)<a$. This combines with $0<ab<1$ to yield

$$\begin{aligned}(1+a)^b &= e^{b\log(1+a)} \leq e^{ab} = 1+ab+\sum_{k=2}^{\infty}\frac{(ab)^k}{k!}\\ &\leq 1+ab+a^2b^2\left(\sum_{k=2}^{\infty}\frac{1}{k!}\right) = 1+ab+(e-2)a^2b^2.\end{aligned}$$

∎

Proof of Theorem 1.1. We partition $\mathcal{A}^m$ into subsets $\{\mathcal{B}_j\}$ and $\{\mathcal{C}_j\}$ such that each $\mathcal{C}_j$ contains a single element while each $\mathcal{B}_j$ contains two elements. The two elements in each $\mathcal{B}_j$ have the property that at some position i one contains the pattern $\tau_{n,k}$ at i while the other contains the pattern $\bar{\tau}_{n,k}$ at i; furthermore, they are identical in positions not reached by these two patterns. We may make the partition so that there are at least $\frac{1}{2}G_m$ subsets $\mathcal{B}_j$'s, and there are at least F_m subsets $\mathcal{C}_j$'s. Since $\Pi_{\lambda_{n,k}}(\tau_{n,k})=\Pi_{\lambda_{n,k}}(\bar{\tau}_{n,k})$, the two elements in each $\mathcal{B}_j$ are $\sim_{\lambda_{n,k}}$ equivalent.

Set $m=n^2$. For any $\alpha>0$,

$$\frac{1}{2^m}\Big(\sum_j|\mathcal{C}_j|^\alpha+\sum_j|\mathcal{B}_j|^\alpha\Big)\geq\frac{1}{2^m}\Big(F_m+\frac{1}{2}G_m\cdot 2^\alpha\Big).$$

It follows from Lemma 3.4 that

$$\frac{1}{2^m}F_m \geq 1 - \frac{m-n}{2^{n-k+1}}, \quad \frac{1}{2^m}G_m \geq \frac{m-n}{2^{n-k+1}} + O(\frac{n^4}{2^{2(n-k+1)}}).$$

Hence

$$\begin{aligned}\frac{1}{2^m}\Big(\sum_j |\mathcal{C}_j|^\alpha + \sum_j |\mathcal{B}_j|^\alpha\Big) &\geq 1 + \frac{(2^{\alpha-1}-1)(m-n)}{2^{n-k+1}} + O\Big(\frac{n^4}{2^{2(n-k+1)}}\Big) \\ &= 1 + \frac{(2^{\alpha-1}-1)\,2^k\,(m-n)}{2^{n+1}} + O\Big(\frac{n^4}{2^{2(n-k+1)}}\Big).\end{aligned}$$

On the other hand,

$$(2\lambda_{n,k})^{(\alpha-1)m} = 1 + (\alpha-1)m \cdot \frac{2^k-1}{2^{n+1}} + O\Big(\frac{n^3}{2^{2n}}\Big).$$

Suppose that $\alpha = 2$. Then $(2^{\alpha-1}-1)\,2^k > (\alpha-1)(2^k-1)$. Hence

$$\frac{1}{2^m}\Big(\sum_j |\mathcal{C}_j|^\alpha + \sum_j |\mathcal{B}_j|^\alpha\Big) > (2\lambda_{n,k})^{(\alpha-1)m}$$

for sufficiently large n. Hence the density of $\nu_{\lambda_{n,k}}$, if it exists, is not in $L^2(\mathbb{R})$ for sufficiently large n.

We now focus on the case $k = 3$. Taking $m = n^2$ and $\alpha = 2$, we have

$$\frac{1}{2^m}F_m \geq 1 - \frac{n^2-n}{2^{n-2}}, \quad \frac{1}{2^m}G_m \geq \frac{m-n}{2^{n-2}}\left[1 - \frac{n^2-n+3}{2^{n-3}}\right].$$

Thus

$$\frac{1}{2^m}(F_m + \frac{1}{2}G_m \cdot 2^2) \geq 1 + \frac{n^2-n}{2^{n-2}} - \frac{2(n^2-n)(n^2-n+3)}{2^{2n-5}}.$$

Assume that $n \geq 12$. Then

$$(2\lambda_{n,3})^{n^2} \leq \left(\frac{2}{2 - \frac{7}{2^n} - \frac{60}{2^{2n}}}\right)^{n^2} = \left(\frac{1}{1 - \frac{7}{2^{n+1}} - \frac{30}{2^{2n}}}\right)^{n^2}.$$

