

# Equidistribution results for sequences of polynomials

Baker, Simon

DOI:

[10.1016/j.jnt.2020.01.003](https://doi.org/10.1016/j.jnt.2020.01.003)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

*Document Version*

Peer reviewed version

*Citation for published version (Harvard):*

Baker, S 2020, 'Equidistribution results for sequences of polynomials', *Journal of Number Theory*, vol. 215, pp. 1-19. <https://doi.org/10.1016/j.jnt.2020.01.003>

[Link to publication on Research at Birmingham portal](#)

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# EQUIDISTRIBUTION RESULTS FOR SEQUENCES OF POLYNOMIALS

SIMON BAKER

ABSTRACT. Let  $(f_n)_{n=1}^\infty$  be a sequence of polynomials and  $\alpha > 1$ . In this paper we study the distribution of the sequence  $(f_n(\alpha))_{n=1}^\infty$  modulo one. We give sufficient conditions for a sequence  $(f_n)_{n=1}^\infty$  to ensure that for Lebesgue almost every  $\alpha > 1$  the sequence  $(f_n(\alpha))_{n=1}^\infty$  has Poissonian pair correlations. In particular, this result implies that for Lebesgue almost every  $\alpha > 1$ , for any  $k \geq 2$  the sequence  $(\alpha^{n^k})_{n=1}^\infty$  has Poissonian pair correlations.

## 1. INTRODUCTION

Given a sequence of real numbers of some number theoretic or dynamical origin, describing its distribution modulo one is a classical problem (see for example [7, 8, 15] for more on this topic). One approach for describing the distribution of a sequence modulo one is to ask whether it is uniformly distributed. In what follows we let  $\{\cdot\}$  denote the fractional part of a real number and  $\|\cdot\|$  denote the distance to the nearest integer. We say that a sequence  $(x_n)_{n=1}^\infty$  is uniformly distributed modulo one if for every pair of real numbers  $u, v$  with  $0 \leq u < v \leq 1$  we have

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq n \leq N : \{x_n\} \in [u, v]\}}{N} = v - u.$$

In recent years there has been much interest in a new approach for describing the finer distributional properties of a sequence modulo one. We say that a sequence  $(x_n)_{n=1}^\infty$  has Poissonian pair correlations if for all  $s > 0$  we have

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq m \neq n \leq N : \|x_n - x_m\| \leq \frac{s}{N}\}}{N} = 2s.$$

The original motivation for investigating whether a sequence has Poissonian pair correlations comes from a connection with quantum physics. For certain quantum systems the discrete energy spectra has the form  $(\{a_n \alpha\})_{n=1}^\infty$  where  $\alpha$  is a constant and  $(a_n)_{n=1}^\infty$  is a

---

*Date:* 2nd January 2020.

*2010 Mathematics Subject Classification.* 11K06.

*Key words and phrases.* Uniform distribution, Poissonian pair correlations.

sequence of integers. The Berry-Tabor conjecture states that the discrete energy spectrum has Poissonian pair correlations except for in certain degenerate cases. This connection inspired several important contributions due to Rudnick, Sarnak, and Zaharescu, see [19, 20, 21]. We refer the reader to [1] and the references therein for more on this connection between quantum physics and the Poissonian pair correlation property.

Much of the recent interest surrounding whether a sequence  $(a_n \alpha)_{n=1}^{\infty}$  has Poissonian pair correlations comes from a connection with additive combinatorics, and more specifically with the so called additive energy of a sequence  $(a_n)_{n=1}^{\infty}$ . This connection was initially observed by Aistleitner et al in [5] and subsequently pursued by several authors. For more on this connection we refer the reader to the survey of Larcher and Stockinger [16] and the references therein. We remark that the sequence  $(n\alpha)_{n=1}^{\infty}$  does not have Poissonian pair correlations for any  $\alpha \in \mathbb{R}$ . This fact can be seen as a consequence of the three gap theorem. For a short proof of this fact we refer the reader to the aforementioned survey of Larcher and Stockinger [16].

An interesting family of sequences is obtained by considering  $(\alpha^n)_{n=1}^{\infty}$  for  $\alpha > 1$ . The main source of motivation behind the present work is a desire to obtain a thorough description of the distribution of these sequences. More generally, we are interested in taking a sequence of polynomials  $(f_n)_{n=1}^{\infty}$ , a real number  $\alpha > 1$ , and studying the distribution of the sequence  $(f_n(\alpha))_{n=1}^{\infty}$  modulo one.

The study of the distributional properties of  $(\alpha^n)_{n=1}^{\infty}$  modulo one dates back to work of Hardy. In [12] he proved that if  $\alpha$  is an algebraic number and  $\lim_{n \rightarrow \infty} \|\alpha^n\| = 0$  then  $\alpha$  is a Pisot number. This result was later obtained independently by Pisot in [17]. Recall that we say a real number  $\alpha > 1$  is a Pisot number if it is an algebraic integer whose Galois conjugates all have modulus strictly less than one. Pisot had previously shown in [18] that there are at most countably many  $\alpha > 1$  satisfying  $\lim_{n \rightarrow \infty} \|\alpha^n\| = 0$ . It is a long-standing open question to determine whether there exist any transcendental numbers satisfying  $\lim_{n \rightarrow \infty} \|\alpha^n\| = 0$ . The main result of this paper builds upon the following theorem due to Koksma.

**Theorem 1.1** ([14]). *For Lebesgue almost every  $\alpha > 1$  the sequence  $(\alpha^n)_{n=1}^{\infty}$  is uniformly distributed modulo one.*

For some recent results on the distribution of the sequence  $(\alpha^n)_{n=1}^{\infty}$  we refer the reader to [2, 6, 9, 10, 11, 13], [8, Chapters 2 and 3], and the references therein.

