

Remarks about inhomogeneous pair correlations

BY FELIPE A. RAMÍREZ

Department of Mathematics and Computer Science, Wesleyan University,

Middletown, Connecticut, U.S.A

e-mail: framirez@wesleyan.edu

(Received 25 January 2021; revised 11 August 2021; accepted 09 August 2021)

For Jorge A. Ramírez (1954–2020)—with Love and Gratitude

Abstract

Given an infinite subset $\mathcal{A} \subseteq \mathbb{N}$, let A denote its smallest N elements. There is a rich and growing literature on the question of whether for typical $\alpha \in [0, 1]$, the pair correlations of the set $\alpha A \pmod{1} \subset [0, 1]$ are asymptotically Poissonian as N increases. We define an inhomogeneous generalisation of the concept of pair correlation, and we consider the corresponding doubly metric question. Many of the results from the usual setting carry over to this new setting. Moreover, the double metricity allows us to establish some new results whose singly metric analogues are missing from the literature.

2020 Mathematics Subject Classification: 11J71, 11K38, 11J83 (Primary);
11K60, 11K31 (Secondary)

1. Metric Poissonian pair correlations

Given a sequence $\mathbf{x} = (x_n)_{n=1}^\infty$ of points on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, a point $\gamma \in \mathbb{T}$, and a real number $s > 0$, we are interested in the asymptotic frequency with which $x_i - x_j$ ($i, j \leq N$) lies in the arc $[\gamma - s/N, \gamma + s/N] \subset \mathbb{T}$. That is, we study the limiting behaviour of

$$F(\gamma, s, N, \mathbf{x}) = \frac{1}{N} \# \left\{ (i, j) \in [N] \times [N] \mid i \neq j, \|x_i - x_j - \gamma\| \leq \frac{s}{N} \right\},$$

where $\|\cdot\|$ denotes distance to $0 \in \mathbb{T}$ and $[N] := \{1, 2, \dots, N\}$. For an increasing sequence of natural numbers $\mathcal{A} = (a_n)_{n=1}^\infty$ and $\alpha, \gamma \in \mathbb{T}$, we denote $F(\alpha, \gamma, s, N, \mathcal{A}) = F(\gamma, s, N, \mathbf{x})$, where the sequence \mathbf{x} is defined by $x_n = a_n \alpha \pmod{1}$.

Much attention is paid to the behaviour of F when $\gamma = 0$. In this case $F(0, s, N, \mathbf{x})$ is called the *pair correlation statistic* of \mathbf{x} . One says that \mathbf{x} has *Poissonian pair correlations (PPC)* if

$$(\forall s > 0) \quad \lim_{N \rightarrow \infty} F(0, s, N, \mathbf{x}) = 2s.$$

Like equidistribution, PPC is a marker of randomness; a sequence of points on the circle which have been chosen independently and uniformly at random will almost surely have Poissonian pair correlations, just as they will almost surely be equidistributed in the circle. In fact, PPC is a stronger feature of randomness than equidistribution is, in the sense that

any sequence which has PPC must also be equidistributed [1, 6, 7, 13]. The converse fails. For example, an orbit of any irrational circle rotation equidistributes, but does not have Poissonian pair correlations. Indeed, the three gaps theorem implies that the gap distribution of the points of such an orbit is far from random.

In the past two decades, one of the questions of greatest interest in this area has been whether a given $\mathcal{A} \subset \mathbb{N}$ has *metric Poissonian pair correlations (MPPC)*, that is, if for Lebesgue almost every $\alpha \in \mathbb{T}$, the sequence $(a_n\alpha \pmod{1})_{n=1}^\infty$ has Poissonian pair correlations. Inspired by a problem in quantum mechanics, Rudnick and Sarnak [14] showed that $(n^k)_{n \geq 1}$ has MPPC whenever $k \geq 2$. But $(n)_{n \geq 1}$ (that is, $\mathcal{A} = \mathbb{N}$) does not have MPPC because for every α the corresponding sequence on \mathbb{T} is an orbit of the circle rotation over angle $2\pi\alpha$, and, as we have mentioned, orbits of circle rotations do not have PPC. To put it informally, the problem for $\mathcal{A} = \mathbb{N}$ arises from the fact that initial strings from the sequence \mathbb{N} have too much additive structure; there are too many different ways to achieve any given $d \in [N] - [N]$ as a difference of two elements of $[N]$, and as a result, $F(\alpha, 0, s, N, \mathbb{N})$ counts events that have been rigged by this extra structure of $[N]$ to occur with non-random regularity. In [3], Aistleitner, Larcher, and Lewko made an important forward stride in the study of MPPC by putting this informal reasoning on a rigorous footing. They connected the pair correlations of \mathcal{A} to the asymptotic behaviour of the additive energy

$$E(A) := \# \left\{ (a, b, c, d) \in A^4 \mid a + b = c + d \right\},$$

where $A := A_N$ denotes the smallest N elements of \mathcal{A} . Specifically, they proved the following theorem.

THEOREM 1.1 ([3, theorem 1]). *For an infinite subset $\mathcal{A} \subset \mathbb{N}$, let A_N denote its smallest N elements. If there exists some $\delta > 0$ such that*

$$E(A_N) \leq N^{3-\delta}$$

for all sufficiently large N , then \mathcal{A} has MPPC.

Remark. The Rudnick–Sarnak result follows from Theorem 1.1. On the other hand, it is easy to see that Theorem 1.1 does not apply to $\mathcal{A} = \mathbb{N}$, since in this case one has $E(A_N) \gg N^3$.

Remark. (on notation) For functions $f, g: \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$ we use $f \ll g$ to mean $f = O(g)$ and $f \gg g$ to mean $g = O(f)$. If both hold, we write $f \asymp g$. By $f \sim g$ we mean that $(f/g) \rightarrow 1$ as the argument increases to ∞ .

Aistleitner, Larcher and Lewko asked whether $E(A_N) = o(N^3)$ is necessary and sufficient for \mathcal{A} to have MPPC. In an appendix to their paper, Bourgain answered the question: it is necessary but not sufficient. This left the question of whether there is some additive energy threshold separating the sets that have MPPC from those that do not. In [4], Bloom, Chow, Gafni and Walker formulated the following question proposing a location for such a threshold:

Question 1.2 ([4, fundamental question 1.7]). *Let $\mathcal{A} \subset \mathbb{N}$ be an infinite set and suppose that*

$$E(A_N) \sim N^3 \psi(N),$$

where $\psi: \mathbb{N} \rightarrow [0, 1]$ is some weakly decreasing function. *Is convergence of the series $\sum \psi(N)/N$ necessary and sufficient for \mathcal{A} to have metric Poissonian pair correlations?*

The answer is no, as we will soon see, but there was good reason to believe otherwise. Previously, Walker had proved that the set of primes does not have MPPC [16]. When \mathcal{A} is the set of primes, one has $E(A_N) \asymp N^3(\log N)^{-1}$. Moreover, Bloom *et al.* constructed sets with

$$E(A_N) \asymp \frac{N^3}{\log N \log \log N}$$

which also do not have MPPC [4]. (And later, Lachmann and Technau constructed sets with

$$E(A_N) \asymp \frac{N^3}{\log N \log \log N \dots \underbrace{\log \dots \log N}_n} \tag{1.1}$$

which do not have MPPC [8].) Furthermore, one can construct for any $\varepsilon > 0$ a set \mathcal{A} having

$$E(A_N) \asymp \frac{N^3}{\log N (\log \log N)^{1+\varepsilon}},$$

such that \mathcal{A} does have MPPC [2]. Finally, since the methods of proof in some of the above results involved Khintchine’s theorem on Diophantine approximation in a crucial way, it was natural to suspect that the condition in Khintchine’s theorem—*i.e.* convergence of the sum $\sum \psi(n)/n$ —is what should define the necessary and sufficient threshold in the Question 1.2.