A direct check yields $(1 - \frac{7}{2^{n+1}} - \frac{30}{2^{2n}})(1 + \frac{7}{2^{n+1}} + \frac{45}{2^{2n}}) \geq 1$. Hence

$$\begin{aligned}(2\lambda_{n,3})^{n^2} &\leq \left(1 + \frac{7}{2^{n+1}} + \frac{45}{2^{2n}}\right)^{n^2} \\ &\leq 1 + n^2\left(\frac{7}{2^{n+1}} + \frac{45}{2^{2n}}\right) + (e-2)n^4\left(\frac{7}{2^{n+1}} + \frac{45}{2^{2n}}\right)^2.\end{aligned}$$

Another direct check shows that

$$\frac{n^2-n}{2^{n-2}} - \frac{2(n^2-n)(n^2-n+3)}{2^{2n-5}} \geq n^2\left(\frac{7}{2^{n+1}} + \frac{45}{2^{2n}}\right) + (e-2)n^4\left(\frac{7}{2^{n+1}} + \frac{45}{2^{2n}}\right)^2$$

for $n \geq 17$. This implies

$$\frac{1}{2^m}\Big(\sum_j |\mathcal{C}_j|^2 + \sum_j |\mathcal{B}_j|^2\Big) > (2\lambda_{n,3})^m$$

for $n \geq 17$. Therefore the density of $\nu_{\lambda_{n,3}}$, if it exists, is not in $L^2(\mathbb{R})$ for $n \geq 17$. ∎

Remark. For $k = 3$, by a more delicate estimate of G_{n^2} and an approximation of $(2\lambda_{n,3})^{n^2}$ by Matlab, we are able to show that the density of $\nu_{\lambda_n,3}$, if it exists, is not in $L^2(\mathbb{R})$ for $n = 15$ and 16.

We prove Theorem 1.4 by setting up a more general theorem concerning Bernoulli convolutions with unbounded densities. Let $P(x) = \sum_{k=0}^{n} \varepsilon_k x^k$ be a $\{0, 1, -1\}$-polynomial of degree n. Denote

$$\mathcal{S}_P = \{0 \leq k \leq n : \ \varepsilon_k \neq 0\}.$$

Then $\mathcal{S}_P$ is a non-empty set of non-negative integers with $0, n \in \mathcal{S}_P$. Denote by d_P the *packing density* of $\mathcal{S}_P$ in $\mathbb{N}$, i.e.,

$$d_P = \limsup_{\ell\to\infty} \frac{u_\ell}{\ell},$$

where u_ℓ is the largest cardinality of the possible sets B such that $\{\mathcal{S}_P + i\}_{i\in B}$ is a family of disjoint subsets of $\{0, 1, \ldots, \ell - 1\}$. For convenience, we call also d_P the *packing density* of P. It is clear $d_P \geq \frac{1}{n+1}$. Hence Theorem 1.4 is a direct corollary of the following Theorem.

Theorem 3.6. *Suppose $\rho > 1$ is any root of a $\{0, 1, -1\}$-polynomial $P(x)$ of degree n and packing density d_P. Suppose that $\rho > 2^{1-d_P}$. Then for $\lambda = \rho^{-1}$ the density of the Bernoulli convolution ν_λ, if it exists, is unbounded.*

Proof. Set $Q(x) = x^n P(x^{-1})$ and write $Q(x) = \sum_{k=0}^{n} \varepsilon_k x^k$. Then λ is a root of Q. Observe that $S_Q = n - S_P$. This means P and Q have the same packing density, $d_Q = d_P$. Define two patterns $\tau, \overline{\tau} : \mathcal{S}_Q \longrightarrow \{0,\ 1\}$ respectively by

$$\tau(k) = \max\{\varepsilon_k, 0\}, \qquad \overline{\tau}(k) = \max\{-\varepsilon_k, 0\}$$

for all $k \in \mathcal{S}_Q$. Then we have $\Pi_\lambda(\tau) = \Pi_\lambda(\overline{\tau})$, since $\tau(k) - \overline{\tau}(k) = \varepsilon_k$ for all $k \in \mathcal{S}_Q$.