In [4] it was shown that if a sequence has Poissonian pair correlations then it is uniformly distributed modulo one. Observe that by our earlier remarks regarding the sequence  $(n\alpha)_{n=1}^{\infty}$ , and the well known fact that  $(n\alpha)_{n=1}^{\infty}$  is uniformly distributed if  $\alpha$  is irrational,

it follows that having Poissonian pair correlations is a stronger property than being uniformly distributed. With this observation and Theorem 1.1 in mind, the following question naturally arises.

**Question 1.2.** Is it true that for Lebesgue almost every  $\alpha > 1$  the sequence  $(\alpha^n)_{n=1}^\infty$  has Poissonian pair correlations?

In this paper we do not answer this question. It is worth mentioning that after completion of this paper, the author and Christoph Aistleitner were able to answer Question 1.2 in the affirmative, see [3]. The arguments used in this paper rely on a second moment method and good estimates on the Lebesgue measure of certain sets. The arguments used in [3] combine some of the techniques introduced in this paper with a martingale approach and techniques from Fourier Analysis. The results from [3] do not imply either Theorem 1.3 or Theorem 1.4, which are the main results of this paper.

The main result of this paper is the following general theorem which gives sufficient conditions for a sequence of polynomials  $(f_n)_{n=1}^\infty$  to ensure that for Lebesgue almost every  $\alpha > 1$  the sequence  $(f_n(\alpha))_{n=1}^\infty$  has Poissonian pair correlations.

**Theorem 1.3.** *Suppose  $(f_n)_{n=1}^\infty$  is a sequence of polynomials satisfying the following properties:*

- (1) *The sequence  $(\deg(f_n))_{n=1}^\infty$  is strictly increasing.*
- (2) *For any  $n_2 > n_1$  the function  $f_{n_2} - f_{n_1} : (1, \infty) \rightarrow \mathbb{R}$  is strictly increasing and convex.*
- (3) *For any  $[a, b] \subset (1, \infty)$ , there exists  $c_{a,b} > 0$  such that for any  $\alpha \in [a, b]$  and  $n_2 > n_1$  we have*

$$(f_{n_2} - f_{n_1})'(\alpha) \geq c_{a,b} \deg(f_{n_2}) \alpha^{\deg(f_{n_2})}.$$

- (4) *For any  $[a, b] \subset (1, \infty)$ , there exists  $C_{a,b} > 1$  such that for any  $\alpha \in [a, b]$  and  $n_2 > n_1$  we have*

$$\frac{\alpha^{\deg(f_{n_2})}}{C_{a,b}} \leq (f_{n_2} - f_{n_1})(\alpha) \leq C_{a,b} \alpha^{\deg(f_{n_2})}.$$

- (5) *For any  $[a, b] \subset (1, \infty)$  and  $C_{a,b}$  as in (4), for  $n_1$  sufficiently large the following inequality is satisfied for all  $n_2 > n_1$*

$$\left( \frac{2 \deg(f_{n_2})}{\deg(f_{n_1})} - 1 \right) \log C_{a,b} + (\deg(f_{n_1}) - \deg(f_{n_2})) \log a - \log \left( \deg(f_{n_2}) \left( \frac{\deg(f_{n_2})}{\deg(f_{n_1})} - 1 \right) \right) \leq -3 \log n_2.$$

*Then for Lebesgue almost every  $\alpha > 1$  the sequence  $(f_n(\alpha))_{n=1}^\infty$  has Poissonian pair correlations.*

The fifth assumption appearing in Theorem 1.3 might seem a little unwieldy. Essentially it is a condition on the growth rate of the sequence  $(\deg(f_n))_{n=1}^\infty$ . Note that it is not satisfied by the sequence  $(n)_{n=1}^\infty$ , which is why we cannot provide an affirmative answer to Question 1.2. However for many natural choices of sequences it is a straightforward exercise to check that this assumption is satisfied. As an example, whenever  $(\deg(f_n))_{n=1}^\infty = (n^k)_{n=1}^\infty$  for some  $k \geq 2$  then this assumption is satisfied. Similarly, if  $(\deg(f_n))_{n=1}^\infty = (n!)_{n=1}^\infty$  then the fifth assumption is satisfied. These observations imply the following theorem which follows from Theorem 1.3.

**Theorem 1.4.** *For Lebesgue almost every  $\alpha > 1$  the sequences  $(\alpha^{n^k})_{n=1}^\infty$  and  $(\alpha^{n^k} + \alpha^{n^{k-1}} + \dots + \alpha + 1)_{n=1}^\infty$  have Poissonian pair correlations for all  $k \geq 2$ . Similarly, for Lebesgue almost every  $\alpha > 1$  the sequence  $(\alpha^{n!})_{n=1}^\infty$  has Poissonian pair correlations.*

**Notation.** Throughout this paper we make use of the standard big  $\mathcal{O}$  notation, i.e.  $X = \mathcal{O}(Y)$  if there exists  $C > 0$  such that  $|X| \leq CY$ . When we want to emphasise a dependence for the underlying constant  $C$  we will include a subscript, i.e.  $X = \mathcal{O}_a(Y)$  if  $|X| \leq C \cdot Y$  for some  $C$  that depends upon  $a$ . Given a sequence of polynomials  $(f_n)_{n=1}^\infty$  we will use the notation  $(d_n)_{n=1}^\infty$  to denote its sequence of degrees. The sequence of polynomials we are referring to will be clear from the context. We let  $\mathcal{L}(\cdot)$  denote the Lebesgue measure.

## 2. PROOF OF THEOREM 1.3

We will repeatedly use the following lemma in our proof of Theorem 1.3.