Regarding necessity, Aistleitner, Lachmann and Technau showed that for every $\varepsilon > 0$ there exists \mathcal{A} having MPPC such that

$$E(A_N) \gg \frac{N^3}{(\log N)^{\frac{3}{4}+\varepsilon}}$$

thereby answering that part of Question 1.2 in the negative, and putting an end to the idea that there is a threshold at all [2]—at least a *two-way* threshold. Sufficiency remains open. The main results of [4] were in support of it. For example, Bloom *et al.* proved the following.

THEOREM 1.3 ([4, theorem 1.4]). *If there is some constant $\xi > 0$ for which*

$$E(A_N) \ll \frac{N^3}{(\log N)^{2+\xi}} \quad \text{and} \quad \delta(A_N) \gg \frac{1}{(\log N)^{2+2\xi}},$$

where $\delta(A_N)$ denotes density, then \mathcal{A} has MPPC.

Later, Bloom and Walker added more evidence for the sufficiency part of Question 1.2 in the form of the following result.

THEOREM 1.4 ([5, theorem 6]). *There exists an absolute positive constant C such that for any $\mathcal{A} \subset \mathbb{N}$,*

$$E(A_N) \ll \frac{N^3}{(\log N)^C} \tag{1.2}$$

implies that \mathcal{A} has MPPC.

Of course, Question 1.2 supposes that that constant can be any $C > 1$. Indeed, for a version of the problem on the d -dimensional torus (where $d \geq 2$), Hinrichs *et al.* have shown that if there exists some $C > 1$ for which (1.2) holds, then \mathcal{A} has “MPPC in dimension d ” [7].

No further progress has been made on the sufficiency part of Question 1.2. One of the results of this note—Theorem 2.2—establishes an inhomogeneous analogue of the sufficiency part of the question. We introduce the inhomogeneous problem in the next section.

2. *Doubly metric Poissonian pair correlations*

As we have mentioned, an independent, identically distributed sequence $\mathbf{x} = (x_n)_n$ of random variables having the uniform distribution on \mathbb{T} will almost surely have Poissonian pair correlations. In fact, it is not hard to show that for any fixed $\gamma \in \mathbb{T}$ such a random sequence $\mathbf{x} \subset \mathbb{T}$ almost surely satisfies

$$(\forall s > 0) \quad \lim_{N \rightarrow \infty} F(\gamma, s, N, \mathbf{x}) = 2s. \tag{2.1}$$

When (2.1) holds for a sequence $\mathbf{x} \in \mathbb{T}$, we will say it has *Poissonian pair correlations with inhomogeneous parameter $\gamma \in \mathbb{T}$ (or γ -PPC)*. Notice then that, by Fubini’s theorem, the randomly generated sequence \mathbf{x} will almost surely have γ -PPC for a full-measure set of $\gamma \in \mathbb{T}$. We will present a proof of the following proposition that serves as template for some of the other results.

PROPOSITION 2.1 *Let γ be chosen randomly and uniformly from \mathbb{T} and $\mathbf{x} = (x_n)_{n=1}^\infty$ be a randomly, uniformly, and independently chosen sequence of points on \mathbb{T} . Then, almost surely, \mathbf{x} has γ -PPC. (That is, (2.1) almost surely holds.)*

With this proposition mind, we propose to study a doubly metric inhomogeneous variant of the notion of metric Poissonian pair correlations. For an infinite subset $\mathcal{A} \subset \mathbb{N}$ and an increasing integer sequence $\{N_t\}$, we will say that \mathcal{A} has *doubly metric Poissonian pair correlations (DMPPC) along the subsequence $\{N_t\} \subset \mathbb{N}$* if for almost all pairs (α, γ) , we have

$$(\forall s > 0) \quad \lim_{t \rightarrow \infty} F(\alpha, \gamma, s, N_t, \mathcal{A}) = 2s. \tag{2.2}$$

If this holds with $\{N_t\} = \mathbb{N}$, then we will just say that \mathcal{A} has DMPPC.

The double metricity of the inhomogeneous set up makes certain calculations easier than in the homogeneous setting. As a result, we are able to confirm the sufficiency part of Question 1.2 for DMPPC.

THEOREM 2.2 *If $\mathcal{A} \subset \mathbb{N}$ is an infinite set such that*

$$E(A_N) \ll N^3 \psi(N),$$

where $\psi: \mathbb{N} \rightarrow [0, 1]$ is a weakly decreasing function such that the series $\sum \psi(N)/N$ converges, then \mathcal{A} has doubly metric Poissonian pair correlations. (That is, (2.2) holds for almost all pairs $(\alpha, \gamma) \in \mathbb{T}^2$.)

Remark. In particular, if there is some $C > 1$ for which (1.2) holds, then \mathcal{A} has DMPPC (a result which in the homogenous setting is only known for higher dimensions).

For functions on the divergence side of Question 1.2, we construct sets $\mathcal{A} \subset \mathbb{N}$ that do not have DMPPC.

THEOREM 2.3 *Suppose $\psi: \mathbb{N} \rightarrow [0, 1]$ is a weakly decreasing function such that there exists $\delta > 0$ for which $N^{3-\delta}\psi(N)$ is increasing, and such that $\sum \psi(N)/N$ diverges. Then there exists an infinite set $\mathcal{A} \subset \mathbb{N}$ such that*

$$E(A_N) \asymp N^3 \psi(N)$$

and such that \mathcal{A} does not have doubly metric Poissonian pair correlations.

Remark. The assumption that $N^{3-\delta}\psi(N)$ is increasing for some $\delta > 0$ is only used to ensure that the sets we construct actually satisfy $E(A_N) \asymp N^3 \psi(N)$. (Notice that since $E(A_N)$ increases to infinity, it is natural that there ought to be extra requirements on ψ besides just divergence of $\sum \psi(N)/N$.) Theorem 2.3 implies, in particular, that there are sets $\mathcal{A} \subset \mathbb{N}$ satisfying (1.1) that do not have DMPPC.

We leave open the full necessity part of Question 1.2 for doubly metric Poissonian pair correlations. It is not entirely clear what the answer should be. The construction of Aistleitner–Lachmann–Technau [2] depends crucially on aspects of homogeneous Diophantine approximation (like continued fractions), so it is not immediately obvious whether a similar strategy would work for the doubly metric problem.

Some of the challenges of establishing DMPPC vanish if we allow ourselves to consider the limit in (2.2) along subsequences $\{N_t\} \subseteq \mathbb{N}$. For DMPPC along sequences, there is the following.

THEOREM 2.4 *Let $\mathcal{A} \subset \mathbb{N}$ be an infinite set. For any subsequence $\{N_t\} \subset \mathbb{N}$ such that*

$$\lim_{t \rightarrow \infty} N_t^{-3} E(A_{N_t}) = 0,$$

there exists a subsequence of $\{N_t\}$ along which \mathcal{A} has doubly metric Poissonian pair correlations.

In particular, Theorem 2.4 implies that if $E(A_N) = o(N^3)$, then any integer sequence has a subsequence along which \mathcal{A} has DMPPC. For example, since we have $E(A_N) \asymp N^3 (\log N)^{-1}$ when \mathcal{A} is the set of prime numbers, Theorem 2.4 leads immediately to the following.

COROLLARY 2.5 *Every increasing integer sequence has a subsequence along which the primes have doubly metric Poissonian pair correlations.*

Combining Theorem 2.4 with Bourgain’s argument in [3] leads to the following necessary and sufficient condition for the existence of a sequence along which \mathcal{A} has DMPPC.

THEOREM 2.6 *For an infinite set $\mathcal{A} \subset \mathbb{N}$, there exists an integer sequence along which \mathcal{A} has doubly metric Poissonian pair correlations if and only if $\liminf_{N \rightarrow \infty} N^{-3} E(A_N) = 0$.*

It would be interesting to know whether Theorems 2.4 and 2.6 also hold in the original homogeneous setting. For example, recall that Walker proved that the primes do not have MPPC [16]. Might it be the case that every increasing integer sequence has a subsequence along which (2.2) holds for almost every $\alpha \in \mathbb{T}$ and with $\gamma = 0$? That is, does Corollary 2.5

hold for MPPC? Or should we take Corollary 2.5 as evidence that the primes actually have DMPPC?