Since $\rho > 2^{1-d_Q}$, by the definition of d_Q, we can find a large number ℓ and a set $B \subset \mathbb{N} \cup \{0\}$ of cardinality u_ℓ such that $\rho > 2^{1-\frac{u_\ell}{\ell}}$ and $\{\mathcal{S}_Q + i\}_{i\in B}$ is a family of disjoint subsets of $\{0, 1, \ldots, \ell - 1\}$.

We construct a subset $\mathcal{B}$ of words in $\mathcal{A}^\ell = \{0,1\}^\ell$, such that $\mathbf{w} \in \mathcal{B}$ if and only if $\mathbf{w}$ contains the pattern τ or $\overline{\tau}$ at each position $i \in B$, and $\mathbf{w}(k) = 0$ for all $k \in \{0,1,\ldots,\ell-1\} \setminus \bigcup_{i\in B} (\mathcal{S}_Q + i)$. In this setting we have $\Pi_\lambda(\mathbf{w}) = \Pi_\lambda(\mathbf{v})$ for all $\mathbf{w},\mathbf{v} \in \mathcal{B}$. Note that the cardinality of $\mathcal{B}$ is just $|\mathcal{B}| = 2^{|B|} = 2^{u_\ell}$. But $\rho > 2^{1-\frac{u_\ell}{\ell}}$. This yields

$$(2/\rho)^\ell = (2\lambda)^\ell < 2^{u_\ell} = |\mathcal{B}|.$$

Therefore by part (c) of Lemma 2.4, the density of ν_λ, if it exists, is not in $L^\alpha(\mathbb{R})$ for sufficient large α. ∎

We finally turn our attention to biased Bernoulli convolutions by proving Theorem 1.5. Let $1/2 < \lambda < 1$ be any root of a $\{0,1,-1\}$-polynomial $P(x) = \sum_{k=0}^n \varepsilon_k\, x^k$. Define two words $\mathbf{u} = i_0 i_1 \cdots i_n, \mathbf{v} = j_0 j_1 \cdots j_n \in \mathcal{A}^{n+1}$ by

$$i_k = \max\{\varepsilon_k, 0\}, \quad j_k = \max\{-\varepsilon_k, 0\}$$

for all $0 \le k \le n$. Then we have $i_k - j_k = \varepsilon_k$ for all $0 \le k \le n$, and thus $\Pi_\lambda(\mathbf{u}) = \Pi_\lambda(\mathbf{v})$.

Now let p be any real number in $(0,1)$. Suppose $\nu_{\lambda,p}$ is the biased Bernoulli convolution associated with λ and p. That is, $\nu_{\lambda,p}$ is the distribution of $\sum_{n=0}^\infty \delta_n \lambda^n$, where $\delta_n = 0$ and $\delta_n = 1$ are chosen independently with probability p and $1-p$, respectively. It is well known that $\nu_{\lambda,p}$ is the self-similar measure satisfying the equation

$$\nu_{\lambda,p} = \sum_{i=0}^{1} p_i\, \nu_{\lambda,p} \circ \phi_i^{-1}, \tag{3.7}$$

where $\phi_0(x) = \lambda\, x$, $\phi_1(x) = \lambda\, x + 1$, $p_0 = p$ and $p_1 = 1-p$.

Iterating (3.7) $n+1$ times we obtain

$$\nu_{\lambda,p} = \sum_{\mathbf{j}\in\mathcal{A}^{n+1}} p_{\mathbf{j}}\, \nu_{\lambda,p} \circ \phi_{\mathbf{j}}^{-1}, \tag{3.8}$$

where $p_{\mathbf{j}} = p_{j_0} p_{j_1} \cdots p_{j_n}$ and $\phi_{\mathbf{j}}(x) = \phi_{j_0} \circ \phi_{j_1} \circ \cdots \circ \phi_{j_n}$ for any $\mathbf{j} = j_0 j_1 \cdots j_n$. Since $\Pi_\lambda(\mathbf{u}) = \Pi_\lambda(\mathbf{v})$, we have $\phi_{\mathbf{u}} = \phi_{\mathbf{v}}$. Thus we can rewrite (3.8) as

$$\nu_{\lambda,p} = (p_{\mathbf{u}} + p_{\mathbf{v}})\nu_{\lambda,p} \circ \phi_{\mathbf{u}}^{-1} + \sum_{\mathbf{j}\in\mathcal{A}^{n+1}\setminus\{\mathbf{u},\mathbf{v}\}} p_{\mathbf{j}}\, \nu_{\lambda,p} \circ \phi_{\mathbf{j}}^{-1}.$$