**Lemma 2.1.** *Let  $f : [a, b] \rightarrow \mathbb{R}$  be a strictly increasing differentiable convex function. If  $I = [c, d]$  or  $I = [c, 1] \cup [0, d]$  for some  $c, d \in [0, 1]$ , then we have*

$$\frac{\mathcal{L}(I)(b-a)}{1 + \mathcal{L}(I)} + \mathcal{O}\left(\frac{\mathcal{L}(I)}{f'(a)}\right) \leq \mathcal{L}(\alpha \in [a, b] : \{f(\alpha)\} \in I) \leq \frac{\mathcal{L}(I)(b-a)}{1 - \mathcal{L}(I)} + \mathcal{O}\left(\frac{\mathcal{L}(I)}{f'(a)}\right).$$

Moreover, if the  $I$  above is such that  $\mathcal{L}(I) = 1/m$  for some  $m \in \mathbb{N}$ , then the above can be strengthened to

$$\mathcal{L}(\alpha \in [a, b] : \{f(\alpha)\} \in I) = \mathcal{L}(I)(b-a) + \mathcal{O}\left(\frac{\mathcal{L}(I)}{f'(a)}\right).$$

*Proof.* We will prove the statement for general  $I$  first. By adding a constant to  $f$  if necessary, we can assume without loss of generality that  $I = [0, c]$  for some  $c \in (0, 1]$ . We start by proving the lower bound. Consider the following collection of intervals:

$$[0, c], [c, 2c], \dots, \left[ \left( \left\lfloor \frac{1}{c} \right\rfloor - 1 \right) c, \left\lfloor \frac{1}{c} \right\rfloor c \right], \left[ \left\lfloor \frac{1}{c} \right\rfloor c, 1 \right].$$

These intervals cover  $[0, 1)$  and there are  $\lfloor \frac{1}{c} \rfloor + 1$  of them. Therefore, there exists an element of this collection, that we will denote by  $J$ , such that

$$(2.1) \quad \mathcal{L}(\alpha \in [a, b] : \{f(\alpha)\} \in J) \geq \frac{b-a}{\lfloor \frac{1}{c} \rfloor + 1} \geq \frac{c(b-a)}{1+c}.$$

Since  $f$  is strictly increasing and convex, the following inequality holds for any interval  $L \subset [f(a), f(b)]$  and  $t \geq 0$ :

$$(2.2) \quad \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in L) \geq \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in L + t).$$

Using that  $f$  is strictly increasing and convex, together with the mean value theorem, we have the following bound. For any interval  $L \subset \mathbb{R}$ , we have

$$(2.3) \quad \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in L) = \mathcal{O}\left(\frac{\mathcal{L}(L)}{f'(a)}\right).$$

Choosing  $t \in [0, 1)$  such that  $J \subseteq [0, c] + t$  we obtain:

$$\begin{aligned} \mathcal{L}(\alpha \in [a, b] : \{f(\alpha)\} \in [0, c]) &= \sum_{M=\lfloor f(a) \rfloor}^{\lfloor f(b) \rfloor} \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in [0, c] + M) \\ &\stackrel{(2.3)}{=} \sum_{M=\lfloor f(a) \rfloor + 1}^{\lfloor f(b) \rfloor - 1} \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in [0, c] + M) + \mathcal{O}\left(\frac{c}{f'(a)}\right) \\ &\stackrel{(2.2)}{\geq} \sum_{M=\lfloor f(a) \rfloor + 1}^{\lfloor f(b) \rfloor - 1} \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in [0, c] + t + M) + \mathcal{O}\left(\frac{c}{f'(a)}\right) \\ &\geq \sum_{M=\lfloor f(a) \rfloor + 1}^{\lfloor f(b) \rfloor - 1} \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in J + M) + \mathcal{O}\left(\frac{c}{f'(a)}\right) \\ &\stackrel{(2.3)}{=} \sum_{M=\lfloor f(a) \rfloor}^{\lfloor f(b) \rfloor} \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in J + M) + \mathcal{O}\left(\frac{c}{f'(a)}\right) \\ &= \mathcal{L}(\alpha \in [a, b] : \{f(\alpha)\} \in J) + \mathcal{O}\left(\frac{c}{f'(a)}\right) \\ &\stackrel{(2.1)}{\geq} \frac{c(b-a)}{1+c} + \mathcal{O}\left(\frac{c}{f'(a)}\right). \end{aligned}$$

This completes the proof of our lower bound. The proof of the upper bound is similar. Note that the upper bound is trivial for  $[0, c]$  such that  $c \geq 1/2$ . As such we restrict our attention to intervals of the form  $[0, c]$  for  $c < 1/2$ . This time we consider the collection of

intervals

$$[0, c], [c, 2c], \dots, \left[ \left( \left\lfloor \frac{1}{c} \right\rfloor - 1 \right) c, 1 \right].$$

These intervals cover  $[0, 1)$  and there are  $\lfloor \frac{1}{c} \rfloor$  of them. Since the Lebesgue measure of the set of  $\alpha$  that are mapped into the intersection of two of these intervals is zero, there exists an element of this collection, that we will denote by  $J'$ , such that

$$\mathcal{L}(\alpha \in [a, b] : \{f(\alpha)\} \in J') \leq \frac{b-a}{\lfloor \frac{1}{c} \rfloor} \leq \frac{c(b-a)}{1-c}.$$

Applying (2.2) in conjunction with (2.3), an analogous argument to that given above yields

$$\mathcal{L}(\alpha \in [a, b] : \{f(\alpha)\} \in [0, c]) \leq \mathcal{L}(\alpha \in [a, b] : f(\alpha) \in J') + \mathcal{O}\left(\frac{c}{f'(a)}\right).$$

Then by the definition of  $J'$  we have

$$\mathcal{L}(\alpha \in [a, b] : \{f(\alpha)\} \in [0, c]) \leq \frac{c(b-a)}{1-c} + \mathcal{O}\left(\frac{c}{f'(a)}\right).$$

This completes our proof of the upper bound.