Finally, It is known that having Poissonian pair correlations is a stronger condition than being equidistributed [1, 6, 7, 13]. We show that the same is true inhomogeneously.

THEOREM 2.7 *If \mathbf{x} is a sequence of points in \mathbb{T} and there exists some γ for which \mathbf{x} has Poissonian pair correlations with inhomogeneous parameter γ , then \mathbf{x} is equidistributed in \mathbb{T} .*

Before moving on to the proofs, we mention a number of questions and speculations that are not pursued here.

What is the relationship between MPPC and DMPPC? Does one imply the other? From the point of view of our results, the most optimistic hope would be for DMPPC to imply MPPC, because then Theorem 2.2 would imply a positive answer to the sufficiency part of Question 1.2 for MPPC.

Another question which has been asked in the homogeneous setting and can as easily be asked in the doubly metric case is that of a zero-one law: If $\mathcal{A} \subset \mathbb{N}$ does not have DMPPC, is it the case that (2.2) fails for almost every pair (α, γ) ? For MPPC, the question is still open, with some progress in the work of Lachmann–Technau [8] and Larcher–Stockinger [9–11]. In fact, Larcher and Stockinger have conjectured that even more is true; namely, when $\mathcal{A} \subset \mathbb{N}$ does not have MPPC, then there is no $(\alpha, 0)$ for which (2.2) holds. The doubly metric version of this conjecture would immediately show that MPPC implies DMPPC. However, that doubly metric statement seems unlikely, and can perhaps be disproved by examining the examples in [2] and showing that they do not have DMPPC. A more likely speculation would be that for every fixed γ , if (2.2) does not hold for almost every α , then it holds for no α .

3. Proof of Proposition 2.1

First, we prove Proposition 2.1. The result itself is not surprising, but we include it anyway because it motivates the definition of DMPPC and because its proof is a template for the later proofs, particularly Theorem 2.2.

Note that a homogeneous version of this argument would give the corresponding result for (homogeneous) Poissonian pair correlations—originally established in [1, 12].

Proof of Proposition 2.1. We will show that for any fixed $s > 0$, we almost surely get

$$\lim_{N \rightarrow \infty} \frac{1}{N} \# \left\{ 1 \leq i \neq j \leq N : \|x_i - x_j - \gamma\| \leq \frac{s}{N} \right\} = 2s.$$

Since $s > 0$ is arbitrary, this must almost surely hold simultaneously for a countable dense set of values for $s > 0$. An approximation argument will do the rest.

Let $s > 0$. For each $i \neq j$ and N , let $\mathbf{1}_{i,j,s/N}$ denote the indicator random variable for

$$\left\{ (\gamma, \mathbf{x}) : \|x_i - x_j - \gamma\| \leq \frac{s}{N} \right\}.$$

Then we can see the pair correlation function as the random variable

$$F(s, N) = \frac{1}{N} \sum_{1 \leq i \neq j \leq N} \mathbf{1}_{i,j,s/N}. \tag{3.1}$$

Our goal is to show that $\mathbb{P}(F(s, N) \rightarrow 2s) = 1$.

We now show that if $(i, j) \neq (k, \ell)$ then $\mathbf{1}_{i,j,\varepsilon}$ and $\mathbf{1}_{k,\ell,\varepsilon'}$ are uncorrelated (hence independent since they are indicators). Indeed,

$$\begin{aligned} & \mathbb{E}(\mathbf{1}_{i,j,\varepsilon}\mathbf{1}_{k,\ell,\varepsilon'}) \\ &= \int_{\mathbb{T}} \int_{\mathbb{T}^4} \left(\sum_{n \in \mathbb{Z}} c(n)e(n(x_i - x_j - \gamma)) \right) \left(\sum_{n \in \mathbb{Z}} c'(n)e(n(x_k - x_\ell - \gamma)) \right) d(x_i, x_j, x_k, x_\ell) d\gamma, \end{aligned}$$

where $c(n)$ and $c'(n)$ are the Fourier coefficients of $\mathbf{1}_{[-\varepsilon,\varepsilon]+\mathbb{Z}}$ and $\mathbf{1}_{[-\varepsilon',\varepsilon']+\mathbb{Z}}$, respectively. Continuing,

$$= \sum_{m,n \in \mathbb{Z}} c(m)c'(n) \int_{\mathbb{T}} \int_{\mathbb{T}^4} e(m(x_i - x_j - \gamma) + n(x_k - x_\ell - \gamma)) d(x_i, x_j, x_k, x_\ell) d\gamma.$$

The integral over γ separates, and is only nonzero when $m = -n$. So we have

$$\mathbb{E}(\mathbf{1}_{i,j,\varepsilon}\mathbf{1}_{k,\ell,\varepsilon'}) = \sum_{n \in \mathbb{Z}} c(n)c'(-n) \int_{\mathbb{T}^4} e(n(x_i - x_j - x_k + x_\ell)) d(x_i, x_j, x_k, x_\ell).$$

Since $(i, j) \neq (k, \ell)$, the integral is 0 unless $n = 0$, so we are left with

$$\mathbb{E}(\mathbf{1}_{i,j,\varepsilon}\mathbf{1}_{k,\ell,\varepsilon'}) = c(0)c'(0) = \mathbb{E}(\mathbf{1}_{i,j,\varepsilon}) \mathbb{E}(\mathbf{1}_{k,\ell,\varepsilon'}),$$

which is what we claimed.

Notice that $\mathbb{E}(\mathbf{1}_{i,j,s/N}) = 2s/N$. Therefore, by (3-1), we have

$$\mathbb{E}(F(s, N)) = \frac{1}{N} \sum_{1 \leq i \neq j \leq N} \frac{2s}{N} = \left(1 - \frac{1}{N}\right) 2s$$

and

$$\begin{aligned} \sigma^2(F(s, N)) &= \frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} \sigma^2(\mathbf{1}_{i,j,s/N}) \\ &= \frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} \left[\frac{2s}{N} - \left(\frac{2s}{N}\right)^2 \right] \ll \frac{s}{N}. \end{aligned} \tag{3-2}$$

In the rest of the proof we will use these last findings to show that the probability that $F(s, N)$ is far from its expected value is small, and that the probability that that happens infinitely often is 0.

Let $\{s_M\}_{M=1}^\infty$ be any sequence converging to s . Notice that, by (3-2), we can write

$$\sigma^2(F(s_M, M^2)) \ll M^{-2}$$

for all sufficiently large M , with an implied constant which may depend on s , but is independent of the sequence $\{s_M\}$, since its terms are eventually bounded uniformly away from ∞ . Then, by Chebyshev's inequality,

$$\mathbb{P} \left[\left| F(s_M, M^2) - \left(1 - \frac{1}{M^2}\right) 2s_M \right| > \frac{1}{M^{1/4}} \right] < \frac{\sigma^2(F(s_M, M^2))}{1/M^{1/2}} \ll \frac{1}{M^{3/2}}$$

holds for all large $M \in \mathbb{N}$. Since the sum $\sum M^{-3/2}$ converges, the Borel–Cantelli Lemma says that we almost surely have

$$\left| F(\gamma, s_M, M^2, \mathbf{x}) - \left(1 - \frac{1}{M^2}\right) 2s_M \right| \leq \frac{1}{M^{1/4}}$$

for all sufficiently large M , which implies that we almost surely have $F(\gamma, s_M, M^2, \mathbf{x}) \rightarrow 2s$ as $M \rightarrow \infty$. Now, notice that for any integer N such that $M^2 \leq N \leq (M + 1)^2$, we have

$$\begin{aligned} \frac{M^2}{(M + 1)^2} F\left(\gamma, s \frac{M^2}{(M + 1)^2}, M^2, \mathbf{x}\right) &\leq F(\gamma, s, N, \mathbf{x}) \\ &\leq \frac{(M + 1)^2}{M^2} F\left(\gamma, \frac{(M + 1)^2}{M^2} s, (M + 1)^2, \mathbf{x}\right). \end{aligned}$$

The left-most and right-most members almost surely converge to $2s$ as M increases, therefore, almost surely, $F(\gamma, s, N, \mathbf{x}) \rightarrow 2s$, as we wanted.