By part (a) of Lemma 2.1, if

$$\frac{(p_{\mathbf{u}} + p_{\mathbf{v}})^{p_{\mathbf{u}}+p_{\mathbf{v}}}}{p_{\mathbf{u}}^{p_{\mathbf{u}}} p_{\mathbf{v}}^{p_{\mathbf{v}}}} \prod_{\mathbf{j}\in\mathcal{A}^{n+1}} p_{\mathbf{j}}^{p_{\mathbf{j}}} > \lambda^{n+1} \tag{3.9}$$

then $\nu_{\lambda,p}$ is singular. Note that

$$\begin{aligned}\log\Big(\prod_{\mathbf{j}\in\mathcal{A}^{n+1}} p_{\mathbf{j}}^{p_{\mathbf{j}}}\Big) &= \sum_{j_0j_1\cdots j_n\in\mathcal{A}^{n+1}} p_{j_0}p_{j_1}\cdots p_{j_n}\left(\log p_{j_0}+\log p_{j_1}+\cdots+\log p_{j_n}\right)\\ &= (n+1)(p\log p+(1-p)\log(1-p)),\end{aligned}$$

the condition (3.9) is equivalent to

$$p\log p+(1-p)\log(1-p)-\log\lambda+g(p)>0, \tag{3.10}$$

where $g(p):=\frac{1}{n+1}[(p_{\mathbf{u}}+p_{\mathbf{v}})\log(p_{\mathbf{u}}+p_{\mathbf{v}})-p_{\mathbf{u}}\log p_{\mathbf{u}}-p_{\mathbf{v}}\log p_{\mathbf{v}}]$.

Proof of Theorem 1.5. Note that the function $g(p)$ in (3.10) is a positive continuous function of p on $(0,1)$. Since $\lambda\in(1/2,1)$, we can always find an interval $I\subset(0,1)$ such that for any $p\in I$,

$$p\log p+(1-p)\log(1-p)-\log\lambda<0,$$

and in the mean time the inequality (3.10) holds, that is, $\nu_{\lambda,p}$ is singular. ∎

4. Appendix

We have mentioned in §1 that the polynomial $P_{n,3}(x)$ given in Theorem 1.1 has about $n/6$ roots outside the unit circle for sufficiently large n. In this appendix we prove a general result that will imply the claim.

Proposition 4.1. *Let $f(z)$ is a rational function in $\mathbb{C}$ such that f has neither zero nor pole on the unit circle $\{z\in\mathbb{C}:\ |z|=1\}$. Assume that $|f(z)|\not\equiv 1$ on the unit circle. Let $a(n),b(n)$ denote the number of zeros (counting multiplicity) of $z^n-f(z)$ outside and inside the unit circle, respectively. Then we have*

$$\lim_{n\to\infty}\frac{a(n)}{n}=\mathcal{L}\{\theta\in[0,1):\ |f(e^{2\pi i\theta})|>1\},\qquad \lim_{n\to\infty}\frac{b(n)}{n}=\mathcal{L}\{\theta\in[0,1):\ |f(e^{2\pi i\theta})|<1\},$$

where $\mathcal{L}$ denotes the Lebesgue measure.

Before proving the proposition we prove the following lemma.

Lemma 4.2. *Let $f(z)$ be analytic and $|f(z)|>1$ (resp. $0<|f(z)|<1$) on a neighborhood U of z_0, where $z_0\in\mathbb{C}$ and $|z_0|=1$. Then there exists an $a>0$ such that*

$$\lim_{n\to\infty}\frac{C_n}{n}\geq\mathcal{L}\left\{\theta\in[0,1):\ |e^{2\pi i\theta}-z_0|<a\right\}, \tag{4.1}$$

where C_n is the number of roots (counting multiplicity) of $z^n - f(z) = 0$ in $\{z \in \mathbb{C} : \ |z| > 1, \ |z - z_0| < a\} \subset U$ (resp. $\{z \in \mathbb{C} : \ |z| < 1, \ |z - z_0| < a\} \subset U$).

Proof. Write $z_0 = e^{2\pi i\theta_0}$. Pick a $\delta > 0$ such that the region

$$\Omega = \left\{z = re^{2\pi i\theta} : \ 1 < r < \delta + 1, |\theta - \theta_0| < \delta\right\}$$

is contained in U and on which $|f(z) - f(z_0)| < |f(z_0)| - 1$.