To deduce the stronger statement when  $\mathcal{L}(I) = 1/m$  for some  $m \in \mathbb{N}$ , notice that the two collections of intervals appearing in the proof of the lower bound and upper bound can both be replaced by the single collection given by the intervals

$$[0, 1/m], [1/m, 2/m], \dots, [(m-1)/m, 1].$$

Importantly this collection consists of exactly  $m$  elements. Repeating the arguments given above for this collection yields the stronger statement. □

The following proposition is the tool that allows us to prove Theorem 1.3.

**Proposition 2.2.** *Let  $(f_n)_{n=1}^\infty$  be a sequence of polynomials satisfying the hypothesis of Theorem 1.3. Then for any  $[a, b] \subset (1, \infty)$  and  $s > 0$  we have*

$$\int_a^b \left( \frac{\#\{1 \leq m \neq n \leq N : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N}\}}{N} - 2s \right)^2 d\alpha = \mathcal{O}_{s,a,b}\left(\frac{1}{N}\right).$$

We split our proof of Proposition 2.2 into a series of lemmas. Throughout this paper  $\chi_A$  denotes the indicator function on a set  $A \subset \mathbb{R}$ . We start by expanding the bracket

appearing within the integral to obtain:

$$\begin{aligned}
(2.4) \quad & \int_a^b \left( \frac{\#\{1 \leq m \neq n \leq N : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N}\}}{N} - 2s \right)^2 d\alpha \\
&= \frac{1}{N^2} \sum_{\substack{1 \leq m \neq n \leq N \\ 1 \leq p \neq q \leq N \\ (m,n) \neq (p,q)}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \\
&+ \frac{1}{N^2} \sum_{1 \leq m \neq n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha \\
&- \frac{4s}{N} \sum_{1 \leq m \neq n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha \\
&+ 4s^2(b-a).
\end{aligned}$$

We will focus on each term on the right hand side of (2.4) individually. It is useful at this point to rewrite the first term as follows:

$$\begin{aligned}
& \frac{1}{N^2} \sum_{\substack{1 \leq m \neq n \leq N \\ 1 \leq p \neq q \leq N \\ (m,n) \neq (p,q)}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \\
&= \frac{4}{N^2} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ (m,n) \neq (p,q)}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \\
&= \frac{4}{N^2} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q \neq n}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \\
&+ \frac{4}{N^2} \sum_{1 \leq m \neq p < n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_p(\alpha)\|) d\alpha \\
(2.5) \quad &= \frac{8}{N^2} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \\
&+ \frac{8}{N^2} \sum_{1 \leq m < p < n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_p(\alpha)\|) d\alpha.
\end{aligned}$$

The behaviour of the two terms on the right hand side of (2.5) is described by the following lemmas.



**Lemma 2.3.** *Suppose  $(f_n)_{n=1}^\infty$  is a sequence of polynomials satisfying the hypothesis of Theorem 1.3. Then for any  $[a, b] \subset (1, \infty)$  and  $s > 0$  we have*

$$\frac{8}{N^2} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \leq 4s^2(b-a) + \mathcal{O}_{s,a,b}\left(\frac{1}{N}\right).$$

*Proof.* To each  $1 \leq p < q \leq N$  and  $M \in [\lfloor f_q(a) - f_p(a) \rfloor, \lceil f_q(b) - f_p(b) \rceil]$  we associate the interval

$$I_{M,q,p} := \left\{ \alpha \in [a, b] : f_q(\alpha) - f_p(\alpha) \in \left[ M - \frac{s}{N}, M + \frac{s}{N} \right] \right\}.$$

This is an interval because the function  $f_q - f_p$  is strictly increasing. We note that

$$(2.6) \quad \left\{ \alpha \in [a, b] : \|f_q(\alpha) - f_p(\alpha)\| \leq \frac{s}{N} \right\} = \bigcup_{M=\lfloor f_q(a)-f_p(a) \rfloor}^{\lceil f_q(b)-f_p(b) \rceil} I_{M,q,p}.$$

We denote the left hand point of each non-empty  $I_{M,q,p}$  by  $c_{M,q,p}$ .

Note that  $\|f_n(\alpha) - f_m(\alpha)\| \in [0, \frac{s}{N}]$  if and only if  $\{f_n(\alpha) - f_m(\alpha)\} \in [0, \frac{s}{N}] \cup [1 - \frac{s}{N}, 1)$ . Therefore by an application of Lemma 2.1 we have

$$(2.7) \quad \begin{aligned} & \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \\ &= \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a)-f_p(a) \rfloor}^{\lceil f_q(b)-f_p(b) \rceil} \int_{I_{M,q,p}} \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha \\ &\leq \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a)-f_p(a) \rfloor}^{\lceil f_q(b)-f_p(b) \rceil} \frac{\mathcal{L}(I_{M,q,p}) 2s/N}{1 - 2s/N} \\ &\quad + \mathcal{O} \left( \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a)-f_p(a) \rfloor}^{\lceil f_q(b)-f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} \right). \end{aligned}$$

We now treat the two terms appearing in (2.7) separately.