4. Proof of Theorem 2.2

For $\mathcal{A} \subset \mathbb{N}$, $s > 0$, $N \in \mathbb{N}$, let us regard $F(s, N, \mathcal{A})$ as a measurable function on the space \mathbb{T}^2 equipped with Lebesgue measure. Specifically, it is the function

$$F(s, N, \mathcal{A}) = \frac{1}{N} \sum_{\substack{(a,b) \in A^2 \\ a \neq b}} \mathbf{1}_{(a-b),s/N},$$

where, for $d \in \mathbb{Z}$ and $\varepsilon > 0$, we use $\mathbf{1}_{d,\varepsilon}$ to denote the indicator of the set

$$\left\{ (\alpha, \gamma) \in \mathbb{T}^2 : \|d\alpha - \gamma\| \leq \varepsilon \right\}.$$

Given $d \in \mathbb{Z}$, let

$$r_A(d) = \# \left\{ (a, b) \in A^2 : a - b = d \right\}$$

be the number of ways to represent d as a difference of two elements of A . In particular, $r(d)$ is non-zero only if $d \in A - A$. It is a simple exercise to verify the following:

$$\sum_{d \in \mathbb{Z}} r_A(d) = N^2 \tag{4.1}$$

$$\sum_{d \in \mathbb{Z}} r_A(d)^2 = E(A) \tag{4.2}$$

$$F(s, N, \mathcal{A}) = \frac{1}{N} \sum_{d \in \mathbb{Z} \setminus \{0\}} r_A(d) \mathbf{1}_{d,s/N}. \tag{4.3}$$

The next lemma shows that the functions $\mathbf{1}_{d,s/N}$ are pairwise independent as random variables on \mathbb{T}^2 .

LEMMA 4.1 *For any sequence $(\varepsilon_d)_{d \in \mathbb{Z}}$ of positive reals, the associated random variables $(\mathbf{1}_{d,\varepsilon_d})_{d \in \mathbb{Z}}$ are pairwise independent.*

Proof. Let $\mathbf{1}_{d_1, \varepsilon_1}, \mathbf{1}_{d_2, \varepsilon_2}$ be any pair of distinct random variables from the sequence. We must show that

$$\int_{\mathbb{T}^2} \mathbf{1}_{d_1, \varepsilon_1} \mathbf{1}_{d_2, \varepsilon_2} = \int_{\mathbb{T}^2} \mathbf{1}_{d_1, \varepsilon_1} \int_{\mathbb{T}^2} \mathbf{1}_{d_2, \varepsilon_2}. \tag{4.4}$$

We rewrite the left-hand side as

$$\begin{aligned} \int_{\mathbb{T}^2} \mathbf{1}_{d_1, \varepsilon_1} \mathbf{1}_{d_2, \varepsilon_2} &= \int_{\mathbb{T}^2} \mathbf{1}_{(-\varepsilon_1, \varepsilon_1) + \mathbb{Z}}(d\alpha - \gamma) \mathbf{1}_{(-\varepsilon_2, \varepsilon_2) + \mathbb{Z}}(d\alpha - \gamma) d\alpha d\gamma \\ &= \int_{\mathbb{T}^2} \left(\sum_{n_1 \in \mathbb{Z}} c_1(n_1) e(n_1(d_1\alpha - \gamma)) \right) \left(\sum_{n_2 \in \mathbb{Z}} c_2(n_2) e(n_2(d_2\alpha - \gamma)) \right) d\alpha d\gamma, \end{aligned} \tag{4.5}$$

where $(c_i(n))_{n \in \mathbb{Z}}$ are the Fourier coefficients of $\mathbf{1}_{(-\varepsilon_i, \varepsilon_i) + \mathbb{Z}}$. We may rewrite (4.5) as

$$\sum_{(n_1, n_2) \in \mathbb{Z}^2} c_1(n_1) c_2(n_2) \int_{\mathbb{T}} e((d_1 n_1 + d_2 n_2)\alpha) d\alpha \underbrace{\int_{\mathbb{T}} e(-(n_1 + n_2)\gamma) d\gamma}_1.$$

The indicated integral is only nonzero if $n_1 + n_2 = 0$, so the expression becomes

$$\sum_{n \in \mathbb{Z}} c_1(n) c_2(-n) \underbrace{\int_{\mathbb{T}} e(n(d_1 - d_2)\alpha) d\alpha}_1.$$

Since $d_1 \neq d_2$, the newly indicated integral is nonzero only if $n = 0$, in which case it is 1, so the expression becomes $c_1(0)c_2(0)$, which is exactly the right-hand side of (4.4).

LEMMA 4.2 *For any infinite set $\mathcal{A} \subset \mathbb{N}$, and for every $s > 0$ and sufficiently large N , we have*

$$\int_{\mathbb{T}^2} F(s, N, \mathcal{A}) = \frac{2(N-1)}{N} s.$$

Proof. As we have seen in (4.3), $F(s, N, \mathcal{A})$ is a linear combination of random variables of the form $\mathbf{1}_{d, s/N}$ where only d varies. Notice that

$$\int_{\mathbb{T}^2} \mathbf{1}_{d, s/N} = \frac{2s}{N}$$

as long as $(2s)/N \leq 1$. Therefore, by (4.3), we have

$$\int F(s, N, \mathcal{A}) = \frac{2s}{N^2} \sum_{d \in \mathbb{Z} \setminus \{0\}} r_{\mathcal{A}}(d) \stackrel{(4.1)}{=} \frac{2s}{N^2} (N^2 - N) = \frac{2(N-1)}{N} s,$$

as claimed.

LEMMA 4.3 *For any infinite subset $\mathcal{A} \subset \mathbb{N}$, $N \in \mathbb{N}$, and $s > 0$, we have*

$$\sigma^2(F(s, N, \mathcal{A})) \ll E(\mathcal{A}) N^{-3} s,$$

where σ^2 denotes variance.

Proof. We have found in Lemma 4.1 that the random variables $\mathbf{1}_{d,s/N}$ are pairwise independent. Therefore,

$$\sigma^2(F(s, N, \mathcal{A})) = \frac{1}{N^2} \sum_{d \in \mathbb{Z} \setminus \{0\}} r_A(d)^2 \sigma^2(\mathbf{1}_{d,s/N}).$$

(The reader should keep in mind that this is really a finite sum, since $r(d)$ is nonzero for only finitely many d .) Now,

$$\sigma^2(\mathbf{1}_{d,s/N}) = \int \mathbf{1}_{d,s/N}^2 - \left(\int \mathbf{1}_{d,s/N} \right)^2 = \begin{cases} \frac{2s}{N} - \left(\frac{2s}{N} \right)^2 & \text{if } N \geq 2s \\ 0 & \text{if } N \leq 2s. \end{cases}$$

Combining, we find

$$\sigma^2(F(s, N, \mathcal{A})) \leq \frac{2s}{N^3} \sum_{d \in \mathbb{Z}} r(d)^2 \stackrel{(4.1)}{=} 2E(A)N^{-3}s,$$

which proves the lemma.

Now we can state the proof of Theorem 2.2.