Let $a = |e^{2\pi i\delta} - 1|$ and select $0 < \delta' < \delta$, $0 < \eta < \delta'$ such that the region

$$\{z = re^{2\pi i\theta} : \ 1 < r < 1 + \eta, \ |\theta - \theta_0| < \delta'\}$$

is contained in $\{z \in \mathbb{C} : |z| > 1, \ |z - z_0| < a\}$.

Let n be a large integer such that $(1 + \eta)^n > 2|f(z_0)|$. Choose θ_1, θ_2 such that

$$\theta_0 - \delta' < \theta_1 < \theta_0 - \delta' + \frac{2}{n}, \quad \theta_0 + \delta' - \frac{2}{n} < \theta_2 < \theta_0 + \delta'$$

and

$$e^{2\pi i n\theta_1} = e^{2\pi i n\theta_2} = -\frac{f(z_0)}{|f(z_0)|}.$$

Then it can be checked directly that $|f(z) - f(z_0)| < |f(z_0)| - 1 \leq |z^n - f(z_0)|$ on the boundary of the region

$$D' = \left\{z = re^{2\pi i\theta} : \ 1 < r < 1 + \eta, \ \theta_1 < \theta < \theta_2\right\}.$$

It follows from Rouche's Theorem that $z^n - f(z)$ has the same number of zeros in D' as does $z^n - f(z_0)$; and this number equals $n(\theta_2 - \theta_1) - 1$ exactly. Therefore the number of zeros of $z^n - f(z)$ in $\{z \in \mathbb{C} : |z| > 1, |z - z_0| < a\}$ is at least $n(\theta_2 - \theta_1) - 1$, and thus no less than $2n\delta' - 5$. The inequality (4.1) now follows by letting $\delta' \to \delta$. ∎

Proof of Proposition 4.1. Since $|f(e^{2\pi i\theta})|$ is real analytic in θ and $|f(e^{2\pi i\theta})| \not\equiv 1$ on the unit circle, the set $\{\theta \in [0, 1) : \ |f(e^{2\pi i\theta})| = 1\}$ is a finite set. Thus to prove Proposition 4.1 it suffices to prove

$$\lim_{n\to\infty} \frac{a(n)}{n} \geq \mathcal{L}\{\theta \in [0, 1) : \ |f(e^{2\pi i\theta})| > 1\}, \qquad \lim_{n\to\infty} \frac{b(n)}{n} \geq \mathcal{L}\{\theta \in [0, 1) : \ |f(e^{2\pi i\theta})| < 1\}.$$

We prove the first inequality here; the second one follows from an essentially identical argument.

Let $U = \{\theta \in [0, 1) : \ |f(e^{2\pi i\theta})| > 1\}$. U is an open set. By Lemma 4.2, for each $z_0 = e^{2\pi i\theta_0} \in U$ and each $\epsilon > 0$ there exists an $0 < \delta < \epsilon$ such that $\{\theta \in [0, 1) : \ |\theta - \theta_0| < \delta\} \subset U$, and in the region $\{z \in \mathbb{C} : \ |z| > 1, \ |z - e^{2\pi i \theta_0}| < |1 - e^{2\pi i \delta}|\}$ the number of zeros $C(n)$ of $z^n - f(z)$ satisfies

$$\lim_{n \to \infty} \frac{C(n)}{n} \geq 2\delta = \mathcal{L}\{\theta \in [0,1) : \ |\theta - \theta_0| < \delta\}.$$

By the Vitali Covering Theorem, for each $\eta > 0$ there exist finitely many disjoint intervals $I_i = [\theta_i - \delta_i, \theta_i + \delta_i] \subset U$ such that $\sum_i \mathcal{L}(I_i) > \mathcal{L}(U) - \eta$, and the number of zeros $C_i(n)$ of $z^n - f(z)$ in the region $\{z \in \mathbb{C} : \ |z| > 1, \ |z - e^{2\pi i \theta_i}| < |1 - e^{2\pi i \delta_i}|\}$ satisfies $\lim_{n \to \infty} \frac{C_i(n)}{n} \geq \mathcal{L}(I_i)$. Now let $C^*(n)$ denote the number of zeros of $z^n - f(z)$ in the disjoint union $\bigcup_i \{z \in \mathbb{C} : \ |z| > 1, \ |z - e^{2\pi i \theta_i}| < |1 - e^{2\pi i \delta_i}|\}$ satisfies

$$\lim_{n \to \infty} \frac{C^*(n)}{n} = \lim_{n \to \infty} \sum_i \frac{C_i(n)}{n} \geq \sum_i \mathcal{L}(I_i) = \mathcal{L}(U) - \eta.$$