**Bounding the first term in (2.7).**

By an application of Lemma 2.1 and (2.6) we have

$$\begin{aligned}
(2.8) \quad & \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{\mathcal{L}(I_{M,q,p}) 2s/N}{1 - 2s/N} \\
&= \frac{2s}{N - 2s} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \mathcal{L}(I_{M,q,p}) \\
&\leq \frac{2s}{N - 2s} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \left( \frac{(b-a)2s/N}{1 - 2s/N} + \mathcal{O}\left(\frac{2s}{N(f_q - f_p)'(a)}\right) \right) \\
&= \frac{4s^2(b-a)}{(N-2s)^2} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} 1 \\
&+ \mathcal{O}\left( \frac{4s^2}{N(N-2s)} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \frac{1}{(f_q - f_p)'(a)} \right).
\end{aligned}$$

By our third assumption we know that  $(f_q - f_p)'(a) \geq c_{a,b} d_q a^{d_q}$  for  $q > p$ . We also know by our first assumption that  $(d_n)_{n=1}^\infty$  is a strictly increasing sequence of natural numbers, therefore  $d_n \geq n$  for all  $n \in \mathbb{N}$ . This implies  $(f_q - f_p)'(a) \geq c_{a,b} q a^q$  for  $q > p$ . Therefore

$$\begin{aligned}
(2.9) \quad & \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \frac{1}{(f_q - f_p)'(a)} = \sum_{n=3}^N \sum_{m=1}^{n-1} \sum_{p=1}^{n-2} \sum_{q=p+1}^{n-1} \frac{1}{(f_q - f_p)'(a)} \\
&= \mathcal{O}_{a,b} \left( \sum_{n=3}^N \sum_{m=1}^{n-1} \sum_{p=1}^{n-2} \sum_{q=p+1}^{n-1} \frac{1}{q a^q} \right) \\
&= \mathcal{O}_{a,b} \left( \sum_{n=3}^N \sum_{m=1}^{n-1} \sum_{p=1}^{n-2} \frac{1}{(p+1) a^{p+1}} \right) \\
&= \mathcal{O}_{a,b} \left( \sum_{n=3}^N \sum_{m=1}^{n-1} 1 \right) \\
&= \mathcal{O}_{a,b}(N^2).
\end{aligned}$$

A straightforward calculation yields

$$(2.10) \quad \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} 1 = \frac{N^4}{8} + \mathcal{O}(N^3).$$

Substituting (2.9) and (2.10) into (2.8), we see that the following holds for the first term in (2.7)

$$\sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{\mathcal{L}(I_{M,q,p}) 2s/N}{1 - 2s/N} \leq \frac{s^2(b-a)N^4}{2(N-2s)^2} + \mathcal{O}_{s,a,b}(N).$$

It is easy to show that

$$\frac{s^2(b-a)N^4}{2(N-2s)^2} = \frac{s^2(b-a)N^2}{2} + \mathcal{O}_{s,a,b}(N).$$

Therefore the following holds for the first term in (2.7)

$$(2.11) \quad \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{\mathcal{L}(I_{M,q,p}) 2s/N}{1 - 2s/N} \leq \frac{s^2(b-a)N^2}{2} + \mathcal{O}_{s,a,b}(N).$$

### Bounding the second term in (2.7).

By the fifth assumption listed in Theorem 1.3, we know that there exists some  $N_1 \in \mathbb{N}$  for which

$$(2.12) \quad \left( \frac{2d_n}{d_q} - 1 \right) \log C_{a,b} + (d_q - d_n) \log a - \log \left( d_n \left( \frac{d_n}{d_q} - 1 \right) \right) \leq -3 \log n$$

whenever  $q \geq N_1$  and  $n > q$ . The equation below describes the error that occurs by restricting the second term in (2.7) to  $q \geq N_1$ . As we will see, this error will be negligible.

We have

$$\begin{aligned}
\sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} &= \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} \\
&+ \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q < N_1 \\ q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} \\
&= \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} \\
&+ \mathcal{O}_{s,a,b} \left( \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q < N_1 \\ q < n}} \frac{1}{N} \right) \\
(2.13) \quad &= \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} \\
&+ \mathcal{O}_{s,a,b}(N).
\end{aligned}$$

In the penultimate equality we used that for any  $q \leq N_1$ , for  $p < q$  and  $m, n$  satisfying  $m < n$  and  $q < n$ , we have

$$\sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{1}{(f_n - f_m)'(c_{M,q,p})} = \mathcal{O}_{a,b}(1).$$

We now bound the first term on the right hand side of (2.13). Recall that by our fourth assumption there exists  $C_{a,b} > 1$  such that  $f_q(\alpha) - f_p(\alpha) \leq C_{a,b} \alpha^{d_q}$  for all  $\alpha \in [a, b]$ . Therefore

$$C_{a,b} c_{M,q,p}^{d_q} \geq M - \frac{s}{N}.$$

Increasing  $C_{a,b}$  if necessary, we may assume without loss of generality that

$$C_{a,b} c_{M,q,p}^{d_q} \geq M + 2$$

holds for all  $c_{M,q,p}$ . Therefore

$$(2.14) \quad c_{M,q,p} \geq \left( \frac{M+2}{C_{a,b}} \right)^{1/d_q}.$$

Using the fact that  $f_n - f_m$  is convex, we see that (2.14) implies

$$(f_n - f_m)'(c_{M,q,p}) \geq (f_n - f_m)' \left( \left( \frac{M+2}{C_{a,b}} \right)^{1/d_q} \right).$$

Therefore

$$(2.15) \quad \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} \\ \leq \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)' \left( \left( \frac{M+2}{C_{a,b}} \right)^{1/d_q} \right)}.$$

We would now like to be able to use our third assumption to assert that

$$(f_n - f_m)' \left( \left( \frac{M+2}{C_{a,b}} \right)^{1/d_q} \right) \geq c_{a,b} d_n \left( \frac{M+2}{C_{a,b}} \right)^{d_n/d_q}.$$

However we cannot apply this assumption directly since  $\left( \frac{M+2}{C_{a,b}} \right)^{1/d_q}$  is not necessarily contained in  $[a, b]$ . However, we know by our fourth assumption that  $f_q(a) - f_p(a) \geq \frac{a^{d_q}}{C_{a,b}}$  and  $f_q(b) - f_p(b) \leq C_{a,b} b^{d_q}$  for all  $p < q$ . Using these facts, together with the property  $\lim_{q \rightarrow \infty} d_q = \infty$ , it follows that for  $q$  sufficiently large, for  $p < q$  and  $M \in [\lfloor f_q(a) - f_p(a) \rfloor, \lceil f_q(b) - f_p(b) \rceil]$ , we have

$$(2.16) \quad \left( \frac{M+2}{C_{a,b}} \right)^{1/d_q} \in \left[ \frac{1+a}{2}, 2b \right].$$