Proof of Theorem 2.2. Let $\varepsilon, s > 0$. We have $E(A_N) \leq N^3 \psi(N)$ where $\psi: \mathbb{N} \rightarrow [0, 1]$ is a weakly decreasing function such that $\sum \psi(N)/N$ converges. Therefore, for any fixed real number $k > 1$, $\sum \psi(\lfloor k^t \rfloor)$ converges. In particular, by Lemma 4.3, $\sum \sigma^2(F(s, \lfloor k^t \rfloor, \mathcal{A}))$ converges. Meanwhile, Chebyshev’s inequality says that the measure of the set of $(\alpha, \gamma) \in \mathbb{T}$ for which

$$\left| F(\alpha, \gamma, s, \lfloor k^t \rfloor, \mathcal{A}) - 2s \left(\frac{\lfloor k^t \rfloor - 1}{\lfloor k^t \rfloor} \right) \right| \geq \varepsilon$$

is no more than $\varepsilon^{-2} \sigma^2(F(s, \lfloor k^t \rfloor, \mathcal{A}))$. Since $\sum \varepsilon^{-2} \sigma^2(F(s, \lfloor k^t \rfloor, \mathcal{A}))$ converges, the Borel–Cantelli lemma tells us that for almost every (α, γ) , there are at most finitely many t for which

$$\left| F(\alpha, \gamma, s, \lfloor k^t \rfloor, \mathcal{A}) - 2s \left(\frac{\lfloor k^t \rfloor - 1}{\lfloor k^t \rfloor} \right) \right| < \varepsilon$$

does not hold. Decreasing ε to 0 along a positive real sequence, we conclude that for almost every (α, γ) ,

$$F(\alpha, \gamma, s, \lfloor k^t \rfloor, \mathcal{A}) \longrightarrow 2s$$

as $t \rightarrow \infty$.

By repeating the argument of the previous paragraph for every rational multiple of s , we may say that for almost every (α, γ) , the following holds:

$$(\forall r \in \mathbb{Q}_{>0}) \quad F(\alpha, \gamma, rs, \lfloor k^t \rfloor, \mathcal{A}) \longrightarrow 2rs \quad (t \rightarrow \infty). \tag{4.6}$$

A review of the definitions reveals that if $k^t < N \leq k^{t+1}$, then

$$\frac{\lfloor k^t \rfloor}{N} F(\alpha, \gamma, s \frac{\lfloor k^t \rfloor}{N}, \lfloor k^t \rfloor) \leq F(\alpha, \gamma, s, N) \leq \frac{\lfloor k^{t+1} \rfloor}{N} F\left(\alpha, \gamma, s \frac{\lfloor k^{t+1} \rfloor}{N}, \lfloor k^{t+1} \rfloor\right).$$

From this, we find that

$$\frac{1}{k} \left(1 - \frac{1}{k^t}\right) F\left(\alpha, \gamma, s/k \left(1 - \frac{1}{k^t}\right), \lfloor k^t \rfloor\right) \leq F(\alpha, \gamma, s, N) \leq kF\left(\alpha, \gamma, sk, \lfloor k^{t+1} \rfloor\right).$$

Now, let $\delta > 0$ be a rational number. For t sufficiently large, we have

$$1 - \delta < \left(1 - \frac{1}{k^t}\right) < 1,$$

hence

$$\frac{1}{k}(1 - \delta)F(\alpha, \gamma, (s/k)(1 - \delta), \lfloor k^t \rfloor) \leq F(\alpha, \gamma, s, N) \leq kF(\alpha, \gamma, sk, \lfloor k^{t+1} \rfloor)$$

holds if N is large enough. Now if $k > 1$ is rational, we may apply (4.6) to show that for almost every (α, γ) ,

$$\frac{2s}{k^2}(1 - \delta)^3 \leq F(\alpha, \gamma, s, N) \leq 2sk^2(1 + \delta)$$

for all sufficiently large N . By taking $k \downarrow 1$ and $\delta \downarrow 0$ along sequences of rational numbers, and intersecting the corresponding full-measure sets of (α, γ) , we see that for almost all (α, γ) ,

$$F(\alpha, \gamma, s, N) \longrightarrow 2s, \quad N \rightarrow \infty. \tag{4.7}$$

Finally, by intersecting countably many full-measure subsets of \mathbb{T}^2 , we can say that for almost all (α, γ) , and all rational $s > 0$, the limit (4.7) holds. This is enough to conclude that it holds for all s .

5. Proof of Theorem 2.3

The proof of Theorem 2.3 is based on Bourgain’s construction in [3]. We require a couple of lemmas. The first lemma is a simple consequence of the fact that circle expanding maps are mixing. It will allow us to define a sequence $(U_N)_N$ of sets which are quasi-independent, meaning that there is some constant $C > 1$ such that $Leb(U_M \cap U_N) \leq C Leb(U_M)Leb(U_N)$ whenever $M \neq N$.

LEMMA 5.1 For any sequence of measurable sets $Q_N \in \mathbb{T}$ it is possible to choose a sequence of positive integers Δ_N so that the sets

$$R_N = \{\alpha \in \mathbb{T} : \Delta_N \alpha \in Q_N\}$$

are pairwise quasi-independent.

Proof. Let $m \geq 2$ be an integer. Recall that the map $f_m: \mathbb{T} \rightarrow \mathbb{T}$ defined by $f_m(\alpha) = m\alpha \pmod{1}$ is measure-preserving and mixing, meaning that for any measurable sets $S, T \subset \mathbb{T}$ we have that $Leb(f_m^{-1}(S)) = Leb(S)$ and also that

$$\lim_{k \rightarrow \infty} Leb(f_m^{-k}(S) \cap T) = Leb(S)Leb(T).$$

Notice that $R_N = f_{\Delta_N}^{-1}(Q_N)$.

We may put $\Delta_1 = 1$, and for each N , choose Δ_N to be a large enough power m^{kN} that

$$\text{Leb}(f_{\Delta_N}^{-1}(Q_N) \cap R_M) = \text{Leb}(f_m^{-kN}(Q_N) \cap R_M) \leq 2\text{Leb}(Q_N)\text{Leb}(R_M) = 2\text{Leb}(R_N)\text{Leb}(R_M)$$

holds for all $M = 1, \dots, N - 1$.

The next lemma is a measure estimate on the set of (α, γ) which simultaneously satisfy an inhomogeneous approximation condition and a homogeneous approximation condition on α .

LEMMA 5.2 For any $M \in \mathbb{N}$ and $0 \leq \sigma, \tau \leq 1/2$, let

$$S = \left\{ (\alpha, \gamma) \in \mathbb{T}^2 : \left| \alpha - \frac{a}{d} \right| \leq \frac{\sigma}{M^2} \text{ for some } d \in [M] \text{ and } (a, d) = 1 \right\}$$

and

$$T = \left\{ (\alpha, \gamma) \in \mathbb{T}^2 : \|d\alpha - \gamma\| \leq \frac{\tau}{M} \text{ for some } d \in [M] \right\}.$$

Then $\text{Leb}(S \cap T) \gg \sigma\tau$.

Proof. Notice that S consists of disjoint vertical strips in \mathbb{T}^2 over the Farey fractions

$$\mathcal{F}_M := \left\{ \frac{a}{d} \in [0, 1] : (a, d) = 1 \right\}.$$

The strip $S_{a/d}$ over $a/d \in \mathcal{F}$ has width $2\sigma/M^2$. Meanwhile, T is a union of strips which wind around the torus, of slopes $1, \dots, M$ and with vertical cross-sections of length $2\tau/M$. Specifically, they are the supports of $\mathbf{1}_{1,\tau}, \dots, \mathbf{1}_{M,\tau}$. The condition $0 < \sigma \leq 1/2$ guarantees that the indicators $\mathbf{1}_{1,\tau}, \dots, \mathbf{1}_{d,\tau}$, when restricted to the strip $S_{a/d}$, are mutually singular. They are supported on non-overlapping parallelograms contained in $S_{a/d}$, each of which has area $(2\sigma/M^2)(2\tau/M)$. Therefore, $\text{Leb}(S_{a/d} \cap T) \geq d4\sigma\tau/M^3$. Summing over $a/d \in \mathcal{F}_M$, we have

$$\text{Leb}(S \cap T) \geq \sum_{d=1}^M \frac{4\sigma\tau}{M^3} d\varphi(d), \tag{5.1}$$

where $\varphi(d)$ denotes the Euler totient function. The growth properties of φ guarantee that $\sum_{d=1}^M d\varphi(d) \gg M^3$, and combining this with (5.1) proves the lemma.