The inequality follows by letting $\eta \longrightarrow 0$. ∎

Lemma 4.3. *Let $k \geq 3$ and $f_k(x) = \frac{x^k - x + 1}{x - 2}$. Then*

$$\mathcal{L}\Big\{\theta \in [0,1) : \ |f_k(e^{2\pi i \theta})| > 1\Big\} > 0.$$

Proof. Notice that $|f_k(e^{i\theta})| = \frac{3 - 2\cos\theta - 2\cos((k-1)\theta) + 2\cos(k\theta)}{5 - 4\cos\theta}$. Set

$$h_k(\theta) = \Big(3 - 2\cos\theta - 2\cos((k-1)\theta) + 2\cos(k\theta)\Big) - \Big(5 - 4\cos\theta\Big).$$

For $\theta_j = \frac{j\pi}{k-1}$ where j is odd we have $h_k(\theta_j) = 0$ (and hence $|f_k(e^{i\theta_j})| = 1$). However, it is easy to check that $h_k'(\theta_j) \neq 0$. Therefore $|f_k(e^{i\theta})| > 1$ in a neighborhood on one side of θ_j, proving the lemma. ∎

We remark that the above proof shows that for $k \geq 4$ we may find $I_k \subset [0,1)$ with $|f_k(e^{i\theta})| > 1$ for $\theta \in I_k$ and $\mathrm{Re}\,(e^{i\theta}) < 0$ by taking a suitable odd j. For $\theta \in I_k$ we have $|2 - e^{i\theta}| \geq \sqrt{5}$.

Lemma 4.4. *For each $k \geq 3$ the polynomial $P_{n,k}(x) = x^n - x^{n-1} - \cdots - x^k - 1$ has exactly one root in the region $\Re(x) > 1$.*

Proof. We use the equality $(x-1)P_{n,k}(x) = F_{n,k}(x)$ where $F_{n,k}(x) = (x-2)x^n - x^k + x - 1$. Set $Q_n(x) = (x-2)x^n$. We show that $F_{n,k}$ has one root in the region $\Re(x) > a$ for any $1 < a < 1.5$ and sufficiently large n (independent of a) by Rouche's Theorem. Note that $|F_{n,k}(x) - Q_n(x)| = |x^k - x + 1|$. On the boundary $\Re(x) = a$ write $x = a + iy$ with $y \in \mathbb{R}$, and one can check that $|x^k - x + 1|^2 = (a^k - a + 1)^2 + y^2 g_k(y)$ where g_k is a polynomial. On the other hand it is a easy calculation that

$$|Q_n(x)|^2 \geq (2-a)^2 a^{2n} + n(2-a)^2 y^2 + na^{2n-2} y^{2n+2}.$$

Comparing the derivatives we see that $(2-a)^2 a^{2n} > (a^k - a + 1)^2$ for $1 < a < 1.5$ and sufficenly large n. Also clear is that

$$n(2-a)^2 y^2 + na^{2n-2} y^{2n+2} \geq y^2 g_k(y)$$

for sufficently large n (independent of a). Finally $|Q_n(x)| > |x^k - x + 1|$ as $|x|$ tends to ∞ whenever $n \geq k$. Therefore Rouche's Theorem applies to show that $F_{n,k}(x)$ and $Q_n(x)$ have the same number of roots in $\Re(x) > a$ for sufficently large n independent of a. It follows that they have the same number of roots in $\Re(x) > 1$ by letting $a \to 1$. Since Q_n has the only root $x = 2$ in the region and $F_{n,k}(x) = (x-2)P_{n,k}$ we conclude that $P_{n,k}$ has exactly one root in $\Re(x) > 1$.