Without loss of generality we can assume that the  $N_1$  we chose earlier was sufficiently large to guarantee (2.16) holds for any  $M \in [\lfloor f_q(a) - f_p(a) \rfloor, \lceil f_q(b) - f_p(b) \rceil]$  for  $q \geq N_1$  and  $p < q$ . In which case we can apply our third assumption for the interval  $[\frac{1+a}{2}, 2b]$  to assert that there exists a constant  $c'_{a,b} > 0$  such that

$$(f_n - f_m)' \left( \left( \frac{M+2}{C_{a,b}} \right)^{1/d_q} \right) \geq c'_{a,b} d_n \left( \frac{M+2}{C_{a,b}} \right)^{d_n/d_q}$$

for any  $M \in [\lfloor f_q(a) - f_p(a) \rfloor, \lceil f_q(b) - f_p(b) \rceil]$  for  $q \geq N_1$  and  $p < q$ . Using this bound in (2.15) we have

$$\begin{aligned}
& \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} \\
&= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{C_{a,b}^{d_n/d_q}}{d_n(M+2)^{d_n/d_q}} \right) \\
&= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \frac{C_{a,b}^{d_n/d_q}}{d_n} \int_{\lfloor f_q(a) - f_p(a) \rfloor - 1}^{\lceil f_q(b) - f_p(b) \rceil - 1} \frac{1}{(x+2)^{d_n/d_q}} dx \right) \\
&= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \frac{(C_{a,b}^{d_n/d_q} (\lfloor f_q(a) - f_p(a) \rfloor + 1)^{1-d_n/d_q}}{d_n(d_n/d_q - 1)} \right) \\
&= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \frac{(C_{a,b}^{d_n/d_q} (f_q(a) - f_p(a))^{1-d_n/d_q}}{d_n(d_n/d_q - 1)} \right) \\
(2.17) \quad &= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \frac{C_{a,b}^{d_n/d_q} (a^{d_q}/C_{a,b})^{1-d_n/d_q}}{d_n(d_n/d_q - 1)} \right).
\end{aligned}$$

In the final line we used our fourth assumption that  $f_q(a) - f_p(a) \geq \frac{a^{d_q}}{C_{a,b}}$ . By (2.12) we know that

$$\frac{C_{a,b}^{d_n/d_q} (a^{d_q}/C_{a,b})^{1-d_n/d_q}}{d_n(d_n/d_q - 1)} \leq \frac{1}{n^3}$$

whenever  $q \geq N_1$  and  $n > q$ . Substituting this bound into (2.17) we obtain

$$\begin{aligned}
\sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} &= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \frac{1}{n^3} \right) \\
&= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{q=N_1}^{N-1} \sum_{p=1}^{q-1} \sum_{n=q+1}^N \sum_{m=1}^{n-1} \frac{1}{n^3} \right) \\
&= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{q=N_1}^{N-1} \sum_{p=1}^{q-1} \sum_{n=q+1}^N \frac{1}{n^2} \right) \\
&= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{q=N_1}^{N-1} \sum_{p=1}^{q-1} \frac{1}{q} \right) \\
&= \mathcal{O}_{s,a,b} \left( \frac{1}{N} \sum_{q=N_1}^{N-1} 1 \right). \\
&= \mathcal{O}_{s,a,b}(1)
\end{aligned}$$

Therefore

$$\sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ N_1 \leq q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} = \mathcal{O}_{s,a,b}(1).$$

Which when combined with (2.13) gives

$$(2.18) \quad \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \sum_{M=\lfloor f_q(a) - f_p(a) \rfloor}^{\lceil f_q(b) - f_p(b) \rceil} \frac{2s}{N(f_n - f_m)'(c_{M,q,p})} = \mathcal{O}_{s,a,b}(N).$$

Substituting (2.11) and (2.18) into (2.7) we obtain the desired inequality:

$$\frac{8}{N^2} \sum_{\substack{1 \leq m < n \leq N \\ 1 \leq p < q \leq N \\ q < n}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \leq 4s^2(b-a) + \mathcal{O}_{s,a,b} \left( \frac{1}{N} \right).$$

□

**Lemma 2.4.** *Suppose  $(f_n)_{n=1}^\infty$  is a sequence of polynomials satisfying the hypothesis of Theorem 1.3. Then for any  $[a, b] \subset (1, \infty)$  and  $s > 0$  we have*

$$\frac{8}{N^2} \sum_{1 \leq m < p < n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_p(\alpha)\|) d\alpha = \mathcal{O}_{s,a,b} \left( \frac{1}{N} \right)$$

*Proof.* Notice that if  $\chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) = 1$  and  $\chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_p(\alpha)\|) = 1$  then  $\chi_{[0, \frac{2s}{N}]}(\|f_p(\alpha) - f_m(\alpha)\|) = 1$ . Therefore

$$(2.19) \quad \begin{aligned} & \sum_{1 \leq m < p < n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_p(\alpha)\|) d\alpha \\ & \leq \sum_{1 \leq m < p < n \leq N} \int_a^b \chi_{[0, \frac{2s}{N}]}(\|f_p(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_p(\alpha)\|) d\alpha. \end{aligned}$$

Importantly  $f_p - f_m$  and  $f_n - f_p$  are polynomials of different degrees. As a consequence of this, the right hand side of (2.19) is in a form where the arguments used in the proof of Lemma 2.3 can be applied. In particular one can define appropriate analogues of the intervals  $I_{M,q,p}$  and the points  $c_{M,q,p}$ . Then by analogous arguments to those given in the proof of Lemma 2.3, it can be shown that

$$(2.20) \quad \sum_{1 \leq m < p < n \leq N} \int_a^b \chi_{[0, \frac{2s}{N}]}(\|f_p(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_p(\alpha)\|) d\alpha = \mathcal{O}_{s,a,b}(N).$$