We now state the following.

Proof of Theorem 2.3. We adapt Bourgain’s arguments from [3, appendix]. First, it is possible to modify them to show that if $\psi(N) \neq o(1)$, then no $\mathcal{A} \subset \mathbb{N}$ satisfying $E(A_N) \asymp N^3\psi(N)$ has DMPPC. We do this here in Theorem 6.2. Therefore, let us assume that $\psi(N) = o(1)$.

Next, notice that we may assume that $\psi(N)^{-1} \in \mathbb{N}$ for every N , for example by replacing $\psi(N)^{-1}$ with $\lfloor \psi(N)^{-1} \rfloor$.

Let $\varepsilon > 0$ be a (small) constant, which we will specify later. Per Lemma 5.1, for each $N \in \mathbb{N}$ let Δ_N be an integer large enough that the sets

$$R_N = \left\{ \alpha \in \mathbb{T} : \|d\Delta_N\alpha\| \leq \frac{\psi(N)\varepsilon}{N} \text{ for some } 0 < d \leq N\varepsilon \right\}$$

are pairwise quasi-independent. For each N , set $S_N = R_N \times \mathbb{T} \subset \mathbb{T}^2$. Let

$$T_N = \left\{ (\alpha, \gamma) \in \mathbb{T}^2 : \|d\Delta_N\alpha - \gamma\| \leq \frac{1}{8N} \text{ for some } 0 < d \leq \frac{N}{20\psi(N)} \right\}.$$

Observe that for all large N the set T_N contains the set $\Delta_N^{-1}T$, where T is from Lemma 5.2, with $\tau = \varepsilon/8$, and that S_N contains $\Delta_N^{-1}S$ with $\sigma = \psi(N)\varepsilon^2$ (the role of M is played by $N\varepsilon$). That lemma tells us then that

$$\text{Leb}(T_N \cap S_N) \gg \psi(N)\varepsilon^3. \tag{5.2}$$

Putting $U_N := T_N \cap S_N$, notice that we have for $M \neq N$

$$\begin{aligned} \text{Leb}(U_M \cap U_N) &\leq \text{Leb}(S_M \cap S_N) \\ &\leq 2\text{Leb}(S_M)\text{Leb}(S_N) \\ &\leq 2 \left(2\psi(M)\varepsilon^2\right) \left(2\psi(N)\varepsilon^2\right) \\ &\stackrel{(5.2)}{\ll} \text{Leb}(U_M)\text{Leb}(U_N). \end{aligned}$$

The implicit constant in this last expression depends on ε , but this is unimportant. What is important now is that $(U_N)_N$ is a sequence of subsets of \mathbb{T}^2 which are pairwise quasi-independent, and have the property that

$$\sum_{t=0}^{\infty} \text{Leb}(U_{2^t}) \gg \sum_{t=0}^{\infty} \psi(2^t)$$

diverges. Therefore, $U_\infty := \limsup_t U_{2^t}$ has positive measure in \mathbb{T}^2 .

We now construct an infinite set $\mathcal{A} \subset \mathbb{N}$ whose additive energy satisfies $E(A_N) \asymp N^3\psi(N)$, and such that for every $(\alpha, \gamma) \in U_\infty$, we have $\limsup_N F(\alpha, \gamma, 1, N, \mathcal{A}) = \infty$. This will prove the theorem.

The set \mathcal{A} will consist of concatenated blocks B_N of integers, each of which is a subset

$$B_N \subset \Delta_N \left[(N\psi(N)^{-1}, 2N\psi(N)^{-1}] \cap \mathbb{N} \right].$$

In view of [3, lemma 6], we may find, for each N , a block B_N with the properties:

- (i) for all $d \in \mathbb{Z} \setminus \{0\}$ we have $r_{B_N}(\Delta_N d) \leq 2N\psi(N)$;
- (ii) for all $d \in \mathbb{Z} \setminus \{0\}$ with $|d| < N/10\psi(N)$ we have $r_{B_N}(\Delta_N d) \geq N\psi(N)/2$;
- (iii) we have $N/2 \leq \#B_N \leq 2N$.

Using (4.2) and the first two properties above, we see that $E(B_N) \asymp N^3\psi(N)$.

Put $\mathcal{A} = \{B_1, B_2, B_4, \dots\}$ as the concatenation of the blocks B_{2^t} , $t \geq 0$. Suppose that

$$\sum_{k=0}^{t-1} \#B_{2^k} < N \leq \sum_{k=0}^t \#B_{2^k},$$

that is, A_N is a truncation of \mathcal{A} in the block B_{2^t} . Clearly, we have $E(A_N) \geq E(B_{2^{t-1}})$, hence

$$E(A_N) \gg \left(2^{t-1}\right)^3 \psi\left(2^{t-1}\right) \gg N^3\psi(N).$$

On the other hand, we may assume that the sequence (Δ_N) is sparse enough that

$$E(A_N) \leq \sum_{k=0}^t E(B_{2^k}),$$

that is, the only contributions to the additive energy come from four-tuples (a,b,c,d) which lie in the same block. This leads to

$$\begin{aligned} E(A_N) &\ll \sum_{k=0}^t (2^k)^3 \psi(2^k) \\ &\ll (2^t)^3 \psi(2^t) \\ &\ll N^3 \psi(N), \end{aligned} \tag{5.3}$$

where (5.3) follows from our assumption that $N^{3-\delta}\psi(N)$ is increasing.¹ Therefore, $E(A_N) \asymp N^3\psi(N)$, as needed.

To estimate the pair correlations, note that for $N = 2^t$, we have

$$\begin{aligned} F(1, \#B_1 + \#B_2 + \#B_4 + \dots + \#B_N, \mathcal{A}) &\geq \frac{1}{4N} \sum_{d \neq 0} r_{B_n}(\Delta_N d) \mathbf{1}_{\Delta_N d, 1/(4N)} \\ &\geq \frac{\psi(N)}{8} \sum_{0 < |d| \leq \frac{N}{10\psi(N)}} \mathbf{1}_{\Delta_N d, 1/(4N)} \\ &\geq \frac{\psi(N)}{8} \sum_{0 < d \leq \frac{N}{10\psi(N)}} \mathbf{1}_{\Delta_N d, 1/(4N)}. \end{aligned}$$

Notice then that for any $\alpha \in U_N$, we will have

$$F(\alpha, 1, \#B_1 + \dots + \#B_N, \mathcal{A}) \geq \frac{1}{160\epsilon}.$$

Since the set of points $(\alpha, \gamma) \in \mathbb{T}^2$ which are contained in infinitely many U_{2^t} 's has positive measure, this implies that

$$\limsup_{N \rightarrow \infty} F(\alpha, \gamma, 1, N, \mathcal{A}) \geq \frac{1}{160\epsilon}.$$

for a positive-measure set of $(\alpha, \gamma) \in \mathbb{T}^2$. If $0 < \epsilon < 1/320$ then this value exceeds 2, therefore \mathcal{A} does not have doubly metric Poissonian pair correlations and the theorem is proved.

6. Proofs of Theorems 2.4 and 2.6

Theorem 2.4 can be deduced from the proof of Theorem 2.2.