Note that for $k = 3$ we have checked that $|Q_n(x)| > |x^3 - x + 1|$ on $\Re(x) = a$ whenever $n > 10$ for $1 < a < 1.5$ (we omit the tedious details). So $P_{n,3}(x)$ has one root in $\Re(x) > 1$ for $n > 10$. ∎

Lemma 4.5. *$P_{n,3}(x)$ has at least two roots in the region $\Omega = \{x = re^{i\theta} : r > 1, \pi/2 \leq \theta < \pi\}$ for $n \geq 13$.*

Proof. We prove that $P_{n,3}(x)$ has at least two roots in Ω for a $n \geq 100$, and check the rest with Matlab. The proof for $n \geq 100$ only uses rather crude estimates, at times with computations performed by Matlab.

Note that $\frac{x-1}{2-x} P_{n,3}(x) = x^n - \frac{x^3-x+1}{2-x}$. We prove that $x^n - \frac{x^3-x+1}{2-x}$ has at least two roots in Ω. Set $F(x) = \frac{x^3-x+1}{2-x}$. Then for $\theta_0 = 1.8$ and $x_0 = e^{i\theta_0}$ we may check that $F(x_0) = c_0$ with $|c_0| > 1.05$. We now construct the small sector neighborhood R_n of x_0, given as

$$R_n := \{x = re^{i\theta} : 1 - \varepsilon_n < r < 1 + \varepsilon_n, \theta_1 \leq \theta < \theta_2\}$$

where $\varepsilon_n = \frac{2}{n}$ and θ_1, θ_2 satisfy $e^{n\theta_1} = e^{n\theta_2} = ic_0$ with $\theta_2 - \theta_1 = \frac{4\pi}{n}$ and $\theta_1 < \theta_0 < \theta_2$. These choices ensure that the polynomial $G_n(x) = x^n - c_0$ has two roots in R_n. Furthermore, we will apply Rouche's theorem to show that $x^n - F(x)$ and $G_n(x)$ have the same number of roots in R_n.

To see Rouche's theorem applies, we have $|x^n - F(x) - G_n(x)| = |F(x) - c_0|$. Let $n \geq 100$. We only need to prove that $|F(x) - c_0| < |G_n(x)|$ on ∂R_n. A rough estimate with the aide of Matlab yields $|F'(x)| = |2x + 2 - \frac{7}{(2-x)^2}| < 3.8$ on R_n. Another simple estimate yields diam $(R_n) < \frac{13.2}{n}$. So

$$|F(x) - c_0| < 3.8\,|x - x_0| < 3.8\,\mathrm{diam}\,(R_n) < \frac{51}{n}.$$

We will estimate $G_n(x)$ on ∂R_n. On the two straight edges of ∂R_n we have $x = re^{i\theta}$ with either $\theta = \theta_1$ pr $\theta = \theta_2$. Hence $x^n = r^n e^{in\theta} = ir^n c_0$. Therefore $|G_n(x)| = |ir^n c_0 - c_0| > |c_0| > 1$. On the outer arc $r = 1 + \varepsilon_n$ with $\varepsilon_n = \frac{2}{n}$ we also have

$$|G_n(x)| > |x|^n - |c_0| = (1 + \varepsilon_n)^n - |c_0| > 1.$$

Finally on the inner arc $r = 1 - \varepsilon_n$ we have

$$|G_n(x)| > |c_0| - |x|^n = |c_0| - (1 - \varepsilon_n)^n > 0.8.$$

It follows that $|G_n(x)| > |F(x) - c_0|$ on ∂R_n for $n \geq 100$ (in fact $n \geq 64$), and Rouche's theorem now applies to show that $x^n - F(x)$ and $G_n(x)$ have the same number of roots in R_n. Clearly, $G_n(x)$ has exactly two roots in R_n, and so does $x^n - F(x)$. But it is easily checked by Matlab that $|F(x)| > 1$ on R_n for $n \geq 100$. Hence the roots of $x^n - F(x)$ in R_n cannot be inside or on the unit circle. Therefore the two roots of $x^n - F(x)$ in R_n must be outside the unit circle. This proves the lemma for $n \geq 100$.

For $13 \leq n < 100$ the lemma is checked using Matlab. We omit the computational details here. ∎

Proof of Theorem 1.2. Assume that $\lambda_{n,k}$ is Pisot. Then $P_{n,k}(x) = f_{n,k}(x) g_{n,k}(x)$ where both $f_{n,k}$ and $g_{n,k}$ are monic polynomials in $\mathbb{Z}x$ with $f_{n,k}$ being the minimal polynomial of $\lambda_{n,k}$. Thus all roots of $P_{n,k}$ that are outside the unit circle are roots of $g_{n,k}$. We derive a contradiction for any k and sufficiently large n.