Substituting (2.20) into (2.19) we obtain

$$\frac{8}{N^2} \sum_{1 \leq m < p < n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_p(\alpha)\|) d\alpha = \mathcal{O}_{s,a,b} \left( \frac{1}{N} \right).$$

□

Substituting the bounds provided by Lemma 2.3 and Lemma 2.4 into (2.5), we see that under the hypothesis of Theorem 1.3, the following holds for the first term in (2.4)

$$(2.21) \quad \frac{1}{N^2} \sum_{\substack{1 \leq m \neq n \leq N \\ 1 \leq p \neq q \leq N \\ (m,n) \neq (p,q)}} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \chi_{[0, \frac{s}{N}]}(\|f_q(\alpha) - f_p(\alpha)\|) d\alpha \leq 4s^2(b-a) + \mathcal{O}_{s,a,b} \left( \frac{1}{N} \right).$$

**Lemma 2.5.** *Suppose  $(f_n)_{n=1}^\infty$  is a sequence of polynomials satisfying the hypothesis of Theorem 1.3. Then for any  $[a, b] \subset (1, \infty)$  and  $s > 0$  we have*

$$\frac{1}{N^2} \sum_{1 \leq m \neq n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha = \mathcal{O}_{s,a,b} \left( \frac{1}{N} \right)$$



and

$$\frac{4s}{N} \sum_{1 \leq m \neq n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha \geq 8s^2(b-a) + \mathcal{O}_{s,a,b}\left(\frac{1}{N}\right).$$

*Proof.* To prove our result it suffices to show that

$$(2.22) \quad \sum_{1 \leq m \neq n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha = \mathcal{O}_{s,a,b}(N)$$

and

$$(2.23) \quad \sum_{1 \leq m \neq n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha \geq 2s(b-a)N + \mathcal{O}_{s,a,b}(1).$$

We start by proving (2.22). Applying Lemma 2.1 together with our first and third assumptions, we see that the following holds:

$$\begin{aligned} \sum_{1 \leq m \neq n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha &= 2 \sum_{1 \leq m < n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) d\alpha \\ &\leq 2 \sum_{1 \leq m < n \leq N} \left( \frac{(b-a)2s/N}{1-2s/N} + \mathcal{O}_{s,a,b}\left(\frac{1}{Nd_n a^{d_n}}\right) \right) \\ &= 2 \sum_{1 \leq m < n \leq N} \frac{(b-a)2s/N}{1-2s/N} + \mathcal{O}_{s,a,b}\left(\sum_{1 \leq m < n \leq N} \frac{1}{Nd_n a^{d_n}}\right) \\ &\leq \frac{4s(b-a)}{N-2s} \sum_{1 \leq m < n \leq N} 1 + \mathcal{O}_{s,a,b}\left(\frac{1}{N}\right) \\ &= \frac{4s(b-a)}{N-2s} \left( \frac{N^2}{2} + \mathcal{O}(N) \right) + \mathcal{O}_{s,a,b}\left(\frac{1}{N}\right) \\ &= \mathcal{O}_{s,a,b}(N). \end{aligned}$$

By a similar argument, this time using the lower bound from Lemma 2.1, it can be shown that

$$\sum_{1 \leq m \neq n \leq N} \int_a^b \chi_{[0, \frac{s}{N}]}(\|f_n(\alpha) - f_m(\alpha)\|) \geq 2s(b-a)N + \mathcal{O}_{s,a,b}(1).$$

□

*Proof of Proposition 2.2.* Proposition 2.2 follows by substituting the bounds provided by (2.21) and Lemma 2.5 into (2.4). □

Equipped with Proposition 2.2 we are now in a position to prove Theorem 1.3.

*Proof of Theorem 1.3.* Let us start by fixing  $[a, b] \subset (1, \infty)$  and let  $s > 0$  be arbitrary. By Proposition 2.2 we know that

$$\int_a^b \left( \frac{\#\{1 \leq m \neq n \leq N^2 : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N^2}\}}{N^2} - 2s \right)^2 d\alpha = \mathcal{O}_{s,a,b} \left( \frac{1}{N^2} \right).$$

Applying Markov's inequality, we have

$$\mathcal{L} \left( \alpha \in [a, b] : \left( \frac{\#\{1 \leq m \neq n \leq N^2 : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N^2}\}}{N^2} - 2s \right)^2 > N^{-1/2} \right) = \mathcal{O}_{s,a,b} \left( \frac{1}{N^{3/2}} \right).$$

Importantly

$$\sum_{N=1}^{\infty} \frac{1}{N^{3/2}} < \infty.$$

Therefore, by the Borel-Cantelli lemma, for Lebesgue almost every  $\alpha \in [a, b]$ , the inequality

$$\left( \frac{\#\{1 \leq m \neq n \leq N^2 : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N^2}\}}{N^2} - 2s \right)^2 > N^{-1/2}$$

is satisfied for at most finitely many values of  $N$ . This implies that for Lebesgue almost every  $\alpha \in [a, b]$  we have

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq m \neq n \leq N^2 : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N^2}\}}{N^2} = 2s.$$

The parameter  $s$  was arbitrary. Therefore by considering a countable dense set of  $s$ , and applying an approximation argument, it can be shown that for Lebesgue almost every  $\alpha \in [a, b]$ , for any  $s > 0$  we have

$$(2.24) \quad \lim_{N \rightarrow \infty} \frac{\#\{1 \leq m \neq n \leq N^2 : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N^2}\}}{N^2} = 2s.$$