Proof of Theorem 2.4. By Lemma 4.3, we have that

$$\sigma^2(F(s, N)) \ll E(A)N^{-3}s.$$

Let $\{N_t\} \subset \mathbb{N}$ be a sequence as in the theorem statement. By passing to a subsequence, we may assume that $\{N_t\}$ increases fast enough that for any $s > 0$, the sum $\sum_t \sigma^2(F(s, N_t))$

¹ This is the only place where that assumption is used.

converges. Let $\varepsilon, s > 0$. The first paragraph of the proof of Theorem 2.2, with the sequence $\{[k^t]\}$ replaced by the sequence $\{N_t\}$, shows that for almost every (α, γ) ,

$$F(\alpha, \gamma, s, N_t, \mathcal{A}) \longrightarrow 2s$$

as $t \rightarrow \infty$. Running the argument for every rational $s > 0$ and intersecting the (countably many) full-probability subsets of phase space, we get that for almost every (α, γ) ,

$$(\forall s \in \mathbb{Q}_{>0}) \quad \lim_{t \rightarrow \infty} F(\alpha, \gamma, s, N_t, \mathcal{A}) = 2s.$$

Therefore, (2.2) holds, as needed.

Half of Theorem 2.6 is Theorem 2.4. The other half follows immediately from an inhomogeneous version of Bourgain’s proof in [3, appendix] that $E(A) = \Omega(N^3)$ precludes MPPC. For completeness, we will carry out the argument in the inhomogeneous setting, following an exposition of Walker [17].

The basis of the argument is the Balog–Szemerédi–Gowers Lemma.

LEMMA 6.1 ([15, section 2.5]). *Let $A \subset \mathbb{Z}$ be a finite set of integers. For any $c > 0$ there exist $c_1, c_2 > 0$ depending only on c such that the following holds. If $E(A) \geq c\#A^3$, then there is a subset $B \subset A$ such that $\#B \geq c_1\#A$ and $\#(B - B) \leq c_2\#A$.*

THEOREM 6.2 ([3, appendix], doubly metric inhomogeneous version). *Suppose that $\{N_t\} \subset \mathbb{N}$ is a sequence such that we have $E(A_{N_t}) \geq cN_t^3$ for some constant $c > 0$ and all large t . Then \mathcal{A} does not have doubly metric Poissonian pair correlations along $\{N_t\}$.*

Proof. For each $t \in \mathbb{N}$ large enough, let B_{N_t} be the subset of A_{N_t} which is guaranteed by Lemma 6.1.

Let $s > 0$ be a fixed real number, to be specified. Let

$$\Omega_t := \left\{ (\alpha, \gamma) \in \mathbb{T}^2 : \|\alpha n - \gamma\| \leq \frac{s}{N_t} \text{ for some } n \in B_{N_t} - B_{N_t} \right\}$$

and notice that

$$\text{Leb}(\Omega_t) \leq \frac{2s}{N_t} \#(B_{N_t} - B_{N_t}) \stackrel{\text{Lem. 6.1}}{\leq} 2sc_2.$$

Notice also that for every $(\alpha, \gamma) \in \mathbb{T}^2 \setminus \Omega_t$ we have

$$F(\alpha, \gamma, s, N_t) = \frac{1}{N_t} \sum_{(a,b) \in A_{N_t} \times A_{N_t} \setminus B_{N_t} \times B_{N_t}} \mathbf{1}_{[0,s/N_t]}(\|\alpha(a-b) - \gamma\|),$$

hence

$$\begin{aligned} \int_{\mathbb{T}^2 \setminus \Omega_t} F(\alpha, \gamma, s, N_t) \, d\alpha \, d\gamma &= \frac{1}{N_t} \sum_{(a,b) \in A_{N_t} \times A_{N_t} \setminus B_{N_t} \times B_{N_t}} \int_{\mathbb{T}^2 \setminus \Omega_t} \mathbf{1}_{[0,s/N_t]}(\|\alpha(a-b) - \gamma\|) \\ &\leq \frac{2s}{N_t^2} \#(A_{N_t} \times A_{N_t} \setminus B_{N_t} \times B_{N_t}) \\ &\stackrel{\text{Lem. 6.1}}{\leq} 2s \left(1 - c_1^2\right). \end{aligned}$$

Now, suppose that the set of $(\alpha, \gamma) \in \mathbb{T}^2 \setminus \Omega_t$ for which $F(\alpha, \gamma, s, N_t) \leq 2s(1 - c_1^2/4)$ has measure less than $c_1^2/4$. Then we would have

$$2s(1 - c_1^2) \geq \int_{\mathbb{T}^2 \setminus \Omega_t} F(\alpha, \gamma, s, N_t) \, d\alpha \, d\gamma > 2s \left(1 - \frac{c_1^2}{4}\right) \left(1 - 2sc_2 - \frac{c_1^2}{4}\right).$$

But for small enough $s > 0$, this cannot possibly hold. Let us now specify $s > 0$ to be small enough. Then there is a set $\Gamma_t \subset \mathbb{T}^2 \setminus \Omega_t$ with $Leb(\Gamma_t) \geq c_1^2/4$ and such that $F(\alpha, \gamma, s, N_t) \leq 2s(1 - c_1^2/4)$ holds for all $(\alpha, \gamma) \in \Gamma_t$. Let $\Gamma := \limsup_{t \rightarrow \infty} \Gamma_t$, and notice that we must have $Leb(\Gamma) > 0$. Since for any $(\alpha, \gamma) \in \Gamma$, we have that

$$\liminf_{t \rightarrow \infty} F(\alpha, \gamma, s, N_t) \leq 2s \left(1 - \frac{c_1^2}{4}\right) < 2s,$$

the theorem is proved.

Proof of Theorem 2.6. For the “if” direction note that if $\liminf N^{-3}E(A) = 0$, then there is a subsequence $\{N_t\} \subset \mathbb{N}$ to which Theorem 2.4 applies.

For the “only if” direction, suppose that $\liminf N^{-3}E(A) > 0$, and let $\{N_t\} \subset \mathbb{N}$ be any subsequence. Then there is some constant $c > 0$ such that $E(A_{N_t}) \geq cN_t^3$ holds for all large t , and Theorem 6.2 implies that \mathcal{A} does not have DMPPC along $\{N_t\}$.

7. Proof of Theorem 2.7

The proof of Theorem 2.7 is adapted from [7]. We will use the following simple lemma stating that if a sequence does not equidistribute in \mathbb{T} , then there are arbitrarily small intervals in \mathbb{T} which are “overrepresented” infinitely often.

LEMMA 7.1 *Suppose $(x_n)_n$ is a sequence of points in \mathbb{T} which does not equidistribute. Then there are arbitrarily small intervals $I \subset \mathbb{T}$ such that*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \# \{n \in [N] \mid x_n \in I\} > Leb(I). \tag{7.1}$$

Proof. For an interval $I \subset \mathbb{T}$ and integer $N \in \mathbb{N}$, let

$$A_N(I) := \frac{1}{N} \# \{n \in [N] \mid x_n \in I\}.$$

That (x_n) does not equidistribute implies that there are arbitrarily short intervals $I \subset \mathbb{T}$ such that $\lim_{N \rightarrow \infty} A_N(I) \neq Leb(I)$. Let I be such an interval, as small as desired. If (7.1) holds then we are done, so let us assume that $\limsup_{N \rightarrow \infty} A_N(I) \leq Leb(I)$. Therefore, it must be that $\liminf_{N \rightarrow \infty} A_N(I) < Leb(I)$, which implies that there is some $\varepsilon > 0$ and infinitely many values of $N \in \mathbb{N}$ for which $A_N(I) \leq Leb(I)(1 - \varepsilon)$.

Now, let $\bigcup_{k=0}^K I_k$ be a partition of \mathbb{T} by intervals where $I_0 = I$ and $Leb(I_k) = (1 - Leb(I))/K$ for each $k = 1, \dots, K$. Notice that for all N we have $\sum_{k=0}^K A_N(I_k) = 1$, so for the infinitely many N for which $A_N(I) \leq Leb(I)(1 - \varepsilon)$ holds we have

$$Leb(I)(1 - \varepsilon) + \sum_{k=1}^K A_N(I_k) \geq 1.$$

This implies that there must be some $\hat{k} := \hat{k}(N) \in [K]$ for which

$$A_N(I_{\hat{k}}) \geq \frac{1 - \text{Leb}(I)(1 - \varepsilon)}{K} = \text{Leb}(I_{\hat{k}}) \left(\frac{1 - \text{Leb}(I)(1 - \varepsilon)}{1 - \text{Leb}(I)} \right).$$

By the pigeonhole principle, there is some \bar{k} such that $\bar{k} = \hat{k}(N)$ for infinitely many N . For this \bar{k} , we have $\limsup_{N \rightarrow \infty} A_N(I_{\bar{k}}) > \text{Leb}(I_{\bar{k}})$.