Note that $P_{n,k}(2) = f_{n,k}(2) g_{n,k}(2) = 2^k - 1$. Therefore $|g_{n,k}(2)| \leq 2^k - 1$. We consider $|g_{n,k}(2)| = |\prod_j (2 - r_j)|$ where r_j are the roots of $g_{n,k}$. By Lemma 4.4 all r_j are in the region $Re(x) < 1$, making $|2 - r_j| > 1$. On the other hand, it follows from Lemmas 4.2 and 4.3 that for any k $g_{n,k}$ has at least $3k$ roots in the region $\{x \in \mathcal{C} : |x| > 1, \Re(x) < 0\}$ for sufficiently large n. With these roots we have $|2 - r_j| > \sqrt{5}$. Hence $|g_{n,k}(2)| > (\sqrt{5})^{3k} > 2^k$.

This is a contradiction. It follows that for each k and sufficiently large n (depending on k) $\lambda_{n,k}$ is not Pisot.

For $k = 3$ and $n \geq 13$ assume that $\lambda_{n,3}$ is Pisot. We prove that $|g_{n,3}(2)| \leq 2^3 - 1$ is impossible. We have already established in Lemma 4.5 that $g_{n,3}$ has at least two roots in the region $\{x = re^{i\theta} : r > 1, \theta \in [\pi/2, \pi)\}$. The congugates of these roots yield two or more roots in $\{x = re^{i\theta} : r > 1, \theta \in [\pi, 3\pi/2)\}$. So $g_{n,3}$ has at least 4 roots in the region $\{x \in \mathcal{C} : |x| > 1, \Re(x) < 0\}$. It follows that $|g_{n,3}(2)| > (\sqrt{5})^4 > 7$, a contradiction. ∎

References

[1] P. Borwein and K. G. Hare, Some computations on the spectra of Pisot and Salem numbers. *Math. Comp.* **71** (2002), 767–780.
[2] P. Erdös, On a family of symmetric Bernoulli convolutions. *Amer. J. Math.* **61** (1939), 974–976.
[3] P. Erdös, On the smoothness properties of a family of Bernoulli convolutions. *Amer. J. Math.* **62** (1940), 180–186.
[4] A. M. Garsia, Arithmetic properties of Bernoulli convolutions. *Trans. Amer. Math. Soc.* **102** (1962), 409–432.
[5] B. Jessen and A. Wintner, Distribution functions and the Riemann zeta function. *Trans. Amer. Math. Soc.* **38** (1935), 48–88.
[6] J.-P. Kahane, Sur la distribution de certaines séries aléatoires. (French) Colloque de Théorie des Nombres (Univ. Bordeaux, Bordeaux, 1969), *Bull. Soc. Math. France*, Mémoire. **25** (1971), 119–122.
[7] R. Kershner, A. Winter, On symmetric Bernoulli convolutions. *Amer. J. Math.* **57** (1935), 541–548. (Taipei, 1998), 35–76, World Sci. Publishing, River Edge, NJ, 1999.
[8] R. D. Mauldin and K. Simon, The equivalence of some Bernoulli convolutions to Lebesgue measure. *Proc. Amer. Math. Soc.* **126** (1998), 2733–2736.
[9] Y. Peres and W. Schlag, Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions. *Duke Math. J.* **102** (2000), 193–251.
[10] Y. Peres, W. Schlag and B. Solomyak, Sixty years of Bernoulli convolutions. Fractal geometry and stochastics, II (Greifswald/Koserow, 1998), 39–65, Progr. Probab., 46, Birkhäuser, Basel, 2000.
[11] Y. Peres and B. Solomyak, Absolute continuity of Bernoulli convolutions, a simple proof. *Math. Res. Lett.* **3** (1996), 231–239.
[12] Y. Peres and B. Solomyak, Self-similar measures and intersections of Cantor sets. *Trans. Amer. Math. Soc.* **350** (1998), 4065–4087.
[13] R. Salem, Algebraic numbers and Fourier analysis. D. C. Heath and Co., Boston, 1963.
[14] B. Solomyak, On the random series $\sum \pm\lambda^n$ (an Erdös problem). *Ann. of Math. (2)* **142** (1995), 611–625.

Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, P. R. China
E-mail address: dfeng@math.tsinghua.edu.cn

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA.
E-mail address: wang@math.gatech.edu