To each  $N \in \mathbb{N}$  we define  $M_N \in \mathbb{N}$  to be the unique integer satisfying the inequalities

$$M_N^2 \leq N < (M_N + 1)^2.$$

Let  $s > 0$  and  $\epsilon > 0$  be arbitrary. Since (2.24) holds for Lebesgue almost every  $\alpha \in [a, b]$  for any  $s > 0$ , we have

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \frac{\#\{1 \leq m \neq n \leq N : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N}\}}{N} \\ & \leq \limsup_{N \rightarrow \infty} \frac{\#\{1 \leq m \neq n \leq (M_N + 1)^2 : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s+\epsilon}{(M_N+1)^2}\}}{M_N^2} \\ & = \limsup_{N \rightarrow \infty} \left( \frac{(M_N + 1)^2}{M_N^2} \right) \frac{\#\{1 \leq m \neq n \leq (M_N + 1)^2 : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s+\epsilon}{(M_N+1)^2}\}}{(M_N + 1)^2} \\ & = 2(s + \epsilon) \end{aligned}$$

for Lebesgue almost every  $\alpha \in [a, b]$ . Similarly it can be shown that for Lebesgue almost every  $\alpha \in [a, b]$ , we have

$$\liminf_{N \rightarrow \infty} \frac{\#\{1 \leq m \neq n \leq N : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N}\}}{N} \geq 2(s - \epsilon).$$

Since  $s$  and  $\epsilon$  were arbitrary, we may conclude that for Lebesgue almost every  $\alpha \in [a, b]$  we have

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq m \neq n \leq N : \|f_n(\alpha) - f_m(\alpha)\| \leq \frac{s}{N}\}}{N} = 2s$$

for any  $s > 0$ . Since the interval  $[a, b]$  was arbitrary this completes our proof.  $\square$

**Acknowledgements.** The author was supported by EPSRC grant EP/M001903/1.

## REFERENCES

- [1] I. Aichinger, C. Aistleitner and G. Larcher, *On Quasi-Energy-Spectra, Pair Correlations of Sequences and Additive Combinatorics*, Contemporary Computational Mathematics – a celebration of the 80th birthday of Ian Sloan (J. Dick, F. Y. Kuo, H. Wozniakowski, eds.), Springer-Verlag, 1–16, 2018.
- [2] C. Aistleitner, *Quantitative uniform distribution results for geometric progressions*, Israel J. Math. 204 (2014), no. 1, 155–197.
- [3] C. Aistleitner, S. Baker, *On the pair correlations of powers of real numbers*, arXiv:1910.01437 [math.NT].
- [4] C. Aistleitner, T. Lachmann, F. Pausinger, *Pair correlations and equidistribution*, J. Number Th. 182 (2018), 206–220.
- [5] C. Aistleitner, G. Larcher, M. Lewko, *Additive energy and the Hausdorff dimension of the exceptional set in metric pair correlation problems*, with an appendix by Jean Bourgain, Israel J. Math. 222 (2017), no. 1, 463–485.
- [6] S. Baker, *On the distribution of powers of real numbers modulo 1*. *Unif. Distrib. Theory* 10 (2015), no. 2, 67–75.

- [7] Y. Bugeaud, *Approximation by algebraic numbers*, Cambridge Tracts in Mathematics, 160. Cambridge University Press, Cambridge, 2004. xvi+274 pp.
- [8] Y. Bugeaud, *Distribution modulo one and Diophantine approximation*, Cambridge Tracts in Mathematics, 193. Cambridge University Press, Cambridge, 2012.
- [9] Y. Bugeaud, L. Liao, M. Rams, *Metrical results on the distribution of fractional parts of powers of real numbers*. Proc. Edinburgh Math. Soc. 62 (2019), 505–521
- [10] Y. Bugeaud, V. Moshchevitin, *On fractional parts of powers of real numbers close to 1*, Math. Z. 271 (2012), no. 3–4, 627–637.
- [11] A. Dubickas, *On the powers of some transcendental numbers*, Bull. Austral. Math. Soc. **76** (2007), no. 3, 433–440.
- [12] G. H. Hardy, *A problem of Diophantine approximation*, J. Indian Math. Soc. **11** (1919), 162–166.
- [13] J.-P. Kahane, *Sur la répartition des puissances modulo 1*, C. R. Math. Acad. Sci. Paris **352** (2014), no. 5, 383–385. (1924), 115–125.
- [14] J. F. Koksma, *Ein mengentheoretischer Satz über die Gleichverteilung modulo Eins*, Compositio Math. 2 (1935), 250–258.
- [15] L. Kuipers, H. Niederreiter, *Uniform distribution of sequences*, Pure and Applied Mathematics. Wiley-Interscience, John Wiley & Sons, New York-London-Sydney, 1974. xiv+390 pp.
- [16] G. Larcher, W. Stockinger, *On Pair Correlation of Sequences*, arXiv:1903.09978 [math.NT].
- [17] C. Pisot, *La répartition modulo 1 et les nombres algébriques*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (2) **7** (1938), no. 3–4, 205–248.
- [18] C. Pisot, *Sur la répartition modulo 1 des puissances successives d'un même nombre*, C.R. Acad. Sci. Paris **204** (1937), 312–314.
- [19] Z. Rudnick, P. Sarnak, *The pair correlation function of fractional parts of polynomials*, Comm. Math. Phys. 194 (1998), no. 1, 61–70.
- [20] Z. Rudnick, P. Sarnak, A. Zaharescu, *The distribution of spacings between the fractional parts of  $n^2\alpha$* , Invent. Math. 145 (2001), no. 1, 37–57.
- [21] Z. Rudnick, A. Zaharescu, *A metric result on the pair correlation of fractional parts of sequences*, Acta Arith. 89 (1999), no. 3, 283–293.

SIMON BAKER: SCHOOL OF MATHEMATICS, UNIVERSITY OF BIRMINGHAM, BIRMINGHAM, B15 2TT, UK

*Email address:* `simonbaker412@gmail.com`