Proof of Theorem 2.7. Assume that the sequence $\mathbf{x} = (x_n)_n$ is not equidistributed in \mathbb{T} . In view of the results of [1, 6, 7, 13], it follows that \mathbf{x} does not have γ -PPC when $\gamma = 0$. So let us fix arbitrarily $\gamma > 0$.

The fact that the sequence is not equidistributed implies, by Lemma 7.1, that there is some arc in \mathbb{T} which is “underrepresented” for infinitely many partial sequences $(x_n)_{n=1}^N$. Furthermore, that arc can be supposed to be as long as we like. In particular, since the properties of equidistribution and γ -PPC are not altered by rotating the entire set \mathcal{A} , this proof will not lose any generality if we assume that there exist $\alpha, \beta > 0$ such that $1 - \alpha < \|\gamma\|$ and such that for infinitely many $N \in \mathbb{N}$ we have

$$\frac{1}{N} \# \{1 \leq n \leq N \mid x_n \in [0, \alpha)\} \leq \beta < \alpha.$$

Let N be one such (large) integer. For $i = 0, \dots, N - 1$, let

$$X_i := \# \left\{ 1 \leq n \leq N \mid x_n \in \left[\frac{i}{N}, \frac{i+1}{N} \right) \right\}$$

so that we have

$$\sum_{i=0}^{\lfloor N\alpha \rfloor - 1} X_i \leq N\beta \quad \text{while} \quad \sum_{i=0}^{N-1} X_i = N. \tag{7.2}$$

Now notice that for $s \in \mathbb{N}$,

$$NF(\gamma, s, N, \mathbf{x}) \leq \sum_{i=0}^{N-1} \sum_{j=-s}^s X_i X_{[i+\lfloor \gamma N \rfloor + j \pmod N]}.$$

The right-hand side defines a quadratic form in the variables X_0, \dots, X_{N-1} which, subject to the constraints imposed by (7.2), reaches its maximum when

$$X_0 = X_1 = \dots = X_{\lfloor N\alpha \rfloor - 1} = \frac{N\beta}{\lfloor N\alpha \rfloor} \quad \text{and} \quad X_{\lfloor N\alpha \rfloor} = \dots = X_{N-1} = \frac{N(1 - \beta)}{N - \lfloor N\alpha \rfloor}. \tag{7.3}$$

Now with a fixed $s \in \mathbb{N}$ and N large enough, we will have $1 - \alpha < \|\gamma\| - s/N$, hence,

$$\begin{aligned} NF(\gamma, s, N, \mathbf{x}) &\leq \sum_{i=0}^{N-1} \sum_{j=-s}^s X_i X_{i+\lfloor \gamma N \rfloor + j} \\ &= \sum_{i=0}^{\lfloor N\alpha \rfloor - 1} X_i \sum_{j=-s}^s X_{i+\lfloor \gamma N \rfloor + j} + \sum_{i=\lfloor N\alpha \rfloor}^{N-1} X_i \sum_{j=-s}^s X_{i+\lfloor \gamma N \rfloor + j} \end{aligned}$$

$$\begin{aligned}
&\stackrel{(7.3)}{\leq} \left(\frac{N\beta}{\lfloor N\alpha \rfloor} \right) \left(\frac{N(1-\beta)}{N - \lfloor N\alpha \rfloor} \right) \sum_{i=\lfloor N\alpha \rfloor}^{N-1} \sum_{j=-s}^s 2 + \left(\frac{N\beta}{\lfloor N\alpha \rfloor} \right)^2 \\
&\times \left(N(2s+1) - \sum_{i=\lfloor N\alpha \rfloor}^{N-1} \sum_{j=-s}^s 2 \right) \sim (2s+1)N \underbrace{\left[\frac{\beta}{\alpha} \left(2 - \frac{\beta}{\alpha} \right) \right]}_{<1}.
\end{aligned}$$

We see that there exists $\theta < 1$ depending only on α and β with the property that for any $s \in \mathbb{N}$ and infinitely many $N \in \mathbb{N}$, we have $F(\gamma, s, N, \mathbf{x}) \leq (2s+1)\theta$. Therefore, if s is large enough, it is impossible that $F(\gamma, s, N, \mathbf{x}) \rightarrow 2s$ as $N \rightarrow \infty$, so \mathbf{x} does not have γ -PPC.

Acknowledgments. I thank Christoph Aistleitner for an illuminating correspondence, and the referee for a careful reading.

REFERENCES

- [1] C. AISTLEITNER, T. LACHMANN and F. PAUSINGER. Pair correlations and equidistribution. *J. Number Theory* **182** (2018), 206–220.
- [2] C. AISTLEITNER, T. LACHMANN and N. TECHNAU. There is no Khintchine threshold for metric pair correlations. *Mathematika* **65**(4) (2019), 929–949.
- [3] C. AISTLEITNER, G. LARCHER and M. LEWKO. Additive energy and the Hausdorff dimension of the exceptional set in metric pair correlation problems. *Israel J. Math.* **222**(1) (2017), 463–485. With an appendix by Jean Bourgain.
- [4] T. F. BLOOM, S. CHOW, A. GAFNI and A. WALKER. Additive energy and the metric Poissonian property. *Mathematika* **64**(3) (2018), 679–700.
- [5] T. F. BLOOM and A. WALKER. GCD sums and sum-product estimates. *Israel J. Math.*, **235**(1) (2020), 1–11.
- [6] S. GREPSTAD and G. LARCHER. On pair correlation and discrepancy. *Arch. Math. (Basel)* **109**(2) (2017), 143–149.
- [7] A. HINRICHS, L. KALTENBÖCK, G. LARCHER, W. STOCKINGER and M. ULLRICH. On a multi-dimensional Poissonian pair correlation concept and uniform distribution. *Monatsh. Math.* **190**(2) (2019), 333–352.
- [8] T. LACHMANN and N. TECHNAU. On exceptional sets in the metric Poissonian pair correlations problem. *Monatsh. Math.* **189**(1) (2019), 137–156.
- [9] G. LARCHER and W. STOCKINGER. *7. On Pair Correlation of Sequences* (De Gruyter, Berlin, Boston, 20 Jan. 2020), 133–146.
- [10] G. LARCHER and W. STOCKINGER. Pair correlation of sequences with maximal additive energy. *Math. Proc. Camb. Phil. Soc.* **168**(2) (2020), 287–293.
- [11] G. LARCHER and W. STOCKINGER. Some negative results related to Poissonian pair correlation problems. *Discrete Math.* **343**(2) (2020), 111656, 11.
- [12] J. MARKLOF. Distribution modulo one and Ratner’s theorem. In *Equidistribution in number theory, an introduction*, volume 237 of *NATO Sci. Ser. II Math. Phys. Chem.* (Springer, Dordrecht, 2007), 217–244.
- [13] J. MARKLOF. Pair correlation and equidistribution on manifolds. *Monatshefte für Mathematik*, **191** (2020) 279–294.
- [14] Z. RUDNICK and P. SARNAK. The pair correlation function of fractional parts of polynomials. *Comm. Math. Phys.* **194**(1) (1998), 61–70.
- [15] T. TAO and V. VU. *Additive combinatorics*. *Camb. Stud. Adv. Math.* (Cambridge University Press, Cambridge, 2006).
- [16] A. WALKER. The primes are not metric Poissonian. *Mathematika* **64**(1) (2018), 230–236.
- [17] A. WALKER. Additive combinatorics: some new techniques for pair correlation problems. Notes for a minicourse (2